Стохастическое программирование и его приложения

Иркутск – 2012

Институт систем энергетики им. Л.А. Мелентьева СО РАН
Институт кибернетики им. В.М. Глушкова НАН Украины
Иркутская государственная сельскохозяйственная академия

Стохастическое программирование и его приложения

Научные редакторы:
член-корреспондент НАН Украины *П.С. Кнопов*
доктор технических наук *В.И. Зоркальцев*

г. Иркутск
2012

УДК 519.856
ББК В 183.4

Стохастическое программирование и его приложения / П.С. Кнопов, В.И. Зоркальцев, Я.М. Иваньо и др. [Электронный ресурс]. – Иркутск: Институт систем энергетики им. Л.А. Мелентьева СО РАН, 2012. – 493с.; объем 700 Мб; 1 опт. компакт-диск (CD-ROM).

ISBN 978-5-93908-110-8.

В книге рассматриваются математические модели функционирования и развития сложных систем (энергетики и сельского хозяйства) в условиях действия случайных факторов, методы решения задач оптимизации в детерминированных и в вероятностных постановках.

Книга рассчитана на специалистов в области стохастической оптимизации, энергетики, экономики, преподавателей и аспирантов вузов.

Табл.: 33. Ил.: 56. Библиогр.: 339 назв.

Рецензенты:

доктор технических наук *А.Ю. Горнов*

доктор физико-математических наук *О.В. Хамисов*

кандидат технических наук *О.В. Войтов*

Утверждено к печати

Ученым советом Института систем энергетики им. Л.А. Мелентьева СО РАН

ISBN 978-5-93908-110-8

(Электронный ресурс)

© П.С. Кнопов, В.И. Зоркальцев,
Я.М. Иваньо и др., 2012

© Российская академия наук, 2012

Содержание

Введение 6

I Математическое программирование 8

1.1 Модели и методы оптимизации в энергетических исследованиях (*Воропай Н.И., Зоркальцев В.И.*) 9

1.2 Субградиентный алгоритм минимизации с преобразованием пространства $\alpha(\varepsilon)$ – алгоритм (*Журбенко Н.Г.*) . . . 36

1.3 Операторы преобразования пространства в квазиньютоновских методах и методах сопряженных градиентов (*Журбенко Н.Г.*) 52

1.4 Обоснование алгоритмов внутренних точек для задач нелинейного программирования (*Зоркальцев В.И., Пержабинский С.М.*) 59

1.5 Сведение задач двухэтапной вероятностной оптимизации с дискретным распределением случайных данных (*Кибзун А.И., Норкин В.И., Наумов А.В.*) 76

1.6 Меры риска в задачах стохастической оптимизации для получения робастных решений (*Кирилюк В.С.*) 104

1.7 Метод эмпирических средних в задачах стохастической оптимизации и оценивания (*Кнопов П.С.*) 125

1.8 Решение двухэтапных задач стохастического программирования большой размерности PNK-методом (*Кузьменко В.Н.*) 150

1.9 Ускоренные модификации субградиентного метода Поляка для овражных выпуклых функций (*Стецюк П.И.*) 160

II Математическое моделирование 185

2.1 Модели несовершенной конкуренции применительно к анализу электроэнергетического рынка Сибири (*Айзен-*

берг Н.И. , Киселёва М.А.) 186
2.2 Модели оптимизации размещения посевов сельскохозяйственных культур в условиях неполной информации (*Астафьева М.Н., Иваньо Я.М.*) 217
2.3 Задачи параметрического программирования с вероятностными и интервальными параметрами применительно к задачам оптимизации сельскохозяйственного производства (*Барсукова М.Н., Иваньо Я.М.*) 229
2.4 Оценка переменных режима ЭЭС методами вероятностного потокораспределения (*Болоев Е.В., Голуб И.И.*) 242
2.5 Модели оптимизации взаимодействия участников в агропромышленных кластерах с вероятностными параметрами (*Бузина Т.С., Иваньо Я.М.*) 265
2.6 Макромодель энергетики и экономического роста (*Горбачук В.М., Пепеляев В.А.*) 278
2.7 О создании энергетических плантаций в России и мире (*Губий Е.В*) 307
2.8 Математические модели для анализа и синтеза надежности при многоуровневом моделирования систем газоснабжения (*Дзюбина Т.В., Илькевич Н.И.*) 317
2.9 Информационно-программное обеспечение задач анализа и надёжности при многоуровневом моделировании газоснабжающих систем (*Дзюбина Т.В., Метёлкин А.М.*) 349
2.10 Имитационная модель поведения однопродуктового рынка (*Зоркальцев В.И., Мокрый И.В.*) 370
2.11 Модели анализа состояний электроэнергетических систем (*Зоркальцев В.И., Пержабинский С.М.*) 381
2.12 Декомпозиция и синтез системной надежности электроэнергетических систем (*Крупенёв Д.С.*) 397
2.13 Теория, методы и модели анализа надежности ЭЭС (*Ко-*

валёв Г.Ф.) 415

2.14 Программно-вычислительный комплекс оценки надежности электроэнергетических систем «Янтарь», алгоритмические особенности (*Лебедева Л.М.*) 437

2.15 Методы комплексного анализа и надежности теплоснабжения (*Стенников В.А., Постников И.В.*) 447

2.16 Оптимальные нормированные векторы конечного продукта и добавленной стоимости в продуктивной модели Леонтьева (*Стецюк П.И.*) 483

ВВЕДЕНИЕ

Функционирование многих технических и экономических систем осуществляется в условиях существенного воздействия случайных, заранее непредусмотренных факторов. Поэтому при математическом моделировании данных систем, в том числе в целях оптимизации их функционирования и развития, очень важно адекватно учитывать априорную неопределенность, в том числе вероятностный характер отдельных показателей и возможных ситуаций. Это требует применения специальных видов математических моделей, учитывающих действия неопределенности и случайных факторов, специальных математических методов для реализации таких моделей, в том числе методов решения задач оптимизации в детерминированных и в стохастических условиях.

В книге представлены результаты разработки, координации совместных усилий в указанном направлении трех научных коллективов: Института кибернетики им. В.М. Глушкова НАН Украины, Института систем энергетики им. Л.А. Мелентьева СО РАН, Иркутской государственной сельскохозяйственной академии. Тематика докладов охватывает ряд важных направлений теории оптимизации:

- методы решения негладких невыпуклых задач стохастической оптимизации, необходимые условия оптимальности, методы вычисления стохастических градиентов;
- методы негладкой оптимизации и их адаптация для решения блочных задач математического программирования;
- методы оптимизации и оценивания параметров стохастических систем при наличии априорной информации относительно области их значений;
- метод внутренних точек для задач математического программирования;
- методы оценивания риска в стохастических моделях, возникающих в различных областях техники, экономики и других сферах человеческой деятельности;

- модели и методы анализа и синтеза надежности функционирования систем энергетики и энергообеспечения;
- модели и методы оптимизации функционирования и развития газоснабжающих систем;
- модели оптимизации сельскохозяйственной деятельности в условиях неопределенных прогнозов природно-метеорологических условий.

В книге представлены материалы, которые обсуждались на Российско-Украинском научном семинаре «Стохастическое программирование и его приложения», прошедшего 23 – 29 июля 2012 года в городе Иркутске. Семинар осуществлялся при финансовой поддержке НАН Украины (грант № 01-01-12 (У)) и РФФИ (грант № 12-01-90550-Укр_г).

Книга подготовлена авторским коллективом в следующем составе: Айзенберг Н.И. (гл. 2.1), Астафьева М.Н. (гл. 2.2), Барсукова М.Н. (гл. 2.3), Болоев Е.В. (гл. 2.4), Бузина Т.С. (гл. 2.5), Воропай Н.И. (гл. 1.1), Голуб И.И. (гл. 2.4), Горбачук В.М. (гл. 2.6), Губий Е.В. (гл. 2.7), Дзюбина Т.В. (гл. 2.8, гл. 2.9), Журбенко Н.Г. (гл. 1.2, гл. 1.3), Зоркальцев В.И. (гл. 1.1, гл. 1.4, гл. 2.10, гл. 2.11), Иваньо Я.М. (гл. 2.2, гл. 2.3, гл. 2.5), Илькевич Н.И. (гл. 2.8), Кибзун А.И. (гл. 1.5), Кирилюк В.С. (гл. 1.6), Киселёва М.А. (гл. 2.1), Кнопов П.С. (гл. 1.7), Ковалёв Г.Ф. (гл. 2.13), Крупенёв Д.С. (гл. 2.12), Кузьменко В.Н. (гл. 1.8), Лебедева Л.М. (гл. 2.14), Метёлкин А.М. (гл. 2.9), Мокрый И.В. (гл. 2.11), Наумов А.В. (гл. 1.5), Норкин В.И. (гл. 1.5), Пержабинский С.М. (гл. 1.4, гл. 2.10), Пепеляев В.А. (гл. 2.6), Постников И.В. (гл. 2.15), Стенников В.А. (гл. 2.15), Стецюк П.И. (гл. 1.9, гл. 2.16).

Раздел I

Математическое программирование

УДК 330+519.6

МОДЕЛИ И МЕТОДЫ ОПТИМИЗАЦИИ В ЭНЕРГЕТИЧЕСКИХ ИССЛЕДОВАНИЯХ

Воропай Н. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: voropai@isem.sei.irk.ru

Зоркальцев В. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: zork@isem.sei.irk.ru

Аннотация. В статье даётся краткая характеристика некоторых моделей оптимизации функционирования и развития систем энергетики, созданных в Институте систем энергетики Сибирского отделения Российской академии наук. Рассматриваются модели оптимизации развития топливно–энергетического комплекса России, вычислительный комплекс, созданный для анализа надёжности электроэнергетических систем, модель потокораспределения в гидравлических системах. Даётся общее представление о разрабатывавшихся в Институте систем энергетики методах оптимизации, в т.ч. метод приведённого градиента, алгоритмы внутренних точек, метод модифицированной функции Лагранжа, алгоритмов отсечений.

Ключевые слова. Модели оптимизации функционирования и развития систем энергетики. Алгоритмы решения задач математического программирования.

Введение

Институт систем энергетики им. Л.А. Мелентьева (ИСЭМ СО РАН) создан в 1960 году. Он расположен в г. Иркутске, в центре Сибири. До 1997 года он назывался Сибирский энергетический институт (СЭИ). ИСЭМ является одним из 8 институтов, составляющих Иркутский научный центр Сибирского отделения Российской Академии наук.

Изначально основным направлением научных исследований Института систем энергетики являлся поиск наилучших вариантов формирования, функционирования и развития сложных энергетических объектов и систем [1]. Основным инструментом в этих исследованиях служат математические модели. Для реализации математических моделей функционирования и развития систем энергетики необходимо применение специальных методов вычислительной математики. Среди них особо важную роль играют методы

оптимизации – алгоритмы выбора наилучших вариантов по заданным критериям среди возможных. Разработка, развитие, экспериментальные и теоретические исследования методов оптимизации, их применение к моделям энергетики составляют также одно из направлений научных исследований института.

В данной статье постараемся познакомить читателей с некоторыми моделями оптимизации, разрабатывавшимися в Институте систем энергетики, а также с оригинальными методами решения задач оптимизации, создававшимися сотрудниками института и получившими широкое признание.

Статья не претендует на полное изложение всех или даже отдельных математических моделей, созданных в институте. Статья не претендует и на полное изложение всех разрабатывавшихся и использовавшихся в институте методов решения задач оптимизации. Наша цель состоит только в том, чтобы дать общее представление о исследованиях и разработках Института систем энергетики по математическому моделированию и методам оптимизации.

Сначала приведём краткую характеристику некоторых математических моделей оптимизации созданных в институте. Затем представим создававшиеся в институте методы оптимизации. Здесь будут рассмотрены четыре класса методов оптимизации. Эти методы мы персонифицируем с авторами исходных их идей. Хотя, конечно, в развитии и реализации этих методов участвовали многие сотрудники института. Необходимо отметить, что создателями рассматриваемых здесь методов оптимизации были как профессиональные математики (В.П. Булатов, И.И. Дикин), так и специалисты-энергетики (Л.А. Крумм, Ш.С. Чурквеидзе). Это отражает полезность симбиоза разных наук, который был заложен еще при основании института.

Модель выбора оптимальных долгосрочных вариантов развития топливно-энергетического комплекса страны

Одной из крупных комплексных разработок Института систем энергетики, потребовавшей привлечения специалистов разного профиля, является создание модели (в нескольких модификациях) для исследования долгосрочных, на 20 и более лет, вариантов развития топливно-энергетического комплекса страны [2, 3]. Модель предназначена для определения наиболее рациональных путей развития отраслей энергетики с учётом существующих взаимосвязей отдельных видов энергоресурсов при их производстве и их взаимозаменяемости при потреблении. Многие объекты энергетического комплекса (электростанции, нефте- и газодобывающие предприятия, нефтеперерабатывающие заводы, системы теплоснабжения городов) являются очень капиталоёмкими, требуют длительного времени для ввода и реконструкции (на предпроектные и проектные исследования, на создание строительной базы и на само строительство). При этом многие энергетические комплексы функционируют длительное время – от 20 до 50 , и более лет. Этим обусловлена необходимость заблаговременной системной проработки вопросов целесообразности сооружения в отдельных районах тех или иных мощностей по добыче, переработке, транспортировке энергоресурсов. Необходимо рассматривать долгосрочные перспективы для оценки эффективности новых мощностей.

Первые, разрабатывавшиеся с конца 60-х годов прошлого века, модели оптимизации топливно-энергетического комплекса относились к классу условно – динамических. Рассматривалось состояние отраслей энергетики в годы $t = 1,\ 2, \ldots, T$, соответствующие последним годам пятилетий. Модель имела вид задачи линейного программирования*:*

$$A^t x^t + C^t y^t = b^t, \tag{1}$$

$$N^t \geq x^t \geq 0, \tag{2}$$

$$M^t \geq y^t \geq 0, \tag{3}$$

$$(c^t, x^t) + (k^t, y^t) \to \min. \qquad (4)$$

Переменные модели составляют векторы (здесь t - рассматриваемый год):

x^t – интенсивности использования располагаемых к году t производственных мощностей;

y^t – прирост между периодами $t-1, t$ новых мощностей.

Заданными являются:

матрицы A^t и C^t – коэффициентов использования существовавших и новых технологий;

вектор заданных потребностей в различных видах энергоресурсов по разным регионам b^t;

располагаемые к году t мощности, составляющие вектор N^t;

максимально возможные вводы новых мощностей, составляющие вектор M^t.

Динамика изменения располагаемых мощностей представляется в виде следующего соотношения

$$N^{t+1} = N^t - R^t + y^t,$$

где R^t – мощности, выводимые из строя за период между годами t и $t+1$. Ограничение (2) означает, что интенсивность использования технологических способов – величина неотрицательная, и она ограничена сверху располагаемыми мощностями. Ограничения (3) означают, что в водимые мощности – неотрицательные величины, непревосходящие заданных сверху ограничений на возможности ввода новых мощностей за период [$t-1$, t].

Соотношение (1) задаёт балансовые условия на производство, транспортировку и потребление отдельных видов энергоресурсов. Матрицы A^t и C^t можно объединить в единую матрицу технологических коэффициентов [A^t, C^t]. Эта матрица имеет двойную блочную структуру – по территориальному и отраслевому принципам.

В модели рассматриваются балансы энергоресурсов применительно к каждому крупному региону. Этих регионов относительно небольшое число: от 12 до 17 в разных вариантах модели развития ТЭК России. Уместно напомнить, что Россия является самой большой в мире по площади страной. Региональные блоки связаны между собой транспортным блоком, описывающим поставки отдельных видов энергоресурсов из одного региона в другой.

В отраслевом разрезе выделяются следующие взаимосвязанные блоки.

1. **Уголь.** В этом блоке описываются процессы добычи, обогащения, перевозок между регионами (транспортом общего пользования) и энергетическое потребление разных видов углей.
2. **Природный газ.** В нём описываются процессы добычи, транспортировки по системе газопроводов и потребления природного газа.
3. **Нефтяная промышленность.** В этом блоке описываются процессы добычи и транспорта нефти, добываемой на разных месторождениях страны.
4. **Нефте- и газопереработка.** В этом блоке укрупнено описываются технологические процессы, используемые на нефте- и газоперерабатывающих заводах для производства различных видов моторного топлива, топочного мазута, химической и других видов продукции из углеводородного сырья.
5. **Электро- и теплоэнергетика.** В этом блоке описываются разные способы производства электроэнергии (в т.ч. на топливных электростанциях, на атомных электростанциях, на гидроэлектростанциях), транспорт электроэнергии по линиям электропередач между регионами. При описании процессов производства электроэнергии учитывается, что многие электростанции одновременно с производством электроэнергии вырабатывают

теплоэнергию для производственных нужд, а также для отопления и горячего водоснабжения населения.

Целевая функция (4) означает, что минимизируются затраты на функционирование и развитие топливно-энергетического комплекса. Компоненты вектора c^t соответствуют переменным текущим издержкам. Компоненты вектора k^t имеют вид:

$$k^t = s^t + Ed^t,$$

где s^t – переменные текущие издержки, d^t – инвестиции на создание единицы производственных мощностей, E – коэффициент эффективности капиталовложений (обычно в расчётах принимается равным 0,12).

Модель анализа надёжности электроэнергетических систем

В начале 70-х гг. в Сибирском энергетическом институте (ныне ИСЭМ СО РАН) под руководством академика Ю.Н. Руденко была разработана методика анализа надёжности электроэнергетических систем (ЭЭС), основанная на методе статистических испытаний (метод Монте-Карло) [4, 5].

Методика анализа надёжности состоит из трёх основных частей:

1. Вероятностный блок, где формируются случайные состояния режимов функционирования ЭЭС с использованием датчиков случайных чисел. Случайность позволяет «проигрывать» разные возможные ситуации при функционировании ЭЭС (случайнее колебания нагрузки, отказы оборудования и др.).

2. Оценка дефицита мощности для каждого выбранного случайного состояния системы. Это осуществляется на основе специальной модели оценки дефицита мощности, имитирующей функционирование электроэнергетической системы при заданных условиях. В результате определяется для каждого узла электроэнергетической системы дефицит электроэнергии (сам факт его возникновения и его объём).

3. Блок вычисления показателей надёжности (в том числе вероятности бездефицитной работы системы и энергоузлов, математического ожидания

дефицита мощности и др.), в котором обрабатывается информация, накопленная в результате многократных оценок случайных состояний ЭЭС.

В рассматриваемой методике анализа надёжности модели оптимизации используются дважды. Во-первых, вид задачи оптимизации имеет модель оценки дефицита мощности. В ней определяются объёмы генерации мощности в каждом узле и её перетоки между узлами в целях обеспечения минимального суммарного дефицита мощности в системе. В качестве модификации такой постановки рассматривается задача минимизации ущербов от дефицита мощности в энергетической системе. Для решения такой задачи требуется дополнительная информация об ущербах от дефицита мощности в узлах энергосистемы в зависимости от объёмов этого дефицита. Для реализации модели оценки дефицита мощности используются алгоритмы метода внутренних точек [6, 7].

Другим объектом оптимизации в рассматриваемой методике является задача выбора средств повышения надёжности электроснабжения. Это обеспечивается за счёт создания дополнительных мощностей по производству электроэнергии и развитию линий электропередач между узлами электроэнергетической системы.

Модели потокораспределения

Одним из приложений теории и методов оптимизации являются гидравлические цепи. Теория гидравлических цепей была создана Л.Я. Хасильевым и А.П. Меренковым [8] в Институте систем энергетики для моделирования функционирования трубопроводных систем, осуществляющих водоснабжение, теплоснабжение, газоснабжение, транспортировку нефти и нефтепродуктов. Эти модели можно представить в виде систем нелинейных уравнений, либо в виде эквивалентных задач математического программирования. В качестве исходного аналога при создании теории гидравлических цепей послужила теория электрических цепей, созданная в XIX веке. Приведём для примера модель классической задачи потокораспределения.

Система моделируемой гидравлической системы описывается ориентированным графом. Пусть m – число узлов, n – число дуг этого графа. Тогда A – матрица инциденции размера $m \times n$ с элементами: $a_{ij} = 1$, если дуга j выходит из узла i; $a_{ij} = -1$, если дуга j входит в узел i; $a_{ij} = 0$, если дуга j не инцидентна узлу i. В каждом столбце матрицы A только два ненулевых элемента: один равен *+1*, другой равен *–1*.

Заданными также являются вектор $b \in R^n$ и вектор $c \in R^n$. Компоненты вектора b характеризуют расходы среды, поставляемые в систему (если $b_i > 0$) или из системы (если $b_i < 0$) для узла $i = 1,...,m$.

Компоненты вектора c характеризуют приращение давления на отдельных дугах. Если $c_j > 0$, то происходит приращение давления (вследствие работы насосов). Если $c_j < 0$, то осуществляется дроселирование давления на дуге $j = 1,...,n$.

Переменные в модели составляют вектора $x \in R^n$, $y \in R^n$, $u \in R^n$. Компоненты векторов x и y характеризуют величины расходов и потерь давления на отдельных дугах. Компоненты вектора u характеризуют давление в отдельных узлах системы.

Классическая задача потокораспределения состоит в отыскании векторов *x, y, u*, удовлетворяющих условиям:

$$Ax = b, \tag{5}$$

$$y - A^T u = c, \tag{6}$$

$$y = f(x). \tag{7}$$

Здесь условие (5) характеризует баланс расходов транспортируемой среды в каждом узле (аналог первого закона Кирхгофа). Условие (6) характеризует баланс давления для каждой дуги. Условие (7) выражает зависимость между расходом x_j и потерей давления y_j по всем

дугам $j = 1,...,n$. Здесь f вектор – функция с компонентами f_j. Часто применяется квадратичная зависимость

$$f_j(x_j) = k_j |x_j| x_j, j = 1,...,n,$$

где k_j – заданный коэффициент. Возможны и другие зависимости. Отметим, что в электрических цепях аналогом (7) является закон Ома.

Потребуем от функций f_j выполнения следующих условий:

- являются возрастающими, $f_j(\alpha) > f_j(\beta)$, если $\alpha > \beta$;

- равны нулю в нуле, $f_j(0) = 0$;

- непрерывные.

В силу указанных свойств каждой функции f_j соответствует обратная ей функция, удовлетворяющая этим же условиям. Обратную функцию обозначим φ_j. Итак, для любого вещественного α

$$\varphi_j(f_j(\alpha)) = \alpha, j = 1,...,n.$$

Из указанных свойств функций f_j и φ_j следует существование для них первообразных:

$$F_j(\beta) = \int_0^\beta f_j(\alpha)d\alpha; \quad \Phi_j(\beta) = \int_0^\beta \varphi_j(\alpha)d\alpha.$$

Отметим, что функции F_j и Φ_j являются строго выпуклыми, имеющими абсолютный минимум в нуле равный нулю.

Введём функции от векторов $x \in R^n$, $y \in R^n$

$$F(x) = \sum_{j=1}^n F_j(x_j), \quad \Phi(y) = \sum_{j=1}^n \Phi_j(y_j).$$

Рассмотрим задачу выпуклого программирования

$$F(x) - \sum c_j x_j \to \min, \tag{8}$$

$$Ax = b. \tag{9}$$

В качестве решения этой задачи наряду с вектором x будет рассматривать вектор множителей Лагранжа u ограничений (9) и вектор y , связанный с x условием (7). Здесь $f(x)=\nabla F(x)$.

Введём двойственную к (8), (9) задачу выпуклого программирования

$$\Phi(y)-(b,u)\to\min, \tag{10}$$

$$y-A^{T}u=c. \tag{11}$$

В качестве решения этой задачи наряду с векторами y и u будет рассматривать вектор множителей Лагранжа x ограничений (11).

Справедливо следующее утверждение [9].

Теорема. *Если система уравнений* (5) *совместна, то система* (5) – (7)*, задача выпуклого программирования* (8)*,* (9) *и задача выпуклого программирования* (10*),* (11) *имеют решения. Причём решения любой из этих задач являются решением остальных двух. Среди решений данных задач векторы x, y имеют единственное значение.*

Приведённая теорема позволяет сводить исходную систему уравнений к задаче выпуклого программирования, для решения которых могут применяться любые известные методы. Отметим, что для существования решения у системы (5) достаточно, если граф, характеризуемый матрицей инциденций A, является связным и выполняется условие

$$\sum_{i=1}^{m} b_i = 0 .$$

Другие модели

Необходимо упомянуть о других моделях оптимизации, создававшихся в ИСЭМ СО РАН: оперативного управления режимами функционирования электроэнергетических систем; оптимизации среднесрочных и долгосрочных программ развития электроэнергетических систем; оптимизации графиков ремонта оборудования электроэнергетики; оптимизации параметров технологических процессов и оборудования по переработке энергоресурсов; оптимизация резервов и запасов топлива [17].

Метод приведенного градиента

Одним из пионеров в разработке специальных методов расчета допустимых и оптимальных режимов электроэнергетических систем был Л.А. Крумм, ныне академик Эстонской Академии наук. Он длительное время работал в ИСЭМ и является одним из основателей электроэнергетического направления исследований института. Особенно большой вклад он внёс в создание моделей и методов оптимизации режимов функционирования электроэнергетических систем.

Л.А. Крумм является одним из создателей класса алгоритмов оптимизации, содержащим методы приведенного градиента в современной терминологии. Его первые работы по обсуждаемым здесь методам оптимизации появились в конце 50-х годов: в 1957 г. – в трудах Таллиннского политехнического института, в 1959 г. – в Известиях СО АН СССР [10, 11].

Л.А. Крумм хорошо понимал, что разработанный им метод может использоваться не только для решения задач расчетов режимов электроэнергетических систем. Поэтому при описании метода он, как правило, использовал общую запись задачи нелинейного программирования. При этом он рассматривал особый класс задач нелинейного программирования – с двусторонними ограничениями-неравенствами на переменные. Это означает, что множество допустимых решений (удовлетворяющих ограничениям задачи) заведомо либо ограниченное, либо пустое (ограничения противоречивые).

Пусть переменные задачи составляют вектор $x \in R^n$. Заданы минимизируемая целевая функция f_0 и функции ограничений f_i, $i = 1, \ldots, m$, а также векторы $\underline{x}$, $\bar{x}$ из R^n. Причем $\underline{x}_j < \bar{x}_j$,

Рассматривается задача

$$f_0(x) \to \min, \tag{12}$$

при ограничениях

$$f_i(x) = 0,\ i = 1, \ldots, m, \tag{13}$$

$$\underline{x} \le x \le \bar{x}\,. \tag{14}$$

Считается, что функции $f_i,\ i = 1, \ldots, m,$ дифференцируемы. Следует отметить, что для моделей расчета электроэнергетических режимов функции f_i существенно нелинейные. Поэтому задача (12) – (14) относится к классу собственно нелинейных и даже невыпуклых задач оптимизации.

Центральная идея метода приведенного градиента состоит в том, чтобы представить ограничения-равенства (13) в виде выражения одних переменных через другие. Пусть вектор переменных x разбит на два вектора: вектор базисных или зависимых переменных $z \in R^m$ и вектор независимых переменных $y \in R^{n-m}$. Для определенности будем считать, что первые $n - m$ компонент вектора x составляют вектор y, а остальные m компонент – вектор z. Причем вектор z представляется в виде вектор-функции от вектора y, которая равносильна ограничениям (13). То есть ограничения (13) приобретают вид

$$x_{i+n-m} - z_i(y) = 0\,,\ i = 1, \ldots, m\,. \tag{15}$$

В результате преобразований (15) задача (11) – (14) приобретает вид

$$g(y) \equiv f_0(y, z(y)) \to \min\,, \tag{16}$$

при ограничениях

$$\underline{x}_j < y_j < \bar{x}_j,\ j = 1, \ldots, n - m\,, \tag{17}$$

$$\underline{x}_{i+n-m} \le z_i(y) \le \bar{x}_{i+n-m}\,,\ i = 1, \ldots, m\,. \tag{18}$$

Эта задача имеет меньшее число переменных, чем задача (12) – (14), и все условия заданы в виде ограничений-неравенств.

Л.А. Крумм особо исследовал случай, когда текущее приближение y^k, где k – номер итерации, является допустимым для задачи (16) – (18). Вектор s^k определяется как направление уменьшения целевой функции от точки y^k в области допустимых решений. Например, вектор s^k можно определить как результат решения задачи

$$(\nabla g(y^k),\, s^k) \to \min, \tag{19}$$

при условиях

$$\underline{x}_j < y_j^k + s_j^k < \bar{x}_j, \;\; j = 1, \ldots, n-m, \tag{20}$$

$$\underline{x}_{i+n-m} \le z_i(y^k + s^k) \le \bar{x}_{i+n-m}, \;\; i = 1, \ldots, m. \tag{21}$$

Затем вычисляем шаг λ_k с тем, чтобы вектор

$$y^{k+1} = y^k + \lambda_k s^k$$

оставался в области допустимых решений задачи (16) – (18) и при этом минимизировал значение целевой функции.

Такой алгоритм в современной терминологии называется методом условного градиента решения задачи (16) – (18). Для исходной задачи (12) – (14) использование преобразования (15) означает, что изложенный алгоритм будет методом приведенного градиента.

Для решения вспомогательной задачи (19) – (21) и ее модификаций могут использоваться разные методы. Одно время для ее решения применяли метод модифицированной функции Лагранжа. Об этом методе, разработанном Ш.С. Чурквеидзе, пойдет речь ниже. В дальнейшем стал использоваться метод внутренних точек.

Метод внутренних точек

Сотрудник Института систем энергетики И.И. Дикин являлся автором оригинального метода внутренних точек решения задач линейного, и нелинейного программирования. Уже в 70-х годах этот метод нашел применение в реализации нескольких моделей СЭИ, в т.ч. анализа надежности электроэнергетических систем, в задачах выбора оптимальных графиков ремонта оборудования [6], с 80-х годов – в моделях термодинамики.

Исходной основой И.И. Дикина в разработке его алгоритма внутренних точек послужила очень любопытная и простая идея академика Л.В. Канторовича по определению рациональной системы цен в случае

неоптимального плана (речь конкретно шла о планах использования посевных площадей, хотя предложенный Л.В. Канторовичем подход вполне годится и для других проблем, в т.ч. для планов развития и ценообразования в энергетике).

Рассмотрим взаимно двойственные задачи линейного программирования

$$c^T x \to \min,\ x \in X, \tag{22}$$

$$b^T u \to \max,\ u \in U, \tag{23}$$

где

$$X = \{x \in R^n : Ax = b,\ x \geq 0\}, \tag{24}$$

$$U = \{u \in R^m : g(u) \equiv A^T u - c \leq 0\} \tag{25}$$

– множества допустимых решений исходной (22) и двойственной (23) задач. Заданными являются векторы $c \in R^n$, $b \in R^m$, матрица A размера $m \times n$. Переменные составляют векторы $x \in R^n$, $u \in R^m$.

Пусть задача (22) моделирует процесс производства и использования некоторых видов продукции, составляющей вектор b. Вектор x состоит из показателей интенсивности технологии, A – матрица технологических коэффициентов. Задача (11) состоит в определении неотрицательных интенсивностей использования технологий, при которых обеспечивается выпуск заданного набора продукции при заданных ресурсах производства. Среди допустимых наборов технологий, составляющих множество X, требуется выбрать набор с минимальными затратами. Вектор c состоит из удельных затрат на единицу каждой технологии.

Двойственная задача (23) выражает проблему определения рациональной системы экономических оценок продукции, составляющих вектор u. Задачи (22), (23) тесно взаимосвязаны. В частности, справедливо следующее утверждение: для того, чтобы вектор $\bar{x} \in X$ был оптимальным решением задачи (22) необходимо и достаточно, чтобы существовал вектор $\bar{u} \in U$, при котором выполняется условие дополняющей нежесткости

$$\bar{x}_j g_j(\bar{u}) = 0, \quad j = 1, \ldots, n. \tag{26}$$

Это условие играет важную роль в экономических приложениях. Компоненты вектора $\bar{u}$ интерпретируются как набор цен, «зовущих к оптимуму», а компоненты вектора $g(\bar{u})$ – в качестве экономических оценок эффективности отдельных технологий.

По каким-то причинам, в т.ч., например, из-за несовершенств линейной модели, может быть выбран неоптимальный с позиций задачи (22) план. Для случая, когда принятое решение $x \in X$ неоптимально, Л.В. Канторович в 1965 г. предложил методику формирования приближенных оценок ресурсов, основанную на минимизации взвешенной суммы квадратов отклонений в условиях дополняющей нежёсткости (26).

Пусть имеется допустимое решение $x^k \in X$ с положительными всеми компонентами $x_j^k > 0$, $j = 1, \ldots, n$. По одному из вариантов методики Л.В. Канторовича для такого решения вектор оценок определяется правилом

$$u^k = \arg\min \Phi_k(u), \; u \in R^m, \tag{27}$$

где

$$\Phi_k(u) = \sum_{j=1}^{n} (x_j^k)^2 (g_j(u))^2 . \tag{28}$$

Если вектор x^k был неоптимальным решением задачи (11), то у вектора $g(u^k)$ будут как положительные, так и отрицательные компоненты. По экономическим соображениям из того, что $g_j(u^k) > 0$ следует необходимость увеличения интенсивности использования данной технологии j в целях уменьшения суммарных затрат. Если $g_j(u^k) < 0$, то технология j имеет отрицательную рентабельность в рамках цен u^k и интенсивность ее использования должна быть сокращена.

На основе приведенного правила получения экономических оценок ресурсов и технологий И.И. Дикиным в 60-х годах был разработан [12, 13]

алгоритм итеративного улучшения решения задачи (22), в котором следующее решение $x^{k+1} \in X$ определяется по формуле

$$x^{k+1} = x^k + \lambda_k s^k$$

при

$$s_j^k = (x_j^k)^2 (g_j(u))^2, \ j = 1, \ldots, n,$$

$$\lambda_k = 1 \Big/ \sqrt{\Phi_k(u^k)}.$$

Первые программные реализации такого «непривычного» на фоне широко использовавшегося в 60-х годах симплекс-метода показали хорошую работоспособность. И.И. Дикиным были доказаны следующие важные факты (при предположении о невырожденности задачи (22)). Если задача (22) имеет оптимальное решение, то последовательность векторов x^k, $k = 1, 2, \ldots$, будет сходиться к одному из оптимальных решений. Причем получаемое оптимальное решение обладает важной особенностью – оно имеет минимальный набор активных ограничений по сравнению с другими оптимальными решениями. Скорость сходимости значения целевой функции $c^T x^k$ к ее оптимальному значению – линейная, т.е. по геометрической прогрессии. Последовательность векторов u^k будет сходиться к оптимальному решению двойственной задачи [14].

Данный алгоритм в 70-х годах успешно развивался в России, в том числе в работах Ю.Г. Евтушенко, В.И. Зоркальцева, В.Г. Жадана. Были проведены интересные исследования непрерывных аналогов алгоритмов. Разработаны новые правила выбора шага λ_k, эффективный способ ввода в область допустимых решений задачи математического программирования [6], двойственные алгоритмы внутренних точек. Уже с 70-х годов данные алгоритмы нашли широкое применение при реализации многих моделей энергетики [6]. С середины 80-х годов это семейство алгоритмов привлекло повышенное внимание во всём мире.

Метод вспомогательной функции и модифицированной функции Лагранжа

В конце 60-х годов сотрудник Института систем энергетики Ш.С. Чурквеидзе предложил использовать для решения разных проблем в оптимизации специальные и, вместе с тем, простые вспомогательные квадратичные функции. С помощью таких добавок он предлагал решать проблемы плохой устойчивости к погрешностям вычислений задач безусловной и условной оптимизации, эффективно учитывать ограничения на переменные.

Фактически им был создан новый, оригинальный метод решения широкого класса задач оптимизации, в т.ч. линейного и нелинейного программирования, который сразу нашел эффективное применение в ИСЭМ при оптимизации краткосрочных и долгосрочных режимов электроэнергетических систем [15, 16, 17]. Этот метод используется и поныне в программно-вычислительном комплексе, предназначенном для выбора оптимальных вариантов долгосрочного развития электроэнергетических систем.

Идеи Ш.С. Чурквеидзе вылились уже в 70-х годах в особое направление методов оптимизации, получившее название метода модифицированной функции Лагранжа. Теоретическими исследованиями и разработками этого метода сначала занимались только российские математики, в т.ч. Б.Т. Поляк, В.В. Третьяков, А.С. Антипин [18, 19]. Затем к этому методу возник повышенный интерес во всём мире.

В настоящее время метод модифицированной функции Лагранжа по праву считается одним из интереснейших в идейном плане и эффективным в вычислительном отношении подходом в теории и методах оптимизации.

Рассмотрим задачу оптимизации при ограничениях в форме равенств. Пусть f_0 – целевая функция, f_i, $i = 1, \ldots, m$ – функции-ограничения от вектора переменных $x \in R^n$. Обсуждается задача

$$f_0(x) \to \min, \tag{29}$$

при ограничениях

$$f_i(x) = 0,\ i = 1, \ldots, m. \tag{30}$$

Будем считать, что все функции дифференцируемые.

Наличие ограничений даже только в форме равенств несколько усложняет проблему поиска экстремума по сравнению с задачами безусловной оптимизации. Как известно, Лагранж предложил общую очень плодотворную идею перехода от задачи (29), (30) к задаче без ограничений. Он предложил добавить к целевой функции f_0 функции ограничений, просуммированные с некоторыми множителями, т.е. перейти к функции

$$L(x, \lambda) = f_0(x) - \sum \lambda_i f_i(x), \tag{31}$$

где λ_i – коэффициенты, называемые множителями Лагранжа, L – функция Лагранжа, зависящая от вектора исходных переменных x и вектора множителей Лагранжа $\lambda \in R^m$.

В результате приравнивания производных функции Лагранжа нулю, получим необходимые условия оптимальности. Равенство частных производных L по переменным λ_i дает собственно не что иное, как выполнение условий (30), так как

$$\nabla_\lambda L(x, \lambda) = f(x),$$

где f – вектор-функция с компонентами $f_i,\ i = 1, \ldots, m$.

Поскольку

$$\nabla_x L(x, \lambda) = \nabla f_0(x) - \sum \lambda_i \nabla f_i(x),$$

то приравнивание частных производных L по x к нулю дает соотношение

$$\nabla f_0(x) = \sum \lambda_i \nabla f_i(x). \tag{32}$$

Итак, равенство частных производных функций Лагранжа нулю означает, что полученное решение x является допустимым, т.е. удовлетворяющим условиям (30). При этом градиент целевой функции, согласно (32), должен находиться в подпространстве, образованном линейными комбинациями градиентов функций-ограничений.

В некоторых случаях условия (30), (32) являются достаточными для того, чтобы данное значение x было оптимальным решением задачи (29), (30). В частности, это имеет место для задач с выпуклой целевой функцией f_0 и линейными функциями-ограничениями.

Центральная идея Чурквеидзе состояла в использовании квадратичных аппроксимаций вида $(\lambda_i - f_i(x))^2$ в качестве добавок к минимизируемым функциям. В данном случае величина λ_i выполняет роль «маяка», к которому должно стремиться значение функции f_i.

В частности, для задачи (29), (30) получаем следующую вспомогательную функцию

$$S(x, \lambda) = f_0(x) + \frac{1}{2}\sum_{i=1}^{m}(\lambda_i - f_i(x))^2 .$$

Несложно увидеть, что данная функция содержит в качестве составной части функцию Лагранжа (31). Действительно,

$$S(x, \lambda) = L(x, \lambda) + \frac{1}{2}\sum_{i=1}^{m}(\lambda_i^2 + (f_i(x))^2) .$$

Появление дополнительных квадратичных составляющих весьма существенно. Они объясняют и использование термина «модифицированная функция Лагранжа».

На рассматриваемом примере хорошо можно пояснить суть метода модифицированной функции Лагранжа. Пусть на итерации k имеется текущее значение вектора переменных x^k и приближенное значение λ^k вектора множителей. Решая задачу безусловной оптимизации, находим вектор

$$x^{k+1} = \arg\min\{S(x, \lambda^{k+1}) + \gamma\sum_{j=1}^{n}(x_j - x_j^k)^2 : x^{k+1} \in R^n\}$$

при некотором положительном γ. На исходной итерации можно использовать любые значения λ^0, x^0. Вместо регуляризирующей

составляющей $\gamma\sum_{j=1}^{n}(x_j - x_j^k)^2$ можно вводить другие квадратичные добавки, как это делал Ш.С.Чурквеидзе, для противодействия «нехорошим свойствам» целевой функции.

Точного решения этой вспомогательной задачи не требуется. Важно, чтобы точность нарастала по итерациям. Чурквеидзе использовал метод покоординатного спуска для решения подобных вспомогательных задач.

На основе полученных значений вектора $f^{k+1} = f(x^{k+1})$ осуществляется корректировка вектора λ. Значение λ_i^k должно быть увеличено, если $f_i^{k+1} < 0$ (что будет стимулировать к увеличению $f(x)$ на следующей итерации) или уменьшено, если $f_i^{k+1} > 0$. В частности, можно положить

$$\lambda^{k+1} = \lambda^k - f^{k+1}.$$

Смысл корректировок λ состоит в том, чтобы достичь допустимого решения вспомогательной задачи, при котором $f(x) = 0$. Пусть

$$\bar{x} = \arg\min S(x, \bar{\lambda})$$

при некотором $\bar{\lambda} \in R^n$. Причем $f(\bar{x}) = 0$, т.е. нет необходимости в корректировке $\bar{\lambda}$. Тогда, как несложно убедиться, вектор $\bar{\lambda}$ будет состоять из множителей Лагранжа решения $\bar{x}$ задачи (29), (30), удовлетворяющего необходимым условиям оптимальности.

Существуют различные варианты изложенного метода, применимые к разным типам задач математического программирования.

Хотелось бы отметить, что не только сам новый метод оптимизации, который может обрастать разными модификациями, является открытием Ш.С. Чурквеидзе. Не менее важным открытием является введенный им новый взгляд на множители Лагранжа.

Методы отсечений и погружений

Большой вклад в развитие теории и методов оптимизации внёс профессор В.П. Булатов. Он являлся создателем лаборатории «Исследования операции» и отдела «Прикладной математики» в институте, в рамках которого велись широко известные в России и во всём мире работы по вычислительной математике. Он был организатором регулярно проводимой раз в 3 года с 1967 года Международной школы-семинара «Методы оптимизации и их приложения». На этой школе побывали многие известные российские и зарубежные специалисты в области прикладной математики.

Основные научные результаты В.П. Булатова сосредоточены в разработке и теоретическом исследовании класса методов отсечения и погружения. Воспользуемся следующей классификацией разрабатывавшихся В.П. Булатовым алгоритмов [20]. Все алгоритмы назовем методами отсечения. Они состоят из двух классов: «методов погружения» и «методов центрированных отсечений».

Методы погружений В.П. Булатов разделял на два типа. Первый из них – метод последовательного погружения надграфика целевой функции. Рассматривается задача

$$\varphi(x) \to \min, \; x \in K \qquad (33)$$

где x – вектор переменных из R^n, $\varphi(x)$ – целевая функция, K – множество допустимых решений. В.П. Булатов обычно рассматривал случай, когда K – компакт, т.е. замкнутое и ограниченное множество в R^n. Хотя некоторые его алгоритмы могут работать и в случае, если K – неограниченное множество.

В методе последовательного погружения надграфика целевой функции считается, что множество допустимых решений обладает «достаточно хорошими свойствами» и явно задано. Например, это множество решений некой системы линейных неравенств. Считается, что все сложности сосредоточены только в целевой функции $\varphi(x)$, которая может быть, например, недифференцируемой, неявно заданной и т.д.

Исходная идея метода последовательного погружения состоит в замене задачи (33) на эквивалентную ей задачу минимизации линейной функции с увеличенным на единицу количеством переменных. Введем функцию

$$f(x,a) = (x) - \alpha, \tag{34}$$

где α – дополнительная переменная. Тогда задача (33) становится равносильной задаче

$$\alpha \to \min, \tag{35}$$

при условиях

$$(x,\ \alpha) \in K_{n+1},\ x \in K, \tag{36}$$

где

$$K_{n+1} = \{(x,\ \alpha):\ f(x,\ \alpha) \le 0\}. \tag{37}$$

Суть метода состоит в том, что в процессе решения используется некоторая аппроксимация множества K_{n+1}.

Пусть на итерации k определена аппроксимация $\widetilde{K}_{n+1}^{k}$. Решается вспомогательная задача

$$\alpha \to \min,$$

при условиях

$$(x,\ \alpha) \in \widetilde{K}_{n+1}^{k},\ x \in K.$$

Пусть x^k, α^k – решение этой задачи. Если это не оптимальное решение задачи (35), (36), то переходим путем отсечения точки $(x^k,\ \alpha^k)$ от $\widetilde{K}_{n+1}^{k}$ к другому множеству $\widetilde{K}_{n+1}^{k+1}$. Отсечение должно быть таковым, чтобы при этом не отбросить оптимальное решение задачи (27), (28).

Одним из способов «отсечений» является введение дополнительных линейных неравенств. Такой подход применялся ранее для решения задач целочисленного линейного программирования (алгоритмы отсечений Гомори) и для решения задач выпуклого программирования (алгоритм Келли). В указанных алгоритмах происходит постепенное «сужение» аппроксимирующего множества

$$\widetilde{K}_{n+1}^1 \supset \widetilde{K}_{n+1}^2 \supset \widetilde{K}_{n+1}^3 \supset \dots \ .$$

В.П. Булатов разрабатывал алгоритмы с линейными отсечениями, в которых указанное включение не выполняется, в т.ч. из-за исключения некоторых введенных на предыдущих итерациях «отсечения» линейных неравенств.

Главное, что В.П. Булатовым была предпринята во многом успешная попытка создания общей теории отсечений, ведущих к оптимальному решению. В практическом плане кроме линейных отсечений интерес представляет рассматривавшийся им случай отсечений с помощью вогнутых функций. Если целевая функция $\varphi(x)$ невыпуклая, то для общего случая трудно предложить эффективный способ нахождения глобального оптимального решения задачи (33). В тех случаях, когда функция $\varphi(x)$ допускает представление в виде суммы выпуклой функции и невыпуклой функции, хорошо аппроксимируемой снизу вогнутой функцией, имеются конструктивные пути нахождения глобального минимума $\varphi(x)$ на компакте K.

Второе, более общее направление в методах «погружения» названо последовательным погружением допустимого множества. Это обобщение рассчитано и на случай, когда в ограничениях задачи присутствуют «плохие» ограничения. Собственно, этот случай можно свести к предыдущему. Пусть у задачи (33) имеются дополнительные ограничения

$$g_i(x) \le 0,\ i = 1, \dots, m,$$

где g_i – некоторые «нехорошие» функции, в т.ч., например, невыпуклые или недифференцируемые. Введем функцию

$$\widetilde{f}(x,\alpha) = \max\{f(x,\alpha),\ g_1(x),\ \dots,\ g_m(x)\}.$$

Затем рассмотрим задачу (35), (36), у которой множество K_{n+1} определено условием

$$K_{n+1} = \{(x,\ \alpha) : \widetilde{f}(x,\ \alpha) \le 0\}.$$

В итоге приходим к рассмотренному первому случаю.

В 1979 году в Докладах академии наук (ДАН) СССР вышла статья математика из Вычислительного центра Российской Академии наук Хачияна Л.Т. «Полиномиальные алгоритмы в линейном программировании». В этой статье рассматривалась модификация предложенного ранее известными математиками А.С. Немировским, Д.Б. Юдиным и Н.З. Шором метода эллипсоидов. Суть метода состояла в том, что допустимая область, содержащая оптимальное решение, погружалась в некий n-мерный эллипсоид. От этого эллипсоида на каждой итерации по специальным правилам отсекалась какая-то часть, не содержащая искомую точку оптимума. Оставшаяся часть погружалась в следующий эллипсоид меньшего объема. В результате сокращений на каждой итерации объемов эллипсоидов, содержащих искомую точку оптимума, за конечно число итераций можно было получить приближение к точке оптимума с любой точностью.

Изначально метод эллипсоидов был разработан для задач выпуклого программирования, частными случаями которого, как известно, являются задачи линейного программирования. Хачиян получил важный теоретический результат – алгоритм эллипсоидов позволяет решать задачи линейного программирования за время, выражающееся полиномиальными функциями от показателей размерности задачи.

В.П. Булатову в самом начале истории результата о возможности решения задач линейного программирования за полиномиальное время пришла в голову простая мысль – не только эллипсоидами можно аппроксимировать множество решений. В итоге его обсуждения с еще двумя ведущими математиками СЭИ (это были И.А. Александров и Е.Г. Анциферов) родился метод «центрированных отсечений», основанный на аппроксимации области, содержащей точку оптимума, n-мерным симплексом. Симплексом называется телесный линейный многогранник в n-мерном пространстве, содержащий ровно $n+1$ вершину. Разделение на каждой итерации многогранника на две части (содержащую и не

содержащую точу оптимума) осуществлялось гиперплоскостью, проходящей через «центр тяжести» многогранника. Отсюда и название алгоритма – метод центрированных отсечений.

В итоге совместных и, затем, индивидуальных исследований все три автора пришли к полиномиальному алгоритму центрированных отсечений с оценками (коэффициентом при степени полинома) лучшим, чем для метода эллипсоидов.

Список литературы

[1] Системные исследования в энергетике: перспективы научных направлений СЭИ – ИСЭМ. Отв. ред. Н.И. Воропай. – Новосибирск: Наука, 2010.

[2] Кузнецов Ю.А., Макаров А.А., Мелентьев Л.А. Система математических моделей для оптимизации перспективных энергетических балансов. Теплоэнергетика, 1966. – № 2.

[3] Модели развития энергетики и согласования решений // Сб. научных трудов под ред. А.А. Макарова. – Иркутск: СЭИ СО АН СССР, 1984.

[4] Могирев В.В. Алгоритм и программа вычисления показателей надежности электроэнергетических систем методом статистического моделирования (программа «Поток») // Методические вопросы исследования надежности больших систем энергетики. – Иркутск: СЭИ, 1975. – Вып. 4. – С.4 – 35.

[5] Иванов В.В., Дикин И.И., Колосок Г.В. Развитие процедур вероятностного моделирования при анализе надежности сложных ЭЭС // Тезисы докладов IX Всесоюз. науч. конф. «Моделирование электроэнергетических систем». – Рига, 1987. – С.344 – 345.

[6] Дикин И.И., Зоркальцев В.И. Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). – Новосибирск: Наука, 1988. – 144 с.

[7] Zorkaltsev, V., Perzhabinsky S. Model for power shortage estimation in electric power systems. Interior point algorithms. International Journal of Energy Optimization and Engineering. (appear)

[8] Хасильев Л.Я., Меренков А.П. Теория гидравлических цепей. – М.: Наука, 1985.

[9] Епифанов С.П., Зоркальцев В.И. Симметрическая двойственность в оптимизации и гидравлические цепи. // Вычислительные технологии, 2009. – № 1.

[10] Крумм Л.А. Методы решения общих уравнений стационарного режима электрической системы с учетом статистических характеристик нагрузок и генераторов при автоматическом регулировании частоты, напряжения и мощности // Труды Таллиннского политехнического института, 1957. – №124.

[11] Крумм Л.А. Две формы более общих уравнений экономичного режима объединенных энергосистем // Известия СО АН СССР, 1959.

[12] Дикин И.И. Итеративное решение задач линейного и квадратичного программирования // Доклады Академии наук, 1967. – Т. 141. – №4. – С. 747 – 748.

[13] Дикин И.И. Итеративные алгоритмы решения задач линейного, квадратичного и выпуклого программирования // Опыт решения экономических задач математическими моделями. – Новосибирск: Наука, 1967. – С. 31 – 38.

[14] Дикин И.И. О сходимости одного итерационного процесса // Управляемые системы. – Новосибирск, 1974. – Вып. 12. – С. 54 – 60.

[15] Сыров Ю.П., Чурквеидзе Ш.С. Некоторые методологические вопросы решения задачи оптимизации управления длительными режимами электроэнергетических систем // Методы управления большими системами. – Иркутск, 1970. – Т.2.

[16] Сыров Ю.П., Чурквеидзе Ш.С., Посекалин В.В. Метод вспомогательных характеристик для решения задачи оптимизации суточных режимов

электроэнергетических систем с использованием метода вспомогательной функции // Труды Иркутского политехнического института. Иркутск, 1971. – Вып. 2.

[17] Чурквеидзе Ш.С., Посекалин В.В. Метод и алгоритм оптимизации перспективных суточных режимов электроэнергетических систем // Математические модели для анализа и экономической оценки вариантов электроэнергетических систем. – Иркутск, 1971.

[18] Поляк Б.Т., Третьяков Н.В. Об одном итерационном методе линейного программирования, его экономической интерпретации // Экономика и математические методы. – Т. VIII. – Вып. 5. – 1972.

[19] Антипин А.С. Методы линейного программирования, основанные на прямой и двойственной модификации функции Лагранжа. – М.: ВЦ АН СССР, 1979. – 74 с.

[20] Булатов В.П. Методы погружения в задачах оптимизации. – Новосибирск: Наука, 1977. – 161 с.

УДК 519.8

СУБГРАДИЕНТНЫЙ АЛГОРИТМ МИНИМИЗАЦИИ С ПРЕОБРАЗОВАНИЕМ ПРОСТРАНСТВА $\alpha(\varepsilon)$ – АЛГОРИТМ

Журбенко Н. Г.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: zhurbnick@yandex.ua

Аннотация. Приводится субградиентный алгоритм минимизации выпуклой функции в конечномерном евклидовом пространстве с преобразованием пространства. Алгоритм основан на процедуре одномерного спуска и является в некотором смысле монотонным. Алгоритм обеспечивает решение задачи с заданной точностью за конечное число итераций. Приводится оценка трудоемкости алгоритма при решении задачи $\varepsilon-$оптимизации.

Ключевые слова. Оптимизация, $\varepsilon-$субградиент, преобразование пространства, отсекающая плоскость, эллипсоид локализации, трудоемкость алгоритма.

Введение

Более 40 лет назад был разработан субградиентный алгоритм минимизации с растяжением пространства в направлении разности двух последовательных градиентов – $r-$алгоритм [1]. Практика использования $r-$алгоритма показывает, что до сих пор он является одним из наиболее эффективных алгоритмов негладкой оптимизации. Однако теоретическое исследование эффективности $r-$алгоритма далеко не закончено (даже для класса выпуклых функций).

В данной статье приводится описание $\alpha(\varepsilon)-$алгоритма – результаты автора по разработке субградиентного алгоритма сравнимой с $r-$алгоритмом эффективностью и с его основными характеристиками: использование процедуры одномерной минимизации и растяжение пространства с большими коэффициентами.

Задача $\bar{\varepsilon}$ –минимизации

Рассматривается задача минимизации ограниченной снизу выпуклой функции $f(x)$ в $n-$мерном евклидовом пространстве R^n:

$f^* = \min\{f(x) | x \in R^n\}$, X^* – множество точек минимума функции $f(x)$. Для заданного числа $\bar{\varepsilon} \geq 0$ введем $\bar{\varepsilon}$–оптимальное множество $X^*(\bar{\varepsilon})$,

$$X^*(\bar{\varepsilon}) = \{x \in R^n | f(x) \leq f^* + \bar{\varepsilon}.$$

Произвольную точку $\tilde{x} \in X^*(\bar{\varepsilon})$ и значение функции $\tilde{f} = f(\tilde{x})$ будем называть решением задачи $\bar{\varepsilon}$–оптимизации ($\bar{\varepsilon}$–решением).

Пусть задан шар начальной локализации $\bar{\varepsilon}$–решения $D(z, R)$: $z \in R^n$ – центр шара, $R > 0$ – его радиус. Будем предполагать, что шар $D(z, R)$ содержит хотя бы одну точку минимума $x^* \in X^*$. Поэтому $D(z, R) \cap X^*(\bar{\varepsilon}) \neq \varnothing$. Далее характеристика множества начальной локализации будет уточнена (см. замечание к утверждению 10).

ε–субградиент

Мы будем использовать несколько отличное от классического определение ε–субградиента [2 – 4]. Вектор $g \in R^n$ называется $(\varepsilon, \tilde{f})$ – субградиентом функции $f(x)$ в точке z, если для $\forall x$ выполняется неравенство:

$$f(x) \geq \tilde{f} + (g, x - z) - \varepsilon, \qquad (1)$$

где $\tilde{f} \geq f^*; \varepsilon \in R^1$. Значение $\tilde{f}$ на k – ой итерации алгоритма решения задачи $\bar{\varepsilon}$ – оптимизации обычно равно полученному рекордному значению функции $f(x)$. Если значение f^* известно априори, то $\tilde{f} = f^*$. Обычное классическое определение ε – субградиента соответствует определению $(\varepsilon, \tilde{f})$ – субградиента (1) при $\tilde{f} = f(z)$, $\varepsilon \geq 0$. Обобщенный градиент (субградиент) функции $f(x)$, в т. z в классическом смысле [5] является $(0, f(z))$–субградиентом: $\partial f(z) = G(0, f(z), z)$.

Через $G(\varepsilon, \tilde{f}, z)$, $\partial f(z)$ будем обозначать множество $(\varepsilon, \tilde{f})$ – субградиентов и множество обобщенных градиентов функции $f(x)$ в т. z соответственно. В дальнейшем предполагается, что имеются алгоритмы вы-

числения $f(x)$ и $g(x) \in \partial f(x)$ в произвольной точке x. Пусть $g \in G(\varepsilon_1, \widetilde{f}_1, z_1)$; $f^* \le f_2 \le f(z_2)$, тогда $g \in G(\varepsilon_2, \widetilde{f}_2, z_2)$, где

$$\varepsilon_2 = \varepsilon_1 + \widetilde{f}_2 - \widetilde{f}_1 - (g, z_2 - z_1). \qquad (2)$$

Таким образом, $(\varepsilon, \widetilde{f})$ – субградиент в некоторой точке z_1 является $(\varepsilon, \widetilde{f})$ – субградиентом в любой другой точке z_2. Формула (2) определяет правило пересчета параметров $(\varepsilon, \widetilde{f})$ – субградиента. $G(\varepsilon, \widetilde{f}, z)$ по своим свойствам аналогично множеству $\partial f(z)$ (выпуклость, замкнутость и т.д., например, $\{0 \in G(\varepsilon, \widetilde{f}, z); \varepsilon \le \bar{\varepsilon}\} \Leftrightarrow \{\widetilde{f} \le f^* + \bar{\varepsilon}\}$).

Обозначим $\widetilde{X}(\bar{\varepsilon}, \widetilde{f}) = \{x \in R^n | f(x) \le \widetilde{f} - \bar{\varepsilon}\}$.

Отметим два простых (но важных для дальнейшего) утверждения.

Утверждение 1. *Если* $\widetilde{X}(\bar{\varepsilon}, \widetilde{f}) = \varnothing$, *то* $\widetilde{f}$ *является* $\bar{\varepsilon}$ *– решением.*

$$(\{\widetilde{X}(\bar{\varepsilon}, \widetilde{f}) = \varnothing\} \Rightarrow \{\forall x : f(x) > \widetilde{f} - \bar{\varepsilon}\} \Rightarrow \{f^* > \widetilde{f} - \bar{\varepsilon}\}).$$

Утверждение 2. *Если* $\widetilde{f} \ge f^* + 2\bar{\varepsilon}$, *то* $X^*(\bar{\varepsilon}) \subset \widetilde{X}(\bar{\varepsilon}, \widetilde{f})$.

$$(\{x \in X^*(\bar{\varepsilon})\} \Rightarrow \{f(x) \le f^* + \bar{\varepsilon}\}. \{f(x) \le f^* + \bar{\varepsilon}; \widetilde{f} \ge f^* + 2\bar{\varepsilon}\} \Rightarrow$$
$$\{x \in \widetilde{X}(\bar{\varepsilon}, \widetilde{f})\}.)$$

Агрегированный ε – субградиент

Пусть $P(z, \eta) = \{x \in R^n | (x - z, \eta) = 0\}$ – плоскость, проходящая через точку z с нормалью $\eta \in R^n, |\eta| > 0$. $P^+(z, \eta) = \{x \in R^n | (x - z, \eta) \ge 0\}$ – полупространство, определяемое (секущей) плоскостью $P(z, \eta)$. Следующее утверждение соответствует обычному построению отсекающих плоскостей "со сдвигом" для множества $\widetilde{X}(\bar{\varepsilon}, \widetilde{f})$. Пусть $g \in G(\varepsilon, \widetilde{f}, z)$; $h = (\bar{\varepsilon} - \varepsilon) / |g|$; $\widetilde{z} = z - hg / |g|$. Легко показать, что тогда $\widetilde{X}(\bar{\varepsilon}, \widetilde{f}) \subset P^+(x - \widetilde{z}, -g)$.

Замечание. Если $\widetilde{f} = f^*$ (задача с известным значением минимума), то субградиент $g \in \partial f(z)$ является (ε, f^*) – субградиент с $\varepsilon = -(f(z) - f^*)$.

Для случая обычной задачи минимизации ($\bar{\varepsilon} = 0$) величина "сдвига" $h = (\bar{\varepsilon} - \varepsilon) / | g | = (f(z) - f^*) / | g |$ в точности соответствует шаговому множителю метода Поляка [6].

Легко видеть, что $P^+(x - \tilde{z}, -g)$ определяется неравенством: $(x - z, g) \le -(\bar{\varepsilon} - \varepsilon),$ а вектор «сдвига» $\tilde{z} - z$ отсекающей плоскости можно определить как результат решения следующей задачи $\min\{1/2 \mid y \mid^2 : (y, g) \le -\tilde{\varepsilon}),$ где $\tilde{\varepsilon} = \bar{\varepsilon} - \varepsilon.$ Обобщим этот результат для множества субградиентов. Пусть $g_i \in G(\varepsilon_i, \tilde{f}, z), i = 1, \ldots, m, \tilde{\varepsilon}_i = \bar{\varepsilon} - \varepsilon_i.$ Тогда $\tilde{X}(\bar{\varepsilon}, \tilde{f})$ содержится в пересечении подпространств: $(x - z, g_i) \le -\tilde{\varepsilon}_i,$ (не исключается, что это пересечение пусто). Вектор «сдвига» y для множества ε–субградиентов $G(\varepsilon_i, \tilde{f}, z)$ из точки z определим решением следующей задачи:

$$\min 1/2 \parallel y \parallel^2 \tag{3}$$

$$(y, g_i) \le -\tilde{\varepsilon}_i; i = 1, \ldots, m. \tag{4}$$

Если система (4) несовместна, то вектор y не определен.

Утверждение 3. *Пусть система ограничений* (4) *– несовместна. Тогда* $\tilde{f}$ *является решением задачи* $\bar{\varepsilon}$*–оптимизации и выпуклая оболочка множества* ε*–субградиентов* $G(\varepsilon_i, \tilde{f}, z)$ *содержит 0.*

Утверждение 4. *Пусть:*

a) $\max\{\tilde{\varepsilon}_i; i = 1, \ldots, m\} > 0;$

b) система (4) совместна;

c) y, λ *– оптимальные значения прямых и двойственных переменных задачи* (3)*,* (4)*;*

d) $g = \sum_{i=1,m} \lambda_i g_i / \sum_{i=1,m} \lambda_i;$

e) $\varepsilon = \sum_{i=1,m} \lambda_i \varepsilon_i / \sum_{i=1,m} \lambda_i.$ (5)

Тогда:

$$g \in G(\varepsilon, \widetilde{f}, z), \tag{6}$$

$$y = -\frac{\bar{\varepsilon} - \varepsilon}{|g|^2} g. \tag{7}$$

Если условие b) не выполнено, то положим $g = 0$.

Вектор g, определяемый в утверждении 4, будем называть *агрегированным* ε–субградиентом множества субградиентов $G(\varepsilon_i, \widetilde{f}, z)$.

Геометрический смысл агрегированного ε–субградиента: агрегированный ε–субградиент принадлежит выпуклой оболочке множества субградиентов $G(\varepsilon_i, \widetilde{f}, z)$ и обеспечивает максимальный сдвиг отсекающей плоскости.

Утверждение 5. *Пусть для всех субградиентов* $\varepsilon_i = \varepsilon < \bar{\varepsilon}, i = 1, \ldots, m$. *Тогда агрегированный субградиент является вектором наименьшей длины выпуклой оболочки множества субградиентов (*$g = Nr\{g_1, \ldots, g_m\}$*).*

Таким образом, для множества ε–субградиентов $G(\varepsilon_i, \widetilde{f}, z)$ с одинаковыми значениями ε_i агрегированный ε–субградиент совпадает с кратчайшим вектором выпуклой оболочки $G(\varepsilon_i, \widetilde{f}, z)$. Именно этот вектор обычно используется при построении ε–субградиентных методов оптимизации. Введенное выше определение агрегированного ε–субградиента полезно в следующих смыслах. Во-первых, при решении задачи (3 – 4) может оказаться, что некоторое значение $\lambda_i > 0$ даже если соответствующее значение $\varepsilon_i > \bar{\varepsilon}$. Обычно такие ε–субградиенты при решении задачи $\bar{\varepsilon}$–оптимизации не используются. Во-вторых, при использовании $g = Nr\{g_1, \ldots, g_m\}$ все ε–субградиенты при выборе направления сдвига из точки z оказываются «равноправными» по отношению к значениям ε_i:

вектор $g = Nr\{g_1,\ldots,g_m\}$ не зависит от ε–параметров. Однако значения ε–параметров существенны для локализации $\bar{\varepsilon}$– решения.

Использование введенного агрегированного ε–субградиента позволяет получать модификации ε–субградиентных методов минимизации по обычной схеме их построения.

На основе утверждения 5 разработан оригинальный алгоритм решения задачи проектирования на политоп (определения вектора $g = Nr\{g_1,\ldots,g_m\}$) [8].

Утверждение 6. *Для заданной точки* z, *направления (спуска)* η $(|\eta|=1)$ *и* $\bar{\varepsilon} > 0$ *можно гарантировать вычисление такого* ε*–субградиента* $g \in G(\varepsilon, \tilde{f}, z)$, *что* $\varepsilon \le \bar{\varepsilon}$ *и* $(g,\eta) \ge 0$.

Алгоритм генерации ε–субградиента утверждения 6 имеет много общего с известными алгоритмами пополнения множеств ε–субградиентов [3], [9].

Локализаторы $\bar{\varepsilon}$–решения

Следующие два утверждения касаются построения специальных эллипсоидов минимального объема, содержащих подмножества шара простой структуры (сегмент и сектор).

Обозначим $W = D(z,R) \bigcap P^+(\tilde{z},\eta)$ – часть шара (сегмент) $D(z,R)$, определяемая секущей плоскостью $P(\tilde{z},\eta_1)$, где $\tilde{z} = z + h\eta; 0 \le h < R, |\eta| = 1$. W^* – эллипсоид минимального объема с центром в точке $\tilde{z}$, содержащий сегмент W, $\tilde{D}^*(z,\tilde{R})$ – шар минимального объема, содержащий эллипсоид W^*, $V(K)$ – объем тела K. $R_\alpha(\eta) = (\alpha - 1)\eta^T\eta + I$ – оператор растяжения пространства [5] по направлению η с коэффициентом α.

Утверждение 7. $\tilde{R} = \sqrt{R^2 - h^2}$; $\tilde{D}^*(z,\tilde{R}) = \Omega W^* = R_{\alpha^*}(\eta)W^*$, *где*

$$\alpha^* = \sqrt{(R+h)/(R-h)};$$

$$q = V(W^*)/V(D(z,\widetilde{R})) = \sqrt{(R-h)(R+h)}\,((R-h)/(R+h))^{n/2}.$$

Обозначим $W = D(z,R) \bigcap P^+(z,\eta_1) \bigcap P^+(z,\eta_2)$ – часть шара $D(z,R)$ (сектор), определяемая секущими плоскостями $P^+(z,\eta_1), P^+(z,\eta_2)$; W^* – эллипсоид минимального объема с центром в т. z, содержащий сектор W, $\widetilde{D}^*(z,\widetilde{R})$ – шар минимального объема, содержащий W^*; $\eta^* = (\eta_2 - \eta_1)/|\eta_2 - \eta_1|$, $\xi^* = (\eta_2 + \eta_1)/|\eta_2 + \eta_1|$, $\{\widetilde{\eta}_i, i = 1,\ldots,n-2\}$ – ортонормированный базис подпространства ортогонального векторам η^*, ξ^*.

Пусть $(\eta_1,\eta_2) = \cos(\varphi) \le 0$, $\psi = \pi - \varphi$ (ψ – угол между плоскостями $P(z,\eta_1), P(z,\eta_2)$).

Утверждение 8.

$$\widetilde{R} = \sqrt{2}\cos(\psi/2)R, \tag{8}$$

$$\widetilde{D}^*(z,\widetilde{R}) = \Omega W^* = \{R_{\alpha^*}(\eta^*)\textstyle\prod_{i=1}^{n-2} R_{\widetilde{\alpha}}(\widetilde{\eta}_i)\}\ W^*, \tag{9}$$

где

$$\alpha^* = ctg(\psi/2),\ \widetilde{\alpha} = \sqrt{2}\cos(\psi/2), \tag{10}$$

$$q = V(W^*)/V(D(z,R)) = \sin(\psi). \tag{11}$$

Заметим, что построение эллипсоида W^* приведено в работе [7]. Приведенный оператор преобразования эллипсоида W^* в шар Ω отличен от используемого в [7] «однорангового эллипсоидального оператора». Существенное отличие состоит в том, что все собственные числа самосопряженного оператора Ω не меньше 1.

Оператор преобразования Ω формулы (9) можно представить в следующем более компактном виде: $\Omega = \widetilde{\alpha} R_{\alpha^*/\widetilde{\alpha}}(\eta^*) R_{1/\widetilde{\alpha}}(\xi^*)$, ($\Omega = R_{\alpha^*}(\eta^*)\prod_{i=1}^{n-2} R_{\widetilde{\alpha}}(\widetilde{\eta}_i) = R_{\alpha^*/\widetilde{\alpha}}(\eta^*) R_{1/\widetilde{\alpha}}(\xi^*)(R_{\widetilde{\alpha}^*}(\eta^*) R_{\widetilde{\alpha}}(\xi^*)\prod_{i=1}^{n-2} R_{\widetilde{\alpha}}(\widetilde{\eta}_i)) = \widetilde{\alpha} R_{\alpha^*/\widetilde{\alpha}}(\eta^*) R_{1/\widetilde{\alpha}}(\xi^*).$)

Из утверждения 8 следует, что объем эллипсоида W^* меньше объема шара $\widetilde{D}^*(z,\widetilde{R})$ только для $\psi < \pi/2$ $(\varphi > \pi/2)$. Поэтому для использования такого эллипсоида в качестве локализации ε–решения необходим алгоритм генерации двух ε–субградиентов, угол между которыми тупой. Описание одного из таких алгоритмов приведено ниже.

Генерация сектора (алгоритм AlgSectorLocation)

Приведем итеративный алгоритм (AlgSectorLocation) генерации (по крайней мере) двух отсекающих плоскостей $P(z,\eta_1)$, $P(z,\eta_2)$ таких, что $(\eta_1,\eta_2) = \cos(\varphi) \le -1+\delta,$ для $\forall \delta > 0.$

Обозначим $E = \{e_1, e_2, \ldots, e_m\}$ – множество векторов $e_i \in R^n, i = 1,\ldots m$, $Nr\{E\}$ – вектор выпуклой оболочки множества E минимальной длины.

Описание алгоритма AlgSectorLocation

1 – ая итерация.

Вычисляем $g_1 \in \partial f(z).$ Если $g_1 = 0,$ то останов (задача минимизации $f(x)$ решена), иначе $e_1 = -g_1 / |g_1|.$

Используя процедуру одномерной минимизации по направлению e_1, вычисляем ε–субградиент $g_2 \in G(\varepsilon, \widetilde{f}, z),$ для которого $\varepsilon \le \bar{\varepsilon},$ $(g_2, e_1) \ge 0$ (см. утверждение 6). Если $g_2 = 0,$ то останов (задача $\bar{\varepsilon}$–минимизации $f(x)$ решена).

Полагаем:

$$G_1 = \{g_1, g_2\},\ e_2 = -g_2/|g_2|,\ E_1 = \{e_1, e_2\},$$

$$p_1 = Nr\{E_1\} = 1/2(e_1+e_2),\ \eta_1 = e_1,\ \eta_2 = e_2,\ \cos(\varphi_1) = (\eta_1,\eta_2) \le 0.$$

Если $p_1 = 0,$ то останов (задача $\bar{\varepsilon}$ – минимизации $f(x)$ решена). Если $\cos(\varphi_1) \le -1+\delta,$ то останов, иначе переход на следующую итерацию.

Пусть на итерации $k\,(k=1,2,\ldots)$ получены множество E_k, $(m_k = |E_k|; 2 \le m_k \le n)$ и вектор $p_k = Nr\{E_k\} \ne 0.$

Итерация $(k+1)$ $(k=1,2\ldots)$.

Используя процедуру одномерной минимизации по направлению p_k, вычисляем ε–субградиент $g_{m+1} \in G(\varepsilon, \tilde{f}, z)$, для которого $\varepsilon \le \bar{\varepsilon}$, $(g_{m+1}, p_k) \ge 0$ (см. утверждение 6). Если $g_{m+1} = 0,$ то останов (задача $\bar{\varepsilon}$–минимизации $f(x)$ решена). Заметим, что в результате одномерной минимизации значение $\tilde{f}$ может уменьшиться. В этом случае производится пересчет $(\varepsilon, \tilde{f})$ параметров векторов множества G_k согласно (2).

Полагаем:

$$\tilde{G}_{k+1} = G_k \bigcup g_{m+1},\ e_{m+1} = -g_{m+1} / |g_{m+1}|,$$

$$\tilde{E}_{k+1} = E_k \bigcup e_{m+1},\ p_{k+1} = Nr\{\tilde{E}_{k+1}\} = \sum_{i \in I} \lambda_i e_i,$$

где I – множество индексов i, для которых $\lambda_i > 0.$

Если $p_{k+1} = 0,$ то останов (задача минимизации $f(x)$ решена). Заметим, что если $p_{k+1} \ne 0,$ то $n \ge |I| \ge 2$ и $\lambda_{m+1} > 0$ (следствие $(g, \eta) \ge 0$).

Полагаем $G_{k+1} = \{g_i \in \tilde{G}_{k+1} \mid \lambda_i > 0\}$, $E_{k+1} = \{e_i \in \tilde{E}_{k+1} \mid \lambda_i > 0\}$ (E_{k+1} – множество «активных» векторов из $\tilde{E}_{k+1}$ относительно операции $Nr\{.\}$, $n \ge |E_{k+1}| \ge 2$). Перенумеруем вектора множества E_{k+1}, таким образом, чтобы $E_{k+1} = \{e_i, i = 1,2\ldots m;\}$, $\lambda_1 \ge \lambda_2 \ge \ldots \lambda_m > 0$ (вектора G_{k+1} перенумеровываются соответственно E_{k+1}).

Полагаем $\eta_1 = e_1$, $\eta_2 = \xi / |\xi|$, где $\xi = \sum_{i=2}^{m} \frac{\lambda_i}{1-\lambda_1} e_i$; $\cos(\varphi_{k+1}) = (\eta_1, \eta_2)$.

Если $\cos(\varphi_{k+1}) \le -1 + \delta,$ то останов, иначе переход на следующую итерацию.

Утверждение 9. *Алгоритм AlgSectorLocation конечен. При останове алгоритма выполняется одно из следующих условий: а) задача* $\bar{\varepsilon}$ *– минимиза-*

ции $f(x)$ *решена; b) получены такие отсекающие плоскости* $P(z,\eta_1), P(z,\eta_2)$, *что* $(\eta_1,\eta_2) = \cos(\varphi_{k+1}) \le -1+\delta$. *Справедливы следующие оценки значений* $|p_{k+1}|$ *и* $\sin(\psi_{k+1})$ $(\psi_{k+1} = \pi - \varphi_{k+1})$

$$|p_{k+1}| \le 1/\sqrt{2+k}, \tag{12}$$

$$\cos(\varphi_{k+1}) = (\eta_1,\eta_2) = \frac{p_{k+1}^2 - \lambda_1}{\sqrt{p_{k+1}^2(1-2\lambda_1)+\lambda_1^2}} \equiv F(p_{k+1}^2,\lambda_1), \tag{13}$$

$$\sin^2(\psi_{k+1}) \le \frac{p_{k+1}^2(1-p_{k+1}^2)}{p_{k+1}^2(1-2/n)+(1/n)^2} \equiv \Phi(p_{k+1}^2) \tag{14}$$

(*Здесь и далее* $p_{k+1}^2 = |p_{k+1}|^2$).

Приведем некоторые свойства формулы (13). Нетрудно видеть, что $\frac{\partial}{\partial\lambda_1} F(p_{k+1}^2,\lambda_1) = \frac{p_{k+1}^4 - p_{k+1}^2}{(p_{k+1}^2(1-2\lambda_1)+\lambda_1^2)^{3/2}} < 0$ для $p_{k+1}^2 < 1$. Так как $\lambda_1 \ge 1/n$, то отсюда следует оценка:

$$\cos(\varphi_{k+1}) \le \frac{p_{k+1}^2 - 1/n}{\sqrt{p_{k+1}^2(1-2/n)+(1/n)^2}}. \tag{15}$$

Пусть $\psi_{k+1} = \pi - \varphi_{k+1}$ (ψ_{k+1} – угол между плоскостями $P(z,\eta_1), P(z,\eta_2)$). Тогда из (15) следует

$$\sin^2(\psi_{k+1}) \le \frac{p_{k+1}^2(1-p_{k+1}^2)}{p_{k+1}^2(1-2/n)+(1/n)^2} \equiv \Phi(p_{k+1}^2). \tag{16}$$

Можно показать, что функция $\Phi(p_{k+1}^2)$ правой части (14) монотонно возрастает для $p_{k+1}^2 \in [0,1/n]$, причем $\Phi(0) = 0$, $\Phi(1/n) = 1$. Поэтому оценка (14) содержательна лишь для $p_{k+1}^2 < 1/n$. Например, $\Phi(1/n^2) = \frac{1}{2}(1+1/n)$. Поэтому, если $p_{k+1}^2 \le 1/n^2$, то $\sin(\psi_{k+1}) \le \frac{\sqrt{2}}{2}(1+1/n)$ $(\approx \frac{\sqrt{2}}{2}$ для $n >> 1)$. Отметим, что при этом, в соответствии с (10), (11), $q = \sin(\psi_{k+1}) \approx 0.7$

(любобытно заметить, что коэффициент уменьшения объма локлизации в методе центров тяжести для $n >> 1$ $q = 1 - 1/n \approx 0.632$); $\alpha^* = ctg(\psi_{k+1}/2) \approx 1.4$; $\widetilde{\alpha} = \sqrt{2}\cos(\psi_{k+1}/2) \approx 1.3$.

Из приведенных оценок (12), (14), следует, что алгоритм AlgSector-Location гарантирует коэффициент $q \approx 0.7$ не более чем за n^2 итераций. Автор, конечно, понимает, что эта оценка удручающе плоха. Однако это *оценка* трудоемкости и можно надеятся, что при реработе алгоритма она практически не будет реализовываться (пример такого алгоритма всем известен – симплекс алгоритм).

$\alpha(\varepsilon)$–алгоритм

В основе предлагаемого ε–субградиентного алгоритма минимизации лежит следующая простая идея. Задано число (параметр алгоритма) q, $0 < q < 1$.

Пусть на итерации k получены: точка x_k, A_k – невырожденное линейное преобразование ($B_k = A_k^{-1}$), $\widetilde{f}_k$ – рекордное значение функции $f(x)$, R_k – параметр эллипсоида локализации E_k множества $\widetilde{X}(\bar{\varepsilon}, \widetilde{f}_k)$: $\widetilde{X}(\bar{\varepsilon}, \widetilde{f}_k) \subset E_k \equiv \{| \, x \in R^n \, || \, A_k(x - x_k) \, | \leq R_k\}$.

Итерация $k+1$ алгоритма производится в преобразованном пространстве $Y_k = A_k X$, где X – исходное пространство переменных. Образом эллипсоида E_k в пространстве Y_k является шар $D(y_k, R_k)$, где $y_k = A_k x_k$. На итерации $k+1$ устанавливается факт решения задачи $2\bar{\varepsilon}$ – оптимизации или строится эллипсоид локализации множества $\widetilde{Y}(\bar{\varepsilon}, \widetilde{f}_{k+1}) \subset \widetilde{E}_k \equiv \{y \in Y_k \, || \, (\Omega_k(y - y_{k+1}) \, | \leq R_{k+1}\}$. При этом $V(\widetilde{E}_{k+1}) \leq qV(D(y_k, R_k))$. Построение этого эллипсоида и параметра $\widetilde{f}_{k+1}$ производится на основании использования процедур одномерной минимизации в соответствии с алгоритмом AlgSectorLocation. При этом на каждой итерации этого алгоритма вычисляются значения коэффициентов умень-

шения объемов эллипсоидов локализации множества $\widetilde{X}(\bar{\varepsilon}, \widetilde{f}_k)$ q_1 и q_2 согласно утверждениям 7 и 8 соответственно. Пусть $q_{k+1} = \min\{q_1, q_2\}$. Если $q_{k+1} \le q$, то алгоритм AlgSectorLocation прекращает работу.

Возможны два случая.

1. Если $q_{k+1} = q_1$, то оператор Ω_{k+1} определяется оператором Ω утверждения (7). Параметры утверждения (7) определяются следующим образом: $\eta = \widetilde{g}_k / |\widetilde{g}_k|$, $h = (\bar{\varepsilon} - \varepsilon)/|\widetilde{g}_k)|$, где $\widetilde{g}_k$ – агрегированный субградиент текущего множества субградиентов алгоритма AlgSectorLocation. Если $h > R_k$, то останов – задача решена. Иначе переход в новую точку (соответствует сдвигу отсекающей плоскости агрегированного субградиента в преобразованном пространстве): $x_{k+1} = x_k - hB_k\widetilde{g}_k / |\widetilde{g}_k|$.

2. Если $q_{k+1} = q_2 > q_1$, то оператор Ω_{k+1} определяется оператором Ω утверждения (8). Параметры утверждения (8) определяются следующим образом: $\eta_1 = e_1$, $\eta_2 = \xi / |\xi|$, где e_1, ξ – текущие вектора алгоритма AlgSectorLocation. Перехода в новую точку не происходит: $x_{k+1} = x_k$.

Преобразование пространства на итерации $k+1$ определяется оператором Ω_{k+1}: $A_{k+1} = \Omega_{k+1}A_k$.

На основании [11] можно доказать следующее утверждение об эффективности приведенного алгоритма.

Утверждение 10. *Для числа итераций k алгоритма, за которые он обеспечивает решение 2ε–оптимизации, справедлива следующая оценка:*

$$k \le \left\rceil n \frac{\ln(1/\gamma)}{\ln(1/q)} \right\lceil,$$

где γ – относительная точность решения задачи ($\gamma = \bar{\varepsilon}/(RC)$, где C оценка сверху норм субградиентов в исходном шаре локализации $D(z, R)$).

Замечания

1. При доказательстве утверждения 10 предполагается, что для шара начальной локализации $\bar{\varepsilon}$–решения $D(z,R)$ выполняется следующие условия. Пусть C–оценка сверху норм субградиентов в точках шара ограничена: $|g(x)| \le C (g(x) \in \partial f(x);\ x \in D(z,R))$. Шар $D(z,R)$ содержит такой шар $d(x^*, r_{\bar{\varepsilon}}) = \{x : | x - x^* | \le r_{\bar{\varepsilon}}\}$ (окрестность некоторой точки минимума $x^* \in X^*$), что все его точки являются $\bar{\varepsilon}$–решениями: $d(x^*, r_{\bar{\varepsilon}}) \subset X^*(\bar{\varepsilon})$. Учитывая ограниченность норм субградиентов, можно считать, что $r_{\bar{\varepsilon}} \ge \bar{\varepsilon} / C$.

2. Поясним смысл используемой относительной точности γ. Из ограниченности норм субградиентов, имеем следующую начальную оценку: $f(z) - f^* \le RC$. Для точки $\bar{\varepsilon}$–решения $\tilde{x} \in X^*(\bar{\varepsilon})$ $f(\tilde{x}) - f^* \le \bar{\varepsilon} = \gamma RC$, начальная оценка улучшена в γ раз.

Численная эффективность α(ε)–алгоритма.

Тестовые задачи:

$$f_1(x) = \sum_{i=1}^{n} \rho_n^{i-1} x_i^2,\ \ f_2(x) = \sum_{i=1}^{n} \rho_n^{i-1} | x_i |,$$

где $\rho_n = 10^{6/(n-1)}$. Степень вытянутости линий уровня («овражности») функций не зависит от размерности и равна 10^6. Начальная точка $x_i = 1.0, i = 1, \ldots n$. Параметр точности решения по функционалу $\bar{\varepsilon} = 10^{-6}$.

Обозначения колонок таблиц: nVarbl – число переменных; qVolum – значение параметра уменьшения объема локализации $\bar{\varepsilon}$ – решения на итерации (параметр алгоритма); nIter – число итераций; nLStep – среднее число применения алгоритма одномерной минимизации на одной итерации; Alpha – среднее значение коэффициента растяжения пространства; r – алгоритм – число итераций r–алгоритма.

Таблица 1

Минимизация функции $f_1(x)$

nVarbl	qVolum	nIter	nLStep	Alpha	r – агоритм
10	0.99	107	1.364	1.661	72
10	0.7	56	3.214	3.035	
20	0.99	195	1.297	1.621	103
20	0.7	86	3.686	2.724	
40	0.99	360	1.236	1.611	157
40	0.7	134	4.127	2.761	
50	0.99	435	1.205	1.6	190
50	0.7	153	4.255	2.759	
100	0.99	711	1.136	1.592	577
100	0.7	243	4.407	2.744	

Таблица 2

Минимизация функции $f_2(x)$

nVarbl	qVolum	nIter	nLStep	Alpha	r – агоритм
10	0.99	413	1.165	1.855	141
10	0.7	133	3.015	3.38	
20	0.99	1274	1.095	1.552	204
20	0.7	289	4.173	2.842	
40	0.99	1930	1.09	1.508	349
40	0.7	374	6.035	2.607	
50	0.99	2594	1.076	1.465	435
50	0.7	455	6.868	2.601	
100	0.9	4062	2.597	1.755	888
100	0.7	1559	9.201	2.547	

Приведенные численные эксперименты демонстрируют достаточно высокую эффективность $\alpha(\varepsilon)$ – алгоритма, сравнимую с эффективностью r – алгоритма [5] (по числу итераций). Результаты выявляют следующие интересные его особенности.

Среднее число применения алгоритма одномерной минимизации на одной итерации оказывается малым в сравнении с гарантированной его оценкой (n^2).

Даже если параметр гарантируемого уменьшения объема локализации ε – решения на каждой итерации задается малым (0.99), тем не менее, среднее значение коэффициента растяжения пространства оказывается существенно больше единицы.

Заключение

Качественная интерпретация $\alpha(\varepsilon)$–алгоритма состоит в следующем. Алгоритм относится к классу методов с преобразованием пространства [5]. На каждой итерации алгоритма преобразование состоит в применении операторов растяжения пространства по ортогональным направлениям. Параметры преобразования определяются построением эллипсоидов локализации ε–решения. Эллипсоиды локализации строятся на основе информации, получаемой в результате применения процедуры одномерной минимизации. На каждой итерации обеспечивается уменьшение объема локализации ε–решения в не менее чем заданное число q раз (параметр алгоритма). Если в результате процедур одномерной минимизации происходит существенное улучшение рекордного значения функции, то преобразование пространства определяется оператором растяжения по направлению агрегированного субградиента. В противном случае, за конечное число одномерных процедур минимизации гарантируется генерации (по крайней мере) двух $\varepsilon-$субградиентов с настолько тупым углом между ними, что он обеспечивает построение эллипсоида локализации $\varepsilon-$решения с требуемым уменьшением объема локализации. При этом преобразование пространства определяется операторами растяжения по $n-1$ ортогональным направлениям. Причем с максимальным коэффициентом производится растяжение по направлению разности двух ортонормированных субградиентов из выпуклой оболочки построенных субградиентов. Таким образом, образно выражаясь, $\alpha(\varepsilon)$–алгоритм объединяет алгоритмы Н.З. Шора с растяжением пространства по субградиенту и по разности двух последовательных субградиентов.

Получена оценка трудоемкости $\alpha(\varepsilon)$–алгоритма для решения задачи $\bar{\varepsilon}$–оптимизации. Численные эксперименты показывают, что (по числу итераций) эффективность $\alpha(\varepsilon)$–алгоритма сравнима с эффективностью

r–алгоритма Н.З. Шора. Однако трудоемкость одной итерации $\alpha(\varepsilon)$–алгоритма существенно больше.

Список литературы

[1] Шор Н.З., Журбенко Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – Киев: Наук. думка, 1971. – №3. – С. 51 – 59.

[2] Brondsted A. and Rockafellar R.T. On the subdifferentiability of convex functions // Proc. Amer. Math. Soc. 16, 1965. – Pp. 605 – 611.

[3] Lemarechal C., Mifflin K. Nonsmooth Optimization. – Oxford: Pergamon Press, 1978. – 180 p.

[4] Демьянов В.Ф., Васильев Л.В. Недифференцируемая оптимизация. – М.: Наука, 1981. – 384 с.

[5] Шор Н.З. Методы минимизации недифференцируемых функций и их применение. – Киев: Наук.думка, 1979. – 200 с.

[6] Поляк Б.Т. Минимизация негладких функционалов // Журн. вычислит. математики и матем. физики, 1969. – Т.9, № 3. – С. 507 – 521.

[7] Стецюк П.И. Ортогонализующие линейные операторы в выпуклом программировании (Часть I) // Кибернетика и системный анализ, 1997. – № 3. – С. 97 – 119.

[8] Журбенко Н.Г. Алгоритм проектирования на политоп // Теорія оптимальних рішень, 2008. – № 7. – С. 125 – 132.

[9] Ржевский С.В. Монотонные методы выпуклого программирования. – Киев: Наук. думка, 1993. – 319 с.

[10] Журбенко Н.Г. Об одном классе методов минимизации с преобразованием пространства // Методы решения экстремальных задач, 1996. – С. 68 – 80.

[11] Журбенко Н.Г. Оценка эффективности одного класса ε–субградиентных методов минимизации с преобразованием пространства // Оптимизация и ее приложения, 1997. – С. 49 – 54.

УДК 519.8

ОПЕРАТОРЫ ПРЕОБРАЗОВАНИЯ ПРОСТРАНСТВА В КВАЗИНЬЮТОНОВСКИХ МЕТОДАХ И МЕТОДАХ СОПРЯЖЕННЫХ ГРАДИЕНТОВ

Журбенко Н. Г.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: zhurbnick@yandex.ua

Аннотация. Приводится интерпретация квазиньютоновского условия и условия сопряженности направлений при построении квазиньютоновских методов и методов сопряженных направлений на основе использования операторов преобразования пространства.

Ключевые слова. Квазиньютоновский метод, сопряженные направления, оператор преобразования, минимизация.

Введение

Эффективным и часто используемым на практике средством решения задач минимизации гладких функций являются алгоритмы, относящиеся к классу квазиньютоновских методов и методов сопряженных направлений. Имеются различные интерпретации и схемы построения этих методов [1 – 3]. Общая схема квазиньютоновских методов и методов сопряженных направлений безусловной минимизации функции $f(x)$ в R^n определяется следующим итеративным процессом генерации приближений $x_k \in R^n$:

$$x_k = x_{k-1} - h_k H_{k-1} g_{k-1}, k = 1,2,\ldots, \quad (1)$$

где h_k – шаговый множитель,

$$h_k = \arg\min_h f(x_{k-1} - h H_{k-1} g_{k-1}); \quad (2)$$

g_{k-1} – градиент функции $f(x)$ в точке x_{k-1}; H_{k-1} – симметричная положительно определенная матрица.

После шага k матрица H_k определяется в форме аддитивной поправки матрицы H_{k-1}: $H_k = H_{k-1} + \nabla H_k$, где ∇H_k – матрица ранга 1 или 2. Выбранная корректировка ∇H_k матрицы H_{k-1} определяет конкретный алгоритм (1).

Известны следующие «градиентные» интерпретации метода (1). Первая из них обычно приводится для квазиньютоновских методов и состоит в следующем. Шаг метода (1) соответствует шагу градиентного метода в пространстве с определяемой матрицей H_{k-1} метрикой, т.е. в пространстве со скалярным произведением $(\cdot,\cdot)_H$: $(x,z)_H = (H_{k-1}^{-1}x, z)$. Поэтому квазиньтоновские методы часто называют методами переменной метрики.

Другая интерпретация метода (1) связана с операцией преобразованием пространства переменных линейным оператором. Представим матрицу H_k в виде $H_k = B_k B_k^T$, где B_k – невырожденная матрица $n \times n.$. Такое представление корректно поскольку матрица H_k невырождена и положительно определена. Положим $A_k = B_k^{-1}$. Пусть Y – «преобразованное» соответствующим матрице A_k линейным оператором пространство: $Y = A_k X$. Градиент $\widetilde{g}(y)$ функции $\widetilde{f}(y) = f(B_k x)$, соответствующей функции $f(x)$ в пространстве Y, равен $B_k^T g(x)$. Поэтому шаг метода (1) соответствует (в исходном пространстве X) шагу градиентного алгоритма в «преобразованном» пространстве Y.

Построение методов рассматриваемого класса обычно основано на основе анализа задачи минимизации квадратичной функции

$$\min\{f(x) = (Cx, x) + (b, x) + c, \tag{3}$$

где $C-$ положительно определенная симметричная матрица (для методов сопряженных направлений можно рассматривать случай $C \geq 0$, однако, для краткости, мы будем считать, что $C > 0$).

Квазиньтоновские методы и методы сопряженных направлений обеспечивают (и должны обеспечивать по определению) решение задачи (3) не более чем за n шагов. В дальнейшем, для краткости, будем считать, что решение обеспечивается за n шагов.

Для квазиньютоновских методов (по определению) должно выполняться условие: $H_n = C^{-1}$. Отсюда следует так называемое квазиньютоновское условие рекуррентного пересчета матрицы H_k:

$$H_k(g_k - g_{k-1}) = -h_k H_{k-1} g_{k-1}. \quad (4)$$

Для методов сопряженных направлений выполнение условия $H_{n+1} = C^{-1}$ не обязательно (если это условие выполняется, то такой метод относится к классу квазиньютоновских). Так в численных схемах простейших алгоритмов сопряженных направлений матрица H_k вообще отсутствует (например, алгоритмы Флетчера-Ривса [4], Полака-Рибьера [4], Поляка Б.Т. [5]). Варианты таких алгоритмов и алгоритмов с несимметричными матрицами мы рассматривать не будем.

Основой построения методов сопряженных направлений является требование $C-$ортогональности (сопряженности) направлений $p_k = H_k g_k$:

$$(Cp_k, p_j) = 0, j = 0,1,\ldots k-1. \quad (5)$$

Наиболее известная методика построения методов сопряженных направлений разработана Х. Хуангом (H. Y. Huang) [1]. Согласно приведенной в работе [3] описанию этой методики, условие $C-$ортогональности (5) будет выполнено (с учетом (2)), если выполнены соотношения:

$$H_k C p_j = \lambda p_j, j = 0,1,\ldots k-1, \quad (6)$$

где λ – произвольная неотрицательная константа.

Также как и для квазиньютоновских методов, после шага k матрица H_k определяется в форме аддитивной поправки.

После описания приведенных выше хорошо и давно известных фактов, нашей задачей является интерпретация квазиньютоновского условия (4) и условия $C-$ортогональности (6) с позиции использования операторов преобразования пространства: матрица H_k представляется в виде $H_k = B_k B_k^T$.

Квадратичной функции $f(x)$ задачи (3) в преобразованном оператором $A_k = B_k^{-1}$ пространстве $Y = A_k X$ имеет вид $\widetilde{f}(y) = (\widetilde{C}_k y, y) + (B_k^T b, x) + c,$ где $\widetilde{C}_k = B_k^T C B_k$ – матрица квадратичной части функции в преобразованном пространстве. Положим: $p_k = H_{k-1} g_{k-1}, \quad \widetilde{p}_k = A_k p_k$ ($\widetilde{p}_k$ – направление в преобразованном пространстве, соответствующее направлению «движения» p_k алгоритма (1) в исходном пространстве).

Пусть матрица преобразования B_k удовлетворяет следующему условию (вектор $\widetilde{p}_k$ является собственным вектором матрицы квадратичной функции в преобразованном пространстве; собственное значение вектора равно λ_k)

$$\widetilde{C}_k \widetilde{p}_k = \lambda_k \widetilde{p}_k. \tag{7}$$

Утверждение 1. *Квазиньютоновское условие* (4) *и условие* (7)*, в котором* $\lambda_k = 1,$ *эквивалентны.*

Естественная форма корректировки матрицы преобразования пространства A_k имеет вид

$$A_k = R_k A_{k-1} \ (B_k = B_{k-1} T_k, \text{ где } T_k = R_k^{-1}). \tag{8}$$

Такая форма соответствует последовательному преобразованию пространства: $Y_k = R_k Y_{k-1}$ ($Y_0 = X$).

Отметим, что представление известных вариантов квазиньютоновских методов (Давидона-Флетчера-Пауэлла, Бройдена-Флетчера-Шенно) в такой форме методов с преобразованием пространства приведено в работе [6].

Шаг (1) соответствует шагу обычного градиентного метода в пространстве Y_{k-1}:

$$y_k = y_{k-1} - h_k \widetilde{g}_{k-1}, \ k = 1,2,\ldots, \tag{9}$$

где $\widetilde{g}_{k-1} = B_{k-1}^T g_{k-1}$ – градиент функции $\widetilde{f}(y) = f(B_{k-1} x)$ в точке $y_{k-1} = B_{k-1} x_{k-1}.$

Пусть $\widetilde{g}_k = B_{k-1}^T g_k$ – градиент функции $\widetilde{f}(y) = f(B_{k-1}x)$ в точке $y_k = B_{k-1}x_k$ пространства Y_{k-1}, $\{\widetilde{g}_{k-1}, \widetilde{g}_k\}$ – двумерное подпространство пространства Y_{k-1} определяемое векторами $\widetilde{g}_{k-1}, \widetilde{g}_k$ Выделим подкласс методов с преобразованием пространства следующим ограничением на выбор операторов R_k: R_k оператор изменяет только подпространство $\{\widetilde{g}_{k-1}, \widetilde{g}_k\}$). Для выделенного подкласса методов с преобразованием пространства справедливо следующее утверждение.

Утверждение 2. *Условия C-ортогональности* (6) *и условие* (7)*, в котором $\lambda_k = \lambda$, эквивалентны.*

Заключение

В заключение сделаем несколько замечаний.

1. Отметим содержательный смысл рассматриваемого класса квазиньютоновских методов и методов сопряженных направлений, генерируемых на основе преобразования пространства. В преобразованном пространстве Y_{k-1} выполняется шаг градиентного алгоритма (9). После этого производится преобразование пространства Y_{k-1} (изменяющее лишь подпространство $\{\widetilde{g}_{k-1}, \widetilde{g}_k\}$) с выполнением условия (7). Содержательный смысл этого условия: *образ* (*в пространстве Y_k) вектора смещения в пространстве Y_{k-1} должен являться собственным вектором матрицы квадратичной функции в пространстве Y_k*. Оказывается, что такой принцип выбора операторов преобразования соответствует квазиньютоновскому условию (4) и известному условию C-ортогональности (6) методов сопряженных направлений. Для квадратичной функции образ траектория метода (1) в пространстве Y_n соответствует движению по (взаимно ортогональным) собственным векторам матрицы $\widetilde{C}_n$.

2. Приведенная интерпретация квазиньютоновского условия и условия C-ортогональности не только интересна своей наглядностью (удивительно, но автор не обнаружил публикацию такой интерпретации). Она позволяет

строить новые модификации методов рассматриваемого класса. Так в работах [7, 8] разработано семейство алгоритмов на основе использования условия (7) и операторов растяжения пространства $R_\alpha(\eta)$ [9]. Как частный случай в это семейство входит предельный вариант r-алгоритма [10] (одного из наиболее эффективных алгоритмов негладкой оптимизации) с бесконечным коэффициентом растяжения.

Список литературы

[1] H. Y. Huang Unified approach to quadratically convergent algorithms for function minimization // J. Optim. Theory and Appl., 1970. – Vol. 5. – Pp. 405 – 423.

[2] Полак Э.Численные методы оптимизации. – М.: Мир, 1974. – 376с.

[3] Пшеничный Б.Н., Данилин Ю.М. Численные методы в экстремальных задачах. – М.: Наука, 1975. – 376 с.

[4] Fletcher R., Reeves C.M. Function minimization by conjugate gradients // Comput. J., 1964. –Vol. 7. – №2. – Pp. 149 – 154.

[5] Поляк Б.Т. Метод сопряженных градиентов в задачах на экстремум // ЖВМ и МФ, 1969. – Вып.9. – №4. – С. 807 – 821.

[6] Стецюк П.И. Линейные операторы в квазиньютоновских методах // Теория и приложения методов оптимизации. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1998. – С. 3 – 8.

[7] Журбенко Н.Г. Построение семейства методов сопряженных направлений на основе использования оператора растяжения пространства // Теория и приложения методов оптимизации. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1998. – С. 12 – 18.

[8] Журбенко Н.Г. Квазиньтоновские алгоритмы минимизации на основе использования оператора растяжения пространства // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1999. – С. 45 – 50.

[9] Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.

[10] Шор Н.З., Журбенко Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – Киев: Наук. Думка, 1971. – №3. – С. 51 – 59.

УДК 519.6:519.85

ОБОСНОВАНИЕ АЛГОРИТМОВ ВНУТРЕННИХ ТОЧЕК ДЛЯ ЗАДАЧ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Зоркальцев В. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: zork@isem.sei.irk.ru

Пержабинский С. М.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: smper@isem.sei.irk.ru

Аннотация. Рассматривается семейство алгоритмов внутренних точек. Алгоритмы предназначены для решения задач математического программирования с нелинейными ограничениями-неравенствами. При поиске направления улучшения решения используются изменяющиеся по итерациям взвешенные евклидовы нормы. Представлены результаты теоретического обоснования алгоритмов при некоторых предположениях (в т.ч. о невырожденности задачи).

Ключевые слова. Метод внутренних точек, взвешенная евклидова норма, линеаризация.

Введение

Известно, что алгоритмы внутренних точек являются высокоэффективными процедурами решения задач математического программирования. Из множества алгоритмов внутренних точек в особый класс выделяются алгоритмы, в которых поиск направления улучшения решения основывается на идее стимулирования движения вдоль границ области допустимых по ограничениям-неравенствам решений. Это обуславливает оригинальность и эффективность этих методов, но в то же время затрудняет их теоретическое обоснование. Пионерные разработки таких алгоритмов были осуществлены в СССР в 60 – 70-х гг. прошлого века С.М. Анцызом, И.И. Дикиным [1, 2], Ю.Г. Евтушенко и В.Г. Жаданом [3, 4], В.И. Зоркальцевым [2, 5, 6]. Алгоритмы внутренних точек обсуждаемого типа успешно используются с 70-х гг. прошлого века при реализации ряда моделей энергетики в Институте систем энергетики им. Л.А. Мелентьева СО

РАН. При этом для моделей с нелинейными ограничениями применяются процедуры итеративной линеаризации.

Постановка задачи

Рассмотрим задачу нелинейного программирования

$$f_0(x) \to \min,\ x \in X, \tag{1}$$

где

$$X = \{x \in R^n :\ \underline{x} \le x \le \bar{x},\ f_i(x) \le 0,\ i = 1, \ldots, m\}. \tag{2}$$

Заданы векторы $\underline{x}$, $\bar{x}$ из R^n (причем $\underline{x}_j < \bar{x}_j$, $j = 1, \ldots, n$) и выпуклые дважды дифференцируемые функции $f_i(x)$, $i = 0, \ldots, m$.

Векторы из X будем называть допустимыми решениями. Подмножество векторов из X, удовлетворяющих ограничениям-неравенствам в строгой форме, обозначим

$$\operatorname{int} X = \left\{x \in R^n : f_i(x) < 0,\ i = 1, \ldots, m,\ \underline{x} < x < \bar{x}\right\}.$$

Будем считать, что $\operatorname{int} X \neq \phi$ и задан вектор x^0 из $\operatorname{int} X$. Следовательно, задача (1), (2) удовлетворяет условию Слейтера. Согласно условиям оптимальности Куна-Таккера, для того, чтобы вектор $x^* \in X$ был оптимальным решением задачи (1), (2), необходимо и достаточно существования векторов $u^* \in R^m$, $\alpha^* \in R^n$, $\beta^* \in R^n$, при которых:

$$\nabla f_0(x^*) + \sum_{i=1}^{m} u_i^* \nabla f_i(x^*) + \alpha^* - \beta^* = 0, \tag{3}$$

$$u_i^* f_i(x^*) = 0,\ i = 1, \ldots, m, \tag{4}$$

$$\beta_j^*(x_j^* - \underline{x}_j) = 0,\ j = 1, \ldots, n, \tag{5}$$

$$\alpha_j^*(\bar{x}_j - x_j^*) = 0,\ j = 1, \ldots, n, \tag{6}$$

$$u^* \ge 0,\ \alpha^* \ge 0,\ \beta^* \ge 0.$$

Допустимое решение $x \in X$ назовем стационарным, если для него выполняются соотношения (3) – (6) при некоторых векторах u, α, β. Если такие векторы единственные, то решение x будем называть невырожденным.

Задачу (1), (2) назовем невырожденной, если у нее любое стационарное решение невырождено. Заметим, что оптимальное решение является всегда стационарным, но не любое стационарное решение является оптимальным.

Описание семейства алгоритмов внутренних точек

В рассматриваемом вычислительном процессе последовательность векторов x^k из $\operatorname{int} X$ вырабатывается по правилу

$$x^{k+1} = x^k + \lambda_k \Delta x^k . \tag{7}$$

Здесь $k = 0, 1, 2, \ldots,$ – номер итерации, Δx^k – вектор спуска, λ_k – шаг корректировки.

Основной проблемой в излагаемом вычислительном процессе является определение направления улучшения решения. Для ее разрешения сначала в точке x^k вычисляется градиент целевой функции

$$c^k = \nabla f_0(x^k) .$$

Если $c^k = 0$, то дальнейшие вычисления прекращаются, т.к. x^k – точка минимума функции $f_0(x)$. Далее будем считать, что $c^k \neq 0$.

После этого находится матрица A_k размера $m \times n$, строками которой являются градиенты функции $f_i(x)$, $i = 1, \ldots, m$. Затем вычисляется диагональная матрица B_k размера $n \times n$ с положительными элементами на главной диагонали.

В работе [8] в качестве B_k использовалась взвешенная сумма матриц вторых производных функции $f_i(x)$, $i = 0, \ldots, m$,

$$B_k = \nabla^2 f_0(x^k) + \sum_{i=1}^{m} v_i^{k-1} \nabla^2 f_i(x^k) , \tag{8}$$

Здесь v_i^{k-1} – приближения на итерации $k-1$ к множителям Лагранжа нелинейных ограничений задачи (1), (2). На начальной итерации $v_i^0 = 1$, $i = 1, \ldots, m$. Экспериментальные исследования [8] такого алгоритма внутренних точек показали его вычислительную эффективность.

Если функции $f_i(x)$, $i = 0, \ldots, m$, сепарабельные, то матрица B_k является диагональной. В этой связи возможна модификация правила (8), в которой для всех i, $i = 0, \ldots, m$, вместо матрицы $\nabla^2 f_i(x^k)$ используется диагональная матрица того же размера с элементами $\dfrac{\partial^2 f_i(x^k)}{\partial x_j^2}$, $j = 1, \ldots, n$, размещенными на главной диагонали.

Затем определяем векторы весовых коэффициентов $d^k \in R^n$, $h^k \in R^m$. Их компоненты должны удовлетворять неравенствам

$$\underline{\sigma}(y_j^k) \le d_j^k \le \overline{\sigma}(y_j^k),\ j = 1, \ldots, n, \tag{9}$$

$$\underline{\sigma}(-f_i^k) \le h_i^k \le \overline{\sigma}(-f_i^k),\ i = 1, \ldots, m. \tag{10}$$

Здесь

$$y_j^k = \min\left\{\overline{x}_j - x_j^k,\ x_j^k - \underline{x}_j\right\},\ j = 1, \ldots, n,$$

$$f_i^k = f_i(x^k),\ i = 1, \ldots, m,$$

$\overline{\sigma}$, $\underline{\sigma}$ – некоторые непрерывные неубывающие функции положительного аргумента, удовлетворяющие двум условиям:

$$0 < \underline{\sigma}(t) \le \overline{\sigma}(t),\ \forall t > 0; \tag{11}$$

$$\overline{\sigma}(t) \le Mt, \tag{12}$$

при некоторых $\varepsilon > 0$, $M > 0$ для всех $t \in (0, \varepsilon]$.

Исследуемый вычислительный процесс представляет семейство алгоритмов, поскольку могут использоваться различные правила вычисления весовых коэффициентов. В частности, можно пользоваться правилом

$$d_j^k = \left(y_j^k\right)^2,\ h_i^k = \left(f_i^k\right)^2,\ j = 1, \ldots, n,\ i = 1, \ldots, m. \tag{13}$$

Другими эффективными, как показали результаты некоторых экспериментальных и теоретических исследований [6], вариантами правил нахождения весовых коэффициентов могут служить следующие

$$d_j^k = \frac{y_j^k}{\min\{\varepsilon, \alpha_j^{k-1}, \beta_j^{k-1}\}},\ h_i^k = \frac{-f_i^k}{\min\{\varepsilon, u_i^{k-1}\}},\ j = 1, \ldots, n,\ i = 1, \ldots, m,$$

при заданном $\varepsilon > 0$. Здесь u_i^{k-1}, α_j^{k-1}, β_j^{k-1} – приближения на итерации $k-1$ к множителям Лагранжа ограничений, $f_i(x) \le 0,\ i = 1, \ldots, m$, $x \le \bar{x}$, $\underline{x} \le x$.

Обозначим

$$D_k = diag(d^k),\ H_k = diag(h^k)$$

диагональные матрицы размера $n \times n$ и $m \times m$ с компонентами векторов d^k и h^k на главной диагонали.

Интервальный и аксиоматический подходы к определению весовых коэффициентов были введены в [5] применительно к алгоритмам внутренних точек для решения задач линейного программирования. Такой подход обобщает имеющиеся способы нахождения весовых коэффициентов и позволяет создавать новые. При этом не обязательно иметь конкретные выражения для функций $\bar{\sigma}$, $\underline{\sigma}$. Достаточно доказать существование таких функций и указанных их свойств.

Вспомогательная задача поиска направления улучшения решения

Вектор Δx^k определим как результат решения вспомогательной задачи минимизации квадратичной выпуклой функции от векторов $\Delta x \in R^n$ и $z \in R^m$

$$(c^k)^T \Delta x + \frac{1}{2}\Delta x^T B_k \Delta x + \frac{1}{2}\Delta x^T D_k^{-1} \Delta x + \frac{1}{2} z^T H_k^{-1} z \to \min, \qquad (14)$$

при линейных ограничениях-равенствах

$$A_k \Delta x - z = 0. \qquad (15)$$

В целевой функции задачи (14), (15) составляющие $\Delta x^T D_k^{-1} \Delta x$, $z^T H_k^{-1} z$ стимулируют при выборе Δx^k к движению «вдоль» границ множества векторов, удовлетворяющих неравенствам $\underline{x} \le x \le \bar{x}$, $f_i(x) \le 0$, $i = 1, \ldots, m$. Составляющая $\Delta x^T B_k \Delta x$ является квадратом взвешенной евклидовой нормы. Если матрица B_k вычисляется по правилу (8), то это слагаемое представляет вторые производные в аппроксимации функций $f_i(x)$, $i = 0, \ldots, m$.

Обозначим u^k вектор множителей Лагранжа этих ограничений. Определим вектор приближений к множителям Лагранжа нелинейных ограничений исходной задачи

$$v_i^k = \max\{0, u_i^k\},\ i = 1, \ldots, m. \tag{16}$$

Множество номеров i, при которых $u_i^k < 0$, обозначим I_k.

Два варианта вычисления направления корректировки

Вспомогательная задача (14), (15) сводится к решению системы линейных уравнений с симметрической положительно определенной матрицей. Возможны два варианта вычислений векторов Δx^k, z^k, u^k.

Один из них основывается на решении системы линейных уравнений с матрицей размера $n \times n$

$$(B_k + D_k^{-1} + A_k^T H_k^{-1} A_k)\Delta x = -c^k. \tag{17}$$

Найдя в результате решения системы (17) вектор Δx^k, определяем векторы $z^k = A_k \Delta x^k$,

$$u^k = H_k^{-1} z^k. \tag{18}$$

Система (17) следует из задачи безусловной минимизации

$$(c^k)^T \Delta x + \frac{1}{2}\Delta x^T B_k \Delta x + \frac{1}{2}\Delta x^T D_k^{-1} \Delta x + \frac{1}{2}\Delta x^T A_k{}^T H_k^{-1} A_k \Delta x \to \min,$$

к которой приходим в результате подстановки вектора z (из условия (15)) в целевую функцию (14).

Второй вариант вычислений основан на решении системы размера $m \times m$

$$(A_k (B_k + D_k^{-1})^{-1} A_k^T + H_k)u = -A_k (B_k + D_k^{-1})^{-1} c^k. \tag{19}$$

Вычислив u^k, находим вектор

$$\Delta x^k = -(B_k + D_k^{-1})^{-1}(A_k^T u^k + c^k) \tag{20}$$

и по формуле (15) – вектор z^k. Система (19) – результат решения задачи (14), (15) методом неопределенных множителей Лагранжа. Второй из двух вариантов вычислений предпочтительней в том случае, когда $m < n$.

Векторы $\alpha^k \in R^n$, $\beta^k \in R^n$, являющиеся приближениями на итерации k к множителям Лагранжа ограничений $x \le \bar{x}$, $\underline{x} \le x$, определим следующим образом

$$\alpha_j^k = \max\{0, -q_j^k\},\ \beta_j^k = \max\{0,\ q_j^k\},\ j = 1, \ldots, n. \qquad (21)$$

Вектор q^k вычисляется по правилу

$$q^k = -(B_k + D_k^{-1})\Delta x^k,$$

что, согласно (20), равносильно следующей формуле

$$q^k = A_k^T u^k + c^k.$$

Нахождение шага корректировки

Шаг корректировки вычисляется следующим образом:

1. $\lambda_1^k = \gamma \max\{\lambda : x^k + \lambda \Delta x^k \in X,\ \lambda \ge 0\}$, $\gamma -$ заданный параметр из интервала $(0, 1)$;

2. если $I_k \ne \phi$, то $\lambda_2^k = \max\{\lambda : f_i(x^k + \lambda \Delta x^k) \le f_i(x^k),\ i \in I_k,\ 0 \le \lambda \le \lambda_1^k\}$;

3. если $I_k = \phi$, то $\lambda_2^k = \lambda_1^k$;

4. $\lambda_k = \arg\min\{f_0(x^k + \lambda \Delta x^k) :\ 0 \le \lambda \le \lambda_2^k\}$.

Правила нахождения шага λ_k гарантируют, что вектор x^{k+1} будет находиться в $\operatorname{int} X$.

Критерий останова

Вычисления заканчиваются при выполнении с заданной точностью $\varepsilon > 0$ для $x^k \in X$, $v^k \ge 0$, $\alpha^k \ge 0$, $\beta^k \ge 0$ ограничений на множители Лагранжа

$$\left\| c^k + A^T v^k + \alpha^k - \beta^k \right\| \le \varepsilon,$$

и условий дополняющей нежесткости (4) – (6)

$$-v_i^k f_i(x^k) \le \varepsilon,\ i = 1, \ldots, m,$$

$$\beta_j^k (x_j^k - \underline{x}_j) \le \varepsilon,\ j = 1, \ldots, n,$$

$$\alpha_j^k(\bar{x}_j - x_j^k) \le \varepsilon,\ j = 1, \ldots, n.$$

Здесь $\|r\| = \left(\sum_{i=1}^{n} r_i\right)^{1/2}$ евклидова (невзвешенная) норма вектора $r \in R^n$, векторы v^k, α^k, β^k находятся по правилам (16), (21).

Свойства вырабатываемых приближений

В силу (17) и положительной определенности матрицы $B_k + D_k^{-1} + A_k^T H_k^{-1} A_k$, для всех k, $k = 0, 1, 2, \ldots$, имеет место неравенство

$$(c^k)^T \Delta x^k < 0. \qquad (22)$$

Из выполнения неравенства (22) и правил определения шага λ_k следует, что

$$f_0(x^{k+1}) < f_0(x^k). \qquad (23)$$

Поскольку целевая функция задачи (1), (2) ограничена снизу на множестве X, то, в силу (23), значения целевой функции сходятся к некоторой величине

$$F = \lim_{k \to \infty} f_0(x^k). \qquad (24)$$

Евклидовы проекции начала координат на линейное многообразие

В применяемой здесь технике доказательства используется установленный в [5] факт ограниченности множества евклидовых проекций начала координат на линейное многообразие. Сначала поясним этот факт.

Пусть L линейное многообразие в R^n. Вектор $s \in L$ назовем опорным вектором многообразия L, если не существует вектора $y \in L$ такого, что $J(y) \subset J(s)$, $J(y) \ne J(s)$, где $J(y) = \{j : y_j \ne 0\}$ – множество номеров ненулевых компонент вектора y. Обозначим $Q(L)$ выпуклую оболочку опорных векторов линейного многообразия L. Поскольку число опорных векторов у линейного многообразия L конечно, то их выпуклая оболочка $Q(L)$ будет ограниченным множеством.

Обозначим

$$r(p) = \arg\min\left\{\sum_{j=1}^{n} p_j (r_j)^2 :\ r \in L\right\}$$

– евклидову проекцию начала координат на линейное многообразие L при использовании евклидовой нормы

$$\|r\|_p = \left(\sum_{j=1}^{n} p_j (r_j)^2 \right)^{1/2}$$

с заданными весовыми коэффициентами $p_j > 0,\ j = 1, \ldots, n$. В [5] доказано следующее утверждение

Лемма 1. *Любая евклидова проекция начала координат R^n на линейное многообразие $L \subset R^n$ находится в выпуклой оболочке опорных векторов этого многообразия, $r(p) \in Q(L)$ при любых $p_j > 0,\ j = 1, \ldots, n$.*

Этот факт служит основой для теоретического обоснования алгоритмов внутренних точек задач линейного программирования [5, 6]. Усложняющим обстоятельством для нелинейных задач является изменение линейных многообразий, рассматриваемых при доказательстве, в зависимости от итеративно меняющейся точки, в которой осуществляется линеаризация. При этом меняются также опорные векторы.

Выделим два линейных многообразия, соответствующих вектору x такому, что $\underline{x} \le x \le \bar{x}$. Линейное многообразие $L_1(x)$ состоит из векторов $\Delta x \in R^n,\ z \in R^m$, удовлетворяющих условиям

$$A(x)\Delta x - z = 0, \tag{25}$$

$$c^T \Delta x = -1. \tag{26}$$

Здесь $A(x)$ – матрица размера $m \times n$, строками которой являются градиенты функций $f_i(x),\ i = 1, \ldots, m$.

Линейное многообразие $L_2(x)$ состоит из векторов $q \in R^n,\ u \in R^m$, удовлетворяющих условию

$$q - \left(A(x)\right)^T u = c. \tag{27}$$

Теоретическое обоснование алгоритмов внутренних точек

Для доказательства сходимости последовательности приближений, вырабатываемых вычислительным процессом (7), к оптимальному решению задачи (1), (2) введем некоторые предположения.

1. Будем считать, что целевая функция линейна: при заданном $c \in R^n$, $c \neq 0$

$$f_0(x) = c^T x .$$

Отметим, что это предположение вводится для удобства доказательства и не является существенным.

2. С изменением вектора x могут меняться значения опорных векторов линейных многообразий $L_1(x)$, $L_2(x)$. Предположим, что объединения по x, удовлетворяющим условию $\underline{x} \le x \le \overline{x}$, множеств опорных векторов многообразий $L_1(x)$ и объединения по x, удовлетворяющим $\underline{x} \le x \le \overline{x}$, множеств опорных векторов многообразий $L_2(x)$ будут ограниченными множествами.

3. Будем считать, что величина шага λ_k на всех итерациях ограничена снизу некоторым положительным числом τ, т.е. существует $\tau > 0$ такое, что

$$\lambda_k \ge \tau , \; k = 0, 1, 2, \ldots . \qquad (28)$$

В случае линейных ограничений это условие обязательно выполняется, поэтому нет необходимости вводить его особо. При наличии нелинейных ограничений выполнение этого условия – открытый вопрос. Поэтому оно, к сожалению авторов, вводится в условия теоремы.

Для доказательства сходимости векторов x^k к оптимальному решению задачи (1), (2) потребуется следующее вспомогательное утверждение, доказываемое стандартным образом.

Лемма 2. *Если положительный ряд* $\sum_{k=0}^{\infty} \alpha_k$ *сходится, последовательность чисел* β_k, $k = 0, 1, 2, \ldots$, *ограниченная, то ряд* $\sum_{k=0}^{\infty} \alpha_k \beta_k$ *сходится.*

С учетом вышеперечисленных фактов перейдем к доказательству следующего утверждения.

Теорема 1. *Пусть*

а) задача (1), (2) *невырождена;*

б) целевая функция линейна;

в) объединения по всем x, *удовлетворяющим условию* $\underline{x} \le x \le \bar{x}$, *множеств опорных векторов линейных многообразий* $L_1(x)$, $L_2(x)$ *ограничены;*

г) существует $\tau > 0$ *такое, что* $\lambda_k \ge \tau$, $k = 1, 2, \ldots$.

Тогда существуют векторы $\tilde{x} \in X$, $\tilde{u} \in R^m$, $\tilde{\alpha} \in R^n$, $\tilde{\beta} \in R^n$ *такие, что*

$$x^k \to \tilde{x},\ u^k \to \tilde{u},\ \alpha^k \to \tilde{\alpha},\ \beta^k \to \tilde{\beta} \text{ при } k \to \infty.$$

Вектор $\tilde{x}$ *является оптимальным решением, а векторы* $\tilde{u}$, $\tilde{\alpha}$, $\tilde{\beta}$ – *множителями Лагранжа ограничений задачи* (1), (2).

Доказательство. Будем нумеровать отдельные фрагменты рассуждений.

1. Вычислительный процесс (7) равносилен следующему

$$x^{k+1} = x^k + \tilde{\lambda}_k \Delta\tilde{x}^k, \tag{29}$$

где

$$\Delta\tilde{x}^k = \frac{-1}{c^T \Delta x^k} \Delta x^k,\ \tilde{\lambda}_k = -\left(c^T \Delta x^k\right)\lambda_k. \tag{30}$$

Положим

$$\tilde{z}^k = \frac{-1}{c^T \Delta x^k} z^k.$$

Из вспомогательной задачи (14), (15) вытекает, что векторы $\Delta\tilde{x}^k$, $\tilde{z}^k$ являются решением задачи:

$$\frac{1}{2}\Delta x^T (B_k + D_k^{-1})\Delta x + \frac{1}{2} z^T H_k^{-1} z \to \min,$$

при ограничениях

$$A_k \Delta x - z = 0,$$

$$c^T \Delta x = -1.$$

Т.е. векторы $\Delta\tilde{x}^k$, $\tilde{z}^k$ являются евклидовыми проекциями начала координат на линейное многообразие $L_1(x^k)$.

Векторы q^k, u^k определяются как результат решения следующей задачи нахождения евклидовой проекции начала координат на линейное многообразие $L_2(x^k)$

$$\frac{1}{2}q^T(B_k + D_k^{-1})^{-1}q + \frac{1}{2}u^T H_k u \to \min,$$

$$q - A_k^T u = c.$$

Из леммы 1 и условия ограниченности объединений по x^k, $k = 0, 1, 2, \ldots$, множеств опорных векторов линейных многообразий $L_1(x)$, $L_2(x)$ следует, что векторы $\Delta\tilde{x}^k$, $\tilde{z}^k$, q^k, u^k будут образовывать ограниченные последовательности: при некоторых $M_1 > 0$, $M_2 > 0$, для всех $k = 0, 1, 2, \ldots$,

$$\left\| \Delta\tilde{x}^k \right\| \le M_1, \left\| \tilde{z}^k \right\| \le M_1, \tag{31}$$

$$\left\| q^k \right\| \le M_2, \left\| u^k \right\| \le M_2, \tag{32}$$

В силу (29), (30) для $k = 0, 1, 2, \ldots$,

$$c^T x^k - c^T x^{k+1} = \tilde{\lambda}_k, \ \tilde{\lambda}_k > 0.$$

Из (24) следует, что положительный ряд $\sum_{k=0}^{\infty} \tilde{\lambda}_k$ является сходящимся

$$\sum_{k=0}^{\infty} \tilde{\lambda}_k = c^T x^0 - F. \tag{33}$$

Из (29), (31), (33) и леммы 2, примененной к процессу (29), получаем, что при некотором $\tilde{x} \in R^n$

$$x^k \to \tilde{x} \text{ при } k \to \infty. \tag{34}$$

При этом $\tilde{x} \in X$, так как на всех итерациях $x^k \in X$ и множество X замкнуто.

Отметим, что в полученном обосновании сходимости векторов x^k условия на выбор весовых коэффициентов не использовались в полной мере.

Использовался только факт положительности значений весовых коэффициентов d_j^k, h_i^k, $j = 1, \ldots, n$, $i = 1, \ldots, m$.

2. Из (22) и (24) следует

$$c^T x^{k+1} - c^T x^k \to 0 \text{ при } k \to \infty.$$

В силу (7)

$$c^T x^{k+1} - c^T x^k = \lambda_k (c^T \Delta x^k).$$

Поэтому

$$\lambda_k (c^T \Delta x^k) \to 0 \text{ при } k \to \infty.$$

Из (28) получаем, что

$$c^T \Delta x^k \to 0 \text{ при } k \to \infty. \qquad (35)$$

Производная в точке оптимума минимизируемой в (14) функции должна быть равна нулю по любому направлению из R^n, в том числе и по направлению Δx^k,

$$-c^T \Delta x^k = (\Delta x^k)^T (B_k + D_k^{-1}) \Delta x^k + (\Delta x^k)^T A_k^T H_k^{-1} A_k \Delta x^k.$$

Из (15) и (35) следует, что

$$(\Delta x^k)^T (B_k + D_k^{-1}) \Delta x^k + (z^k)^T H_k^{-1} z^k \to 0 \text{ при } k \to \infty. \qquad (36)$$

Согласно (18), (20), (27), (36) при $k \to \infty$

$$(q^k)^T (B_k + D_k^{-1})^{-1} q^k \to 0, \qquad (37)$$

$$(u^k)^T H_k u^k \to 0. \qquad (38)$$

Из (32) следует, что предельные при $k \to \infty$ значения последовательностей векторов u^k, q^k образуют ограниченные множества, которые обозначим $\mathrm{Lim} u^k$, $\mathrm{Lim} q^k$.

В силу (9) – (12) последовательности векторов d^k, h^k являются ограниченными и, следовательно, имеют ограниченные множества предельных при $k \to \infty$ значений, которые обозначим $\mathrm{Lim} d^k$, $\mathrm{Lim} h^k$.

Для любого $\tilde{u} \in \mathrm{Lim} u^k$ существует, согласно (38), вектор $\tilde{h} \in \mathrm{Lim} h^k$ такой, что

если $\widetilde{u}_i \neq 0$, то $\widetilde{h}_i = 0$, если $\widetilde{h}_i \neq 0$, то $\widetilde{u}_i = 0$, $i = 1, \ldots, m$.

Аналогично, в силу (37), для любого $\widetilde{q} \in \operatorname{Lim} q^k$ существует вектор $\widetilde{d} \in \operatorname{Lim} d^k$ такой, что

если $\widetilde{q}_j \neq 0$, то $\widetilde{d}_j = 0$, если $\widetilde{d}_j \neq 0$, то $\widetilde{q}_j = 0$, $j = 1, \ldots, n$.

Учитывая условия на выбор весовых коэффициентов (9) – (12), получаем, что для вектора $\widetilde{x} \in X$ при любых векторах $\widetilde{u} \in \operatorname{Lim} u^k$, $\widetilde{q} \in \operatorname{Lim} q^k$ выполняются условия стационарности.

Поскольку задача (1), (2) невырожденная, то для стационарного решения $\widetilde{x} \in X$ векторы q и u, удовлетворяющие условиям (37) и (38) соответственно, единственны. Поэтому последовательность векторов u^k будет сходиться к одному вектору $\widetilde{u} = \lim\limits_{k \to \infty} u^k$. Соответственно, последовательность векторов q^k будет сходиться к единственному вектору $\widetilde{q} = \lim\limits_{k \to \infty} q^k$.

3. Предположим, что у вектора $\widetilde{u}$ есть отрицательные компоненты. Если $\widetilde{u}_i < 0$ для некоторого i, то начиная с некоторой итерации k_0, на всех последующих $u_i^k < 0$. Согласно (18), $z_i^k = h_i^k u_i^k$ и $h_i^k > 0$, то на всех последующих итерациях $z_i^k < 0$. Из (15) следует, что

$$(\nabla f_i(x^k))^T \Delta x^k < 0. \qquad (39)$$

В силу (39) и условий на выбор шага λ_k, для всех $k \geq k_0$ имеет место неравенство $f_i(x^k + \lambda_k \Delta x^k) < f_i(x^k)$.

Получаем, что значение $f_i(x^k)$ для данного i будет убывать по итерациям. Поэтому $\widetilde{h}_i > 0$. Неравенства $\widetilde{h}_i > 0$ и $\widetilde{u}_i < 0$ противоречат условию (38). Тем самым доказано, что $\widetilde{u} \geq 0$.

Если для некоторого i $\widetilde{u}_i > 0$, то, начиная с некоторой итерации k_0, для всех $k \geq k_0$ значение $f_i(x^k)$ для данного i будет возрастать. Следовательно,

значение $f_i(\tilde{x})$ либо меньше нуля, либо равно нулю. В силу (38), $\tilde{h}_i = 0$. Отсюда, согласно (10), следует, что $\underline{\sigma}(-f_i(\tilde{x})) = 0$. Это означает, что $f_i(\tilde{x}) = 0$. Аналогично, если для некоторого i $f_i(\tilde{x}) < 0$, то $\tilde{h}_i > 0$. Следовательно, $\tilde{u}_i = 0$. Тем самым доказано выполнение условий дополняющей нежесткости (3) для векторов $\tilde{x}$, $\tilde{u}$.

Неотрицательность компонент векторов $\tilde{\alpha}$, $\tilde{\beta}$ следует из правила (21). Покажем, что для векторов $\tilde{x}$, $\tilde{\alpha}$, $\tilde{\beta}$ выполняются условия дополняющей нежесткости (4), (5).

Множества номеров положительных, отрицательных и нулевых компонент $\tilde{q}$ обозначим Q_+, Q_-, Q_0, соответственно.

Пусть $j \in Q_+$, тогда $\tilde{q}_j > 0$ и $\tilde{\alpha}_j = 0$, $\tilde{\beta}_j = \tilde{q}_j$. Поскольку $\tilde{q}_j > 0$, то начиная с некоторой итерации $\hat{k}$, на всех последующих $q_j^k > 0$. Согласно (20),

$$\Delta x_j^k = -\frac{d_j^k}{b_j^k d_j^k + 1} q_j^k. \tag{40}$$

Поскольку $q_j^k > 0$, $d_j^k > 0$, $b_j^k \geq 0$, то, в силу (40), $\Delta x_j^k < 0$ для всех $k \geq \hat{k}$. Из (7) и неравенства $\lambda_k > 0$ получаем, что величина x_j^k для данного j будет убывать по итерациям. Поэтому

$$\bar{x}_j - \tilde{x}_j > 0. \tag{41}$$

Из (37) и положительности q_j^k следует, что для данного номера j последовательность величин d_j^k сходится к нулю при $k \to \infty$. Это означает, в силу определения весовых коэффициентов d_j^k и (41), что $(x_j^k - \underline{x}_j) \to 0$ при $k \to \infty$. Отсюда следует, что условия (4), (5) выполняются для $j \in Q_+$.

Пусть $j \in Q_-$, тогда $\tilde{q}_j < 0$ и $\tilde{\alpha}_j = \tilde{q}_j$, $\tilde{\beta}_j = 0$. Если $\tilde{q}_j < 0$, то начиная с некоторой итерации $\hat{k}$, на всех последующих $q_j^k < 0$. Поскольку $q_j^k < 0$,

$d_j^k > 0$, $b_j^k \geq 0$, то, согласно (40), $\Delta x_j^k > 0$ для всех $k \geq \hat{k}$. Из (7) и неравенства $\lambda_k > 0$ получаем, что величина x_j^k для данного j будет возрастать по итерациям. Поэтому

$$\widetilde{x}_j - \underline{x}_j > 0. \qquad (42)$$

Из (37) и отрицательности q_j^k следует, что для данного номера j последовательность величин d_j^k сходится к нулю при $k \to \infty$. Это означает, в силу определения весовых коэффициентов d_j^k и (42), что $(\overline{x}_j - x_j^k) \to 0$ при $k \to \infty$. Отсюда следует, что для $j \in Q_-$ выполняются условия (4), (5).

Пусть $j \in Q_0$, тогда $\widetilde{q}_j = 0$ и, следовательно, $\widetilde{\alpha}_j = 0$, $\widetilde{\beta}_j = 0$. Это означает, что для $j \in Q_0$ имеют место (4), (5).

Итак, доказано, что для векторов $\widetilde{x}$, $\widetilde{u}$, $\widetilde{\alpha}$, $\widetilde{\beta}$ выполняются условия дополняющей нежесткости (3) – (5). Теорема доказана.

Заключение

Представлено описание нового семейства алгоритмов внутренних точек для решения задач оптимизации с нелинейными ограничениями. Основной вычислительной проблемой в алгоритмах является нахождение на каждой итерации направления корректировки решения. Эта задача представляется в виде задачи безусловной минимизации квадратичной выпуклой функции от n или m переменных. Такая альтернатива может быть полезна в вычислительном отношении.

Осуществлено теоретическое обоснование сходимости последовательности приближений, вырабатываемых алгоритмами внутренних точек из описанного класса, к стационарному и оптимальному решению задачи (1), (2). Весовые коэффициенты в этих алгоритмах удовлетворяют условиям (9) – (12). Отметим, что для доказательства сходимости последовательности приближений, вырабатываемых вычислительным процессом (7), имеющиеся стандартные техники

доказательства, описанные, например, у У.И. Зангвилла [7], не годятся. Это связано с тем, что при поиске направления улучшения решения используются изменяющиеся по итерациям нормы. Доказательство теоремы 1 является некоторым переложением техники доказательства сходимости, предложенной в [5] для теоретического обоснования алгоритмов внутренних точек в линейном программировании.

Список литературы

[1] Дикин И.И. Итеративное решение задач линейного и квадратичного программирования // Доклады АН СССР, 1967, том 174. – С.747–748.

[2] Дикин И.И., Зоркальцев В.И. Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). – Новосибирск: Наука, 1980. – 144с.

[3] Евтушенко Ю.Г., Жадан В.Г. Численные методы решения некоторых задач исследования операций. // Журнал вычислительной математики и математической физики.– 1973. – Т. 13. – № 3. – С. 583–597.

[4] Евтушенко Ю.Г., Жадан В.Г. Барьерно-проективные методы решения задач нелинейного программирования // Журнал вычислительной математики и математической физики.– 1994. –Т. 34. –№ 5. – С. 669–684.

[5] Зоркальцев В.И. Относительно внутренняя точка оптимальных решений. Сыктывкар: Коми фил. АН СССР, 1984. – 48с.

[6] Зоркальцев В.И. Метод относительно внутренних точек. – Сыктывкар: Коми филиал АН СССР, 1986. – 18с.

[7] Зангвилл У.И. Нелинейное программирование. – М.: Советское радио, 1973. – 312с.

[8] Пержабинский С.М. Алгоритм внутренних точек, использующий квадратичные аппроксимации. // Современные технологии. Системный анализ. Моделирование. – Иркутск: ИрГУПС. – 2008. – №3(19). – С.97–101.

УДК 519.856

СВЕДЕНИЕ ЗАДАЧ ДВУХЭТАПНОЙ ВЕРОЯТНОСТНОЙ ОПТИМИЗАЦИИ С ДИСКРЕТНЫМ РАСПРЕДЕЛЕНИЕМ СЛУЧАЙНЫХ ДАННЫХ К ЗАДАЧАМ ЧАСТИЧНО ЦЕЛОЧИСЛЕННОГО ПРОГРАММИРОВАНИЯ[1]

Кибзун А. И.

Московский авиационный институт

E-mail: kibzun@mail.ru

Наумов А. В.

Московский авиационный институт

E-mail: naumovav@mail.ru

Норкин В. И.

Институт кибернетики В.М. Глушкова НАН Украины

E-mail: norkin@ipnet.kiev.ua

Аннотация. В работе рассматриваются модели двухэтапного стохастического программирования с квантильным критерием, а также модели с вероятностным ограничением на случайные значения целевой функции второго этапа. Такие модели позволяют формализовать требования к надежности и безопасности оптимизируемой системы, а также оптимизировать функционирование системы в экстремальных условиях. Предлагается способ эквивалентного преобразования таких моделей при дискретном распределении случайных параметров к задачам частично целочисленного программирования большой размерности. Число дополнительных целочисленных (булевых) переменных в этой задаче равно числу возможных значений вектора случайных параметров. Полученные смешанные задачи предполагается решать мощными стандартными компьютерными программами дискретной оптимизации. В качестве иллюстрации подхода приводятся результаты численных экспериментов на задаче небольшой размерности.

Ключевые слова. Стохастическое программирование, двухэтапные задачи, квантильное программирование, вероятностные ограничения, детерминированный эквивалент, частично целочисленные задачи оптимизации, дискретное программирование.

[1] Работа выполнена при частичной поддержке Государственного фонда фундаментальных исследований Украины в рамках совместного российско-украинского проекта Ф40.1/016 (2011-2012) и Российского фонда фундаментальных исследований (проекты 11-07-90407-Укр-ф-а, 11-07-00315-а).

Введение

Традиционно в задачах стохастического программирования оптимизируется среднее значение показателя качества управления, зависящего от случайного параметра [1, 2]. Наряду со средним значением можно и нужно использовать другие показатели качества управления, например, медиану или другие квантили распределения [3 – 5]. Такие критерии качества используются, например, в задачах управления летательными аппаратами [3]. Однако задача оптимизации функции квантили является более сложной, чем оптимизация среднего значения, т.к. функция квантили может быть невыпуклой и разрывной. Для приближенной оптимизации квантили часто используют оптимизацию ее верхней оценки, полученную доверительным методом [3 – 5], [7], или с помощью CVaR (conditional value at risk) функций [11]. Другие методы приближенной и глобальной оптимизации функции квантили (применительно к задачам портфельной оптимизации) указаны в Wozabal и др. [12].

В работах [13, 14] некоторые задачи квантильной оптимизации с дискретным распределением случайных параметров сведены к задачам частично целочисленного программирования. Наиболее общие результаты в этом направлении получены в статье [6]. Ранее прием сведения задач с вероятностными ограничениями, тесно связанных с задачей оптимизации функции квантили, к задачам частично целочисленного программирования применялся в работах [14 – 18], [20].

Двухэтапные модели стохастического программирования с целевой функцией в виде математического ожидания являются одними из основных в стохастическом программировании [1, 2], [21, 22]. Двухэтапная модель стохастического программирования моделирует ситуации принятия решений в условиях неопределенности, когда вначале (на первом этапе) в условиях стохастической неопределенности принимается детерминированное решение в расчете на то, что когда ситуация прояснится (на втором этапе), будет принято дополнительное

оптимальное корректирующее решение. Решение первого этапа оптимизируется по сумме математического ожидания целевых функций первого и второго этапов.

С другой стороны двухэтапные модели стохастического программирования являются частным случаем двухуровневых кооперативных моделей принятия оптимальных решений, в которых целевая функция верхнего уровня аддитивно включает среднее значение целевой функции нижнего уровня.

Например, применительно к оптимизации энергетических систем задачи двухэтапного стохастического программирования можно интерпретировать следующим образом. Переменные первого этапа могут описывать функционирование длительно действующих мощных энергетических установок, стоимости запуска и остановки которых весьма велики. Переменные второго этапа могут описывать производство энергии мобильными установками, запуск и остановка которых относительно дешевы, и которые могут сравнительно легко адаптироваться к пиковым нагрузкам. В качестве целевой функции задачи (при постоянных ценах на энергию) может выступать средняя стоимость производства энергии на всех видах установок. Целевая функция и переменные второго этапа могут описывать количество произведенной энергии на мобильных установках. В постановку задачи могут быть включены ограничения снизу на вероятность обеспечения пиковых потребностей в энергии. Для построения стохастической модели энергетической системы необходимы вероятностные сценарии потребности в энергии.

Вместо функции математического ожидания в задаче стохастического программирования иногда имеет смысл использовать функцию квантили случайного оптимального значения задачи. Это позволяет оптимизировать функционирование стохастической системы в экстремальных условиях. В работах [7 – 10], [23] двухэтапные задачи стохастического программирования с квантиль-

ной целевой функцией решались с помощью доверительного метода из работ [4, 5].

В настоящей статье мы сводим задачи квантильного двухэтапного стохастического программирования, а также задачи двухэтапного стохастического программирования с вероятностными ограничениями при дискретном распределении случайных данных к задачам частично целочисленного программирования и, таким образом, распространим результаты из [6] на общие двухэтапные задачи квантильной и вероятностной оптимизации с дискретным распределением. Для решения возникающих при этом частично целочисленных задач используется современный стандартный пакет IBM ILOG CPLEX [24] решения частично целочисленных задач линейного программирования.

Достоинства данного подхода к решению вероятностных и квантильных задач стохастического программирования заключаются в следующем:

- функции вероятности и квантили позволяют формализовать требования к надежности и безопасности оптимизируемой системы, а также оптимизировать функционирование системы в экстремальных условиях;
- подход допускает наличие как непрерывных, так и дискретных переменных в первоначальной формулировке задачи;
- проблема невыпуклости и разрывности вероятностных и квантильных функций задачи перекладывается на методы дискретной оптимизации;
- в случае непрерывного распределения случайных данных его всегда можно аппроксимировать дискретным, например, эмпирическими распределением с достаточно большим множеством реализаций;
- полученные эквивалентные задачи частично целочисленного программирования можно решать мощными современными программными средствами дискретной оптимизации даже при весьма большом количестве переменных задачи и возможных сценариев для случайных данных.

Обозначения и предварительные сведения

Пусть (Ω, Σ, P) – некоторое вероятностное пространство, $X : \Omega \to X(\Omega) \subseteq R^m$ – векторная случайная величина со значениями в множестве $X(\Omega)$, $U \subseteq R^n$ – множество допустимых стратегий оптимизации, $\Phi : U \times X(\Omega) \to R^1$ и $Q : U \times X(\Omega) \to R^1$ – функции, борелевские по $x \in X(\Omega)$ функции для всех $u \in U$. Определим функции математического ожидания

$$f_1(u) \overset{def}{=} \mathrm{M}\,[\Phi(u, X)] = \int_{\Omega} \Phi(u, X(\omega)) d\mathrm{P}(\omega),$$

вероятности

$$P_{\varphi}(u) \overset{def}{=} \mathrm{P}\,\{\Phi(u, X) \le \varphi, Q(u, X) \le 0\},$$

и квантили

$$\varphi_{\alpha}(u) \overset{def}{=} \inf\,\{\varphi \in R^1 : P_{\varphi}(u) \ge \alpha\} = \min\{\varphi \in R^1 : \mathrm{P}\,\{\Phi(u, X) \le \varphi, Q(u, X) \le 0\} \ge \alpha\}, \quad (1)$$

где α – параметр, $0 < \alpha < P^*$, $\mathrm{P}\,\{\cdot\}$ обозначает вероятность события в скобках (по определению, $\mathrm{P}\,\{\varnothing\} = 0$), M – знак математического ожидания,

$$P^* \overset{def}{=} \sup_{u \in U} \mathrm{P}\,\{Q(u, X) \le 0\}.$$

Заметим, что функция $P_{\varphi}(\cdot)$ определена на всем множестве U. Свойства функций вероятности детально изучены в [2], [4, 5], [25, 26], а квантили – в [4, 5], [27]. Так если $\Phi(u, x), Q(u, x)$ полунепрерывны снизу по u при каждом x, то функция $P_{\varphi}(u)$ полунепрерывна сверху по совокупности переменных (u, φ) [28]. Кроме того, функция $P_{\varphi}(u)$ монотонна (не убывает) по φ, и непрерывна справа, поэтому инфинум в определении $\varphi_{\alpha}(u)$ достигается. Функция квантили является специальным случаем маргинальной функции (функции минимума),

поэтому при сделанных предположениях она полунепрерывна снизу по (u,α) [29].

В силу непрерывности вероятностной меры имеет место

$$\lim_{\varphi\to+\infty} \mathrm{P}\{\Phi(u,X)\le\varphi,\ Q(u,X)\le 0\}=\mathrm{P}\{Q(u,X)\le 0\}, \tag{2}$$

$$\lim_{\varphi\to+\infty} \mathrm{P}\{\Phi(u,X)\le\varphi,\ Q(u,X)\le 0\}=0, \tag{3}$$

Задача оптимизации функции квантили

Простейшая задача (одноэтапного) стохастического программирования [1, 2] имеет вид:

$$f_1(u) \overset{def}{=} \mathrm{M}\,[\Phi(u,X)]\to\inf_{u\in U}\Phi(u,x)$$

в которой минимизируется среднее значение случайного показателя $\Phi(u,X)$ на множестве U значений детерминированного параметра u. Вместо среднего значения в качестве целевой функции можно использовать медиану распределения случайной величины $\Phi(u,X)$ или другую ее квантиль уровня α, $0<\alpha<P^*$. Задача минимизации функции квантили [4], [27], [30] имеет вид

$$\varphi_\alpha(u) \overset{def}{=} \min\{\{\varphi\in R^1 : \mathrm{P}\{\Phi(u,X)\le\varphi, Q(u,X)\le 0\}\ge\alpha\}\}\to\inf_{u\in U}. \tag{4}$$

Известно, что она эквивалентна следующей задаче [2]:

$$\begin{aligned} &\varphi\to\inf_{\varphi\in R^1,\,u\in u},\\ &P_\varphi(u)=\mathrm{P}\{\Phi(u,X)-\varphi\le 0, Q(u,X)\le 0\}\ge\alpha. \end{aligned} \tag{5}$$

В дальнейшем мы будем рассматривать и другие эквивалентные оптимизационные задачи, при этом эквивалентность будет пониматься в следующем смысле. Рассмотрим общую задачу математического программирования.

Определение 1. *Под задачей математического программирования будем понимать задачу минимизации целевой функции* $\Phi: U\to R^1$, *определенной на неко-*

тором допустимом множестве U точек (стратегий и т.п.), в формальной записи

$$\Phi(u) \to \inf_{u \in U} \tag{6}$$

Элементы $u \in U$ будем называть допустимыми решениями задачи. Множество U может быть пустым, тогда будем говорить, что задача не имеет допустимых решений.

Определение 2. *Нижнюю грань (конечную или бесконечную) $\Phi(u)$ на U,*

$$\varphi^* = \inf_{u \in U} \Phi(u)$$

будем называть оптимальным значением целевой функции задачи (6)*. Если $\varphi^* > -\infty$ и существует допустимая точка $u^* \in U$ такая, что $\varphi^* = \Phi(u^*)$, то будем говорить, что оптимальное значение задачи* (6) *достигается, а точку u^* будем называть оптимальным решением задачи. В этом случае будем также писать $\varphi^* = \min \Phi(u^*)$. В противном случае, т.е. если $\varphi^* = -\infty$ или не существует точки $u^* \in U$ такой, что $\varphi^* = \Phi(u^*)$, то будем говорить, что оптимальное значение задачи не достигается.*

Определение 3. *Две задачи оптимизации вида (6) эквивалентны, если выполнены все следующие условия:*

а) либо обе эти задачи имеют допустимые решения (с конечными значениями целевых функций), либо обе не имеют таких решений;

б) если эти задачи имеют допустимые решения, то оптимальные значения их целевых функций (конечные или бесконечные) совпадают;

в) если оптимальные значения их целевых функций конечны, то эти значения в обеих задачах либо достигаются, либо не достигаются;

г) если оптимальные значения достигаются, то по оптимальному решению одной задачи с помощью явно указанного алгоритма за конечное число шагов восстанавливается оптимальное решение другой.

Заметим, что если оптимальное решение задачи u^* существует, то это означает, что оптимальное значение $\varphi^* = \Phi(u^*)$ критериальной функции достигается.

Отметим, что введенное отношение эквивалентности оптимизационных задач, очевидно, транзитивно.

Доказательство эквивалентности оптимизационных задач будет основано на установлении специального соответствия между множествами допустимых стратегий задач, а именно, для каждой допустимой стратегии одной задачи мы указываем допустимую стратегию другой с таким же или меньшим значением целевой функции.

Лемма 1. *Две оптимизационные задачи*

$$\Phi_1(u) \to \inf_{u \in U_1} \text{ и } \Phi_2(u) \to \inf_{u \in U_2} \tag{7}$$

эквивалентны в смысле определения 3*, если известны алгоритмы (отображения)* $A_1 : U_1 \to U_2$ *и* $A_2 : U_2 \to U_1$*, которые для каждой допустимой стратегии одной задачи указывают допустимую стратегию другой задачи с таким же или меньшим значением целевой функции, т.е. для любых* $u_1 \in U_1$ *и* $u_2 \in U_2$ *выполнено* $\Phi_2(A_1(u_1)) \le \Phi_1(u_1)$ *и* $\Phi_1(A_2(u_2)) \le \Phi_2(u_2)$.

При этом, если u_1^* *– оптимальное решение первой задачи, то* $A_1(u^*_1)$ *– оптимальное решение второй задачи и* $\Phi_1(u^*_1) = \Phi_2(A_1(u^*_1))$. *Наоборот, если* u_2^* *– оптимальное решение второй задачи, то* $A_2(u^*_2)$ *– оптимальное решение первой задачи и* $\Phi_2(u^*_2) = \Phi_1(A_2(u^*_2))$.

Доказательство леммы приведено в [6]. Подобный лемме 1 способ доказательства эквивалентности задач применялся, например, в [13, 14], [31, 32]. Таким образом, для доказательства эквивалентности, вообще говоря, нет необходимости предполагать существование решения исходной задачи, его существование или несуществование может быть установлено в ходе решения эквивалентной задачи. Если известно, что одна из задач имеет оптимальное решение, то достаточные условия эквивалентности оптимизационных задач из леммы 1 могут быть несколько ослаблены.

Лемма 2. *Предположим, что одна из задач* (7) *(первая) имеет оптимальное решение* $u_1^* \in U_1$ *и существует допустимая точка* $u_2' \in U_2$ *другой задачи такая, что* $\Phi_2(u_2') \le \Phi_1(u^*_1)$. *Предположим также, что известен алгоритм (отображение)* $A_2 : U_2 \to U_1$ *который для каждой допустимой точки второй задачи указывает допустимую точку первой задачи с таким же или меньшим значением целевой функции, т.е. для любого* $u_2 \in U_2$ *выполнено* $\Phi_1(A_2(u_2)) \le \Phi_2(u_2)$. *Тогда и другая задача имеет оптимальные решения, причем* $u_2' \in U_2$ *является одним из таких оптимальных решений, а* $A_2(u_2')$ *является оптимальным решением первой задачи, и оптимальные значения обеих задач совпадают. Таким образом, рассматриваемые оптимизационные задачи эквивалентны.*

Доказательство. В силу оптимальности точки u_1^* имеем

$$\Phi_1(u_1^*) \ge \Phi_2(u_2') \ge \Phi_1(A(u_1^*)) \ge \Phi_1(u_1^*),$$

т.е. $\Phi_1(u_1^*) = \Phi_1(A(u_1^*)) = \Phi_2(u_2')$ и, таким образом, $A_2(u_2')$ - оптимальное решение первой задачи. Покажем, что u_2' является оптимальным решением второй задачи. Предположим противное. Тогда существует другая допустимая точка $u_2'' \in U_2$ с меньшим значением целевой функции, $\Phi_2(u_2'') < \Phi_2(u_2')$. Но по пред-

положению существует допустимая точка первой задачи $A_2(u_2'') \in U_1$ такая, что $\Phi_1(A(u_2'')) \le \Phi_2(u_2'') < \Phi_2(u_2') \le \Phi_1(u_1^*)$, что противоречит оптимальности точки u_1^*. Лемма доказана.

Предположение А. *Случайная величина X принимает конечное число значений, т.е. $X(\Omega) = \{x_1, x_2, ..., x_K\}$, с вероятностями $p_1 > 0, p_2 > 0, ..., p_K > 0$, $\sum_{k=1}^{K} p_k = 1$.*

Предположение Б. *Заданы функции $\mu_1(u,x)$ и $\mu_2(u,x)$ такие, что для любых $u \in U$, $x \in X(\Omega)$ выполнено*

$$-\infty < \mu_1(u,x) \le \inf_{x \in X(\Omega)} \Phi(u,x), \tag{8}$$

$$-\infty < \mu_2(u,x) \le \max\left\{0, \inf_{x \in X(\Omega)} Q(u,x)\right\}. \tag{9}$$

Условие (9) означает, что либо $\mu_2(u,x) \le 0$, либо, $\mu_2(u,x) \le \inf_{x \in X(\Omega)} Q(u,x)$

Рассмотрим следующую задачу частично целочисленного программирования:

$$\begin{cases} \varphi \to \inf\limits_{\varphi \in R^1, u \in U, w_1, ..., w_K}, \\ \Phi(u,x_k) - \varphi,, (\Phi(u,x_k) - \mu_1(u,x))w_k,\ k = \overline{1,K}, \\ Q(u,x_k) - \varphi,, (Q(u,x_k) - \mu_2(u,x))w_k,\ k = \overline{1,K}, \\ \sum_{k=1}^{K} w_k p_k,, 1-\alpha, \\ w_k \in \{0,1\}, k = \overline{1,K} \end{cases} \tag{10}$$

В работе [6] получен следующий результат о возможности сведения задачи квантильной оптимизации к задачам частично целочисленного программирования.

Теорема 1. *Если выполнены предположения А, Б, то частично целочисленная задача оптимизации* (10) *эквивалентна в смысле определения 3 каждой из задач*

(4), (5). *При этом если* (φ^*, u^*, w_1^*, . . . , w_K^*)– *оптимальное решение задачи* (10), *то* u^*, (φ^*, u^*) – *оптимальные решения соответственно задач* (4), (5).

Замечание 1. Задача (10) в общем случае является задачей нелинейного частично целочисленного программирования, даже если функции Φ, Q линейны по u. Однако, в силу определенной произвольности в выборе функций μ_1 , μ_2 , часто их можно выбрать так, что задача (10) оказывается выпуклой (или даже линейной) частично целочисленной, для решения которой можно применять методы сокращения перебора, реализованные в пакетах программ частично-целочисленного программирования, например, IBM ILOG CPLEX [24]. В общем случае решение задачи (10) сводится к перебору вариантов-подмножеств $I_1 \subseteq \{1,...,k\}$ таких, что $\sum_{k\in I_1} p_k \le 1-\alpha$ решению соответствующих подзадач математического программирования вида

$$\begin{cases} \varphi \to \inf\limits_{\varphi\in R^1, u\in U}, \\ \Phi(u,x_k)\le \varphi,\ k\in\{1,2,...,K\}\setminus I_1; \\ \mu_1(u,x)\le\varphi,\ k\in I_1; \\ Q(u,x_k)\le \varphi,\ k\in\{1,2,...,K\}\setminus I_1; \\ \mu_2(u,x)\le 0, k\in I_1. \end{cases} \qquad (11)$$

и выбору варианта с минимальным значением целевой функции. Если множество U выпукло и функции Φ, Q, μ_1, μ_2 выпуклы по $u\in U$, то (11) – задача выпуклого программирования. Если, кроме того, Φ, Q, μ_1, μ_2 кусочно-линейны по u, а U задается линейными ограничениями, то (11) сводится к задаче линейного программирования.

Замечание 2. Результат теоремы 1 может быть обобщен на задачи квантильной оптимизации с функциями дискретного максимума, а именно, пусть в задачах (4), (5) функции Φ, Q имеют вид

$$\Phi(u,x)=\max_{i\in I}\Phi_i(u,x),\ Q(u,x)=\max_{j\in J}Q_j(u,x) \qquad (12)$$

где I,J конечные множества индексов. Предположим, что существуют функции, $\mu_{i1}(u,x)$, $\mu_{j2}(u,x)$, удовлетворяющие при каждом $u\in U, x\in X(\Omega)$ и $i\in I,\ j\in J$ условиям

$$-\infty<\mu_{1i}(u,x)\le\inf_{x\in X(\Omega)}\Phi_i(u,x), \qquad (13)$$

$$-\infty<\mu_{2j}(u,x)\le\max\left\{0,\ \inf_{x\in X(\Omega)}Q_j(u,x)\right\}. \qquad (14)$$

Составим следующую задачу частично целочисленного программирования:

$$\begin{cases}\varphi\to\inf\limits_{\varphi\in R^1,u\in U,w_1,\dots,w_K},\\ \Phi_i(u,x_k)-\varphi,,(\Phi_i(u,x_k)-\mu_{1i}(u,x_k))w_k,\ i\in I,\ k=\overline{1,K},\\ Q_j(u,x_k)-\varphi,,(Q_j(u,x_k)-\mu_{2j}(u,x_k))w_k,\ j\in J,\ k=\overline{1,K},\\ \sum_{k=1}^{K}w_k p_k,,1-\alpha,\\ w_k\in\{0,1\},k=\overline{1,K}\end{cases} \qquad (15)$$

Следствие 1. *Задачи* (4), (5) *при условиях* (12), (13), (14) *эквивалентны задаче частично целочисленного программирования* (15).

Замечание 3. Хотя задачи (10), (15) не линейны по (u,w_k), но в силу достаточной свободы в выборе функций $\mu_{1,}\ \mu_2,\mu_{1i},\mu_{2j}$ последние можно подобрать так, что коэффициенты при w_k не будут зависеть от u и, таким образом, задачи (10), (15) станут линейными по w_k. Линейность по w_k позволяет использовать непрерывную релаксацию задач (10), (15) для получения оценок снизу оптимальных значений этих задач. А если при этом функции Φ,Q являются функциями максимума из линейных функций и множество U полиэдрально, то мы приходим к задаче (15) частично целочисленного линейного программирования, которую можно решить стандартными программными пакетами оптимизации, на-

пример, IBM ILOG CPLEX [24]. Ниже следующие примеры иллюстрируют эти возможности.

Пример 1. Сделаем следующее предположение.

Предположение Б'. *Пусть существуют (известны) константа* μ *и функции* $M(x)$ *и* $N(x)$ *такие, что функции* Φ, Q *из* (4), (10) *удовлетворяют условию*

$$-\infty < \mu(u,x) \le \inf_{u\in U,\, x\in X(\Omega)} \Phi(u,x), \tag{16}$$

и для всех $x \in X(\Omega)$ *выполнено*

$$\sup_{u\in U} \Phi(u,x) \le M(x) < \infty, \ \sup_{u\in U} Q(u,x) \le N(x) < \infty \tag{17}$$

Возьмем функции

$$\mu_1(u,x) = \Phi(u,x) - M(x) + \mu, \ \mu_2(u,x) = Q(u,x) - N(x) \tag{18}$$

Они удовлетворяют Предположению Б. Действительно, для любого $u \in U$ в силу (16), (17) выполнено

$$\mu_1(u,x) \le \mu \inf_{u\in U,\, x\in X(\Omega)} \Phi(u,x) \le \inf_{u\in U} \Phi(u,x),$$
$$\mu_2(u,x) = Q(u,x) - N(x) \le 0.$$

Функции $M(x), N(x)$ и число μ заведомо существуют, если функции $\Phi(u,x), Q(u,x)$, непрерывны по u при каждом $x \in X(\Omega)$ и множество U компактно.

Подставляя (18) в (10), приходим к следующей линейной по булевым переменным задаче частично целочисленного программирования, эквивалентной при предположениях А, Б' (16), (17) исходным задачам (4), (5):

$$\begin{cases} \varphi \to \inf\limits_{\varphi\in R^1,\, u\in U,\, w_1\in\{0,1\},\, w_K\in\{0,1\}}, \\ \sum_{k=1}^{K} p_k w_k \le 1-\alpha, \\ \Phi(u,x_k) - \varphi \le (M(x_k) - \mu) w_k, \ k = 1,2,...,K, \\ Q(u,x_k) \le N(x_k) w_k, k = 1,2,...K. \end{cases} \tag{19}$$

Пример 2. Рассмотрим частный случай задачи (4), когда функции $\Phi(u,x)$ и $Q(u,x)$ сепарабельны и имеют вид

$$\Phi(u,x)=\Phi_1(u)+\Phi_2(x),\ Q(u,x)=Q_1(u)+Q_2(x)$$

Предположим, что существует константа μ_3 такая, что

$$-\infty<\mu_3\le\min\left\{\inf_{x\in X(\Omega)}\Phi_2(x),\ \inf_{x\in X(\Omega)}Q_2(x)\right\}.$$

Возьмем

$$\mu_1(u,x)=\Phi_1(u)+\mu_3,\ \mu_1(u,x)=Q_1(u)+\mu_3. \qquad (20)$$

Эти функции удовлетворяют Предположению Б. Действительно,

$$\mu_1(u,x)=\Phi_1(u)+\mu_3\le+\inf_{x\in X(\Omega)}\Phi_2(x)\le\inf_{x\in X(\Omega)}(\ \Phi_1(u)+\Phi_2(x)).$$

Аналогично проверяется второе условие Б. Подставляя (20) в (10), приходим еще к одной задаче частично целочисленного программирования, эквивалентной при сделанных предположениях задаче (4):

$$\begin{cases}\varphi\to\inf\limits_{\varphi\in R^1,\,u\in U,\,w_1,\ldots,\,w_k},\\ \Phi_1(u)+\Phi_2(x_k)-w_k(\Phi_2(x_k)-\mu_3),\ \varphi,k=\overline{1,K},\\ Q_1(u)+Q_2(x_k)-w_k(Q_2(x_k)-\mu_3),\ 0,\ k=\overline{1,K},\\ \sum_{k=1}^{K}w_k p_k\ ,1-\alpha,\\ w_k\in\{0,1\},k=\overline{1,K}.\end{cases} \qquad (21)$$

Заметим, что если функции $\Phi_1(u)$ и $Q_1(u)$ выпуклы по $u\in U$, то задача (21), является задачей смешанного выпуклого программирования.

Пример 3. В работе [14] рассмотрен еще один частный случай задачи (4), когда функции $\Phi(u,x)$ и $Q(u,x)$ являются кусочно-линейными выпуклыми функциями максимума вида

$$\Phi(u,x)=\max_{i=1,\dots,k_1}(A_{1i}u+B_{1i}x+b_{1i}),$$
$$Q(u,x)=\max_{i=1,\dots,k_2}(A_{2j}u+B2jx+b2j), \tag{22}$$

где $u\in[0,\bar{u}]^n, x\in R^m, A_{1i}, B_{1i}, A_{2j}, B_{2j}$ – строки матриц A_1, A_2, B_1, B_2 соответственно; b_{1i}, b_{2j} – компоненты векторов b_1, b_2. Пусть μ_3 определяется следующим образом

$$\mu_3, \min\{\min_{i=\overline{1,k_1},\,k=\overline{1,K}} B_{1i}x_k, \min_{j=\overline{1,k_2},\,k=\overline{1,K}} B_{2j}x_k\}.$$

В указанной работе доказано, что в этом случае при многогранном выпуклом множестве U задача (5) эквивалентна следующей задаче частично целочисленного линейного программирования:

$$\begin{cases}\varphi\to \inf\limits_{\varphi\in R^1,\,u\in U,\,w_1,\dots,\,w_k},\\ A_{1i}u+\mu_3+w_k(B_{1i}x_k-\mu_3)+b_{1i},\varphi, i=\overline{1,k_1}, k=\overline{1,K},\\ A_{2j}u+\mu_3+w_k(B_{2j}x_k-\mu_3)+b_{2j},\varphi, j=\overline{1,k_2}, k=\overline{1,K},\\ \sum_{k=1}^{K} w_k p_k \dots \alpha,\\ w_k\in\{0,1\}, k=\overline{1,K}.\end{cases} \tag{23}$$

Данная задача является частным случаем (15).

Замечание 4. Для проверки корректности постановки задач (4), (5) в силу (2) необходимо найти

$$P^*=\sup \mathrm{P}\{Q(u,X)\le 0\}. \tag{24}$$

Задача максимизации вероятности (24) в предположении (17) сводится к следующей задаче частично целочисленного программирования [6], [14], [20]:

$$\sum_{k=1}^{K} p_k w_k \to \sup_{u\in U,\,w_1\in\{0,1\},\dots,w_K\in\{0,1\}},$$
$$Q(u,x_k)\le(Q(u,x_k)-\mu_2(u,x_k))w_k, k=1,\dots,K. \tag{25}$$

Задачи двухэтапного квантильного стохастического программирования

Двухэтапная задача стохастического программирования с критерием в форме математического ожидания имеет вид [1]:

$$f_1(u)+\mathrm{E}[\Phi(u,X)]\to \min_{u\in U},$$

где

$$\Phi(u,x)=\begin{cases} f_2^*(u,x)=\inf\limits_{v\in W(u,x)} f_2(u,v,x),\ W(u,x)\neq\varnothing, \\ +\infty,\ W(u,x)\neq\varnothing, \end{cases} \qquad (26)$$
$$W(u,x)=\{v\in V(x): Q_2(u,v,x)\le 0\},$$

$U\subseteq R^n$ множество допустимых стратегий первого этапа; $u\in U$ – детерминированная стратегия первого этапа; $f_1(u)$ целевая функция первого этапа; X случайный параметр, принимающий (в данной статье) конечное множество значений $X(\Omega)=\{x_1,...,x_K\}$ с вероятностями $p_1,...,p_K$ (предположение А); $V(x)\subset R^m$ множество допустимых стратегий второго этапа при $x=X$, $V(x)\subset R^m$ стратегия второго этапа (коррекция); $f_2(u,v,x)$ целевая функция второго этапа; $Q_2(u,v,x)$ – функция в ограничениях второго этапа; Е знак математического ожидания.

Отметим, что множество допустимых стратегий в двухэтапной задаче может быть уже, чем U, поскольку оно включает неявное ограничение $\mathrm{E}[\Phi(u,X)]<+\infty$. Неинтегрируемость величины $\Phi(u,X)$ может возникать как в силу неинтегрируемости случайных параметров задачи, так и в силу возможных бесконечных значений $\Phi(u,X)$. Например, такая ситуация возникает, если распределение случайных параметров имеет "тяжелые хвосты".

Вместо среднего $\mathrm{E}[\Phi(u,X)]$ в двухэтапной задаче может использоваться медиана или другая квантиль случайной величины $\Phi(u,X)$ [7 – 10].

Определим функции вероятности и квантили:

$$P_\varphi(u)=\mathrm{P}\{\Phi(u,X)\le\varphi\},\ \varphi_\alpha(u)=\min_\varphi\{\varphi:\ P_\varphi(u)\ge\alpha\},0<\alpha<1.$$

Двухэтапная задача квантильной оптимизации имеет вид:

$$\{f_1(u)+\varphi_\alpha(u)\}\to\min_{u\in U}. \tag{27}$$

Она эквивалентна следующей задаче

$$\begin{gathered} f_1(u)+\varphi\to\min_{u\in U,\ \varphi\in R^1}, \\ \mathrm{P}\{\Phi(u,X)\le\varphi\}\ge\alpha \end{gathered} \tag{28}$$

Задача (28) с фиксированным параметром φ, например, $\varphi=0$, и $\Phi(u,X)$ из (26) имеет самостоятельный интерес и называется задачей двухэтапного стохастического программирования с вероятностным ограничением (на случайные значения целевой функции второго этапа).

Лемма 3. *Задачи* (27) *и* (28) *эквивалентны.*

Доказательство. Пусть u' – допустимое решение задачи (27), тогда $\varphi'=\varphi_\alpha(u')$ конечно. В силу свойств квантили

$$\mathrm{P}\{\Phi(u',X)\le\varphi'\}=\mathrm{P}\{\Phi(u',X)\le\varphi_\alpha(u')\}\ge\alpha.$$

Таким образом, (φ',u') допустимое решение задачи (28) с тем же самым значением $f_1(u')+\varphi_\alpha(u')$ целевой функции.

Обратно, пусть (φ',u') – допустимое решение задачи (28), т.е $u'\in U$, $\mathrm{P}\{\Phi(u',X)\le\varphi'\}\ge\alpha>0$ и $-\infty<f_1(u')+\varphi'<+\infty$. В силу непрерывности вероятностной меры $\lim\limits_{\varphi\to-\infty}\mathrm{P}\{\Phi(u',X)\le\varphi'\}=0$. Поэтому $-\infty<\varphi_\alpha(u')+\varphi'$ и, следовательно, для значений целевых функций задач (28), (27) имеем неравенство $f'=f_1(u')+\varphi'\ge f_1(u')+\varphi_\alpha(u')>-\infty$, т.е. u' – допустимое решение для задачи (27). Таким образом, в силу леммы 1 задачи (27) и (28) эквивалентны. Лемма доказана.

Рассмотрим также следующую задачу

$$f_1(u)+\varphi \to \min_{u\in U,\,\varphi\in R^1,\,v_1\in V(x_k),\ldots,v_K\in V(x_K)},$$
$$\sum_{k=1}^{K} p_k\, I\{f_2(u,v_k,x_k)\le\varphi, Q_2(u,v_k,x_k)\le 0\}\ge\alpha, \qquad (29)$$

где $I\{\cdot\}$ индикатор условий в фигурных скобках, равный единице, если все условия в скобках выполнены, и равный нулю в противном случае.

Предположение В *(о существовании решения задачи второго этапа).* *Если множество* $W(u,x)=\{v\in V(x): Q_2(u,v,x)\le 0\}$ *не пусто, то* inf *в* (24) *достигается и, таким образом, существует* $v(u,x)$ *такое, что* $\Phi(u,x)=f_2(u,v(u,x),x)$.

Лемма 4. *При предположениях А, В задачи* (28) *и* (29) *эквивалентны.*

Доказательство. Пусть (φ',u') – допустимое решение задачи (28),т.е. $u'\in U$, $\mathrm{P}\{\Phi(u',X)\le\varphi'\}\ge\alpha$, где функция $\Phi(u,x)$ определена в (26). Обозначим I'_α – множество индексов k таких, что $\Phi(u',x_k)\le\varphi'$. Очевидно, $\sum_{k\in I'_\alpha} p_k\ge\alpha$. Для $k\in I'_\alpha$ множество $W(u',x_k)$ не пусто и в силу предположения. В существуют $v'_k\in V(x_k)$ такие, что $f_2(u',v'_k,x_k)=\Phi(u',x_k)\le\varphi'$ и $Q_2(u',v'_k,x_k)\le 0$. Для $k\notin I'_\alpha$ выберем произвольные $v'_k\in V(x_k)$. Очевидно, набор $(\varphi',u',v'_1,\ldots,v'_K)$ является допустимым для задачи (29) с тем же самым значением $f'=f_1(u')+\varphi'$ целевой функции.

Обратно, пусть набор $(\varphi',u',\{v'_k\}_1^K)$ является допустимым для задачи (29). Покажем, что (φ',u') является допустимым решением для задачи (28). Обозначим с помощью I'_α множество индексов k таких, что $v'_k\in V(x_k)$, $f_2(u',v'_k,x_k)\le\varphi'$ и $Q_2(u',v'_k,x_k)\le 0$. Очевидно, $\sum_{k\in I'_\alpha} p_k\ge\alpha$. Для $k\in I'_\alpha$ выполнено $\Phi(u',x_k)\le f_2(u',v'_k,x_k)=\le\varphi'$, поэтому $\mathrm{P}\{\Phi(u',X)\le\varphi'\}\ge\sum_{k\in I'_\alpha} p_k\ge\alpha$, и, таким образом, (f',φ',u') допустимое решение для задачи (28) с тем же самым

значением $f' = f_1(u') + \varphi'$ целевой функции. Следовательно, в силу леммы 1 задачи (28) и (29) эквивалентны. Лемма доказана.

В силу транзитивности отношения эквивалентности все задачи (27), (28) и (29) эквивалентны при предположениях А, В. Задача (29) того же типа, что (5), для которой возможность сведения к частично целочисленному эквиваленту (19) при предположениях Б' показана в примере 1. Сделаем следующие дополнительные предположения относительно задачи (27) и, следовательно, (29).

Предположения Г.

Г1. $\inf\limits_{u,v,x} \{f_2(u,v,x): u \in U, v \in V(x), x \in X(\Omega), Q_2(u,v,x) \le 0\} \ge \mu > -\infty$.

Г2. $\sup\limits_{u \in U, v \in V(x)} f_2(u,v,x) \le M_2(x) < \infty, \ldots, \sup\limits_{u \in U, v \in V(x)} Q_2(u,v,x) \le N_2(x) < \infty$.

В частности, условие Г2 выполнено, если функции $f_2(u,v,x)$, $Q_2(u,v,x)$ полунепрерывны сверху по (u,v) на компактном множестве $U \times V(x)$ для каждого $x \in X(\Omega)$ и можно найти (численно или аналитически) оценки вида Г2 для каждого $x \in X(\Omega)$.

Предположения Г'.

Г1'. *Оптимальное значение функции квантили в* (27) *ограничено снизу (известной) константой* $\mu \to -\infty$, *т.е. для любого оптимального решения* u^* *задачи* (27) *выполнено* $\varphi(u^*) \ge \mu \to -\infty$.

Г2'. *Существуют (известны) функции* $M_2(x), N_2(x)$ *при* $x \in X(\Omega)$ *такие, что для любой оптимальной стратегии* u^* *в задаче* (27) *и соответствующей ей оптимальной стратегии* $v^*(x)$ *в задаче второго этапа* (26) *для всех* x *таких, что* $\Phi(u^*, x) < +\infty$, *выполнено*

$$f_2(u', v^*(x), x) \le M_2(x) < \infty, \quad Q_2(u', v^*(x), x) \le N_2(x) < \infty .$$

Условие Г1' заведомо выполнено, если выполнено условие Г1. Условие Г2' выполнено, если выполнено Г2.

Обозначим $w_k \in \{0,1\},\ k = 1, 2, \ldots, K$ – набор булевых переменных. Общая двухэтапная дискретная задача в квантильной постановке при дискретном распределении случайных данных может быть эквивалентно представлена в виде:

$$\begin{cases} f_1(u) + \varphi \to \min\limits_{u \in U, \varphi \in R^1, v_1 \in V(x_1), \ldots, v_K \in V(x_K), w_1 \in \{0,1\}, \ldots, w_K \in \{0,1\}}, \\ \sum_{k=1}^{K} p_k w_k \le 1 - \alpha\ , \\ f_2(u, v_k, x_k) - \varphi \le (M(x_k) - \mu) w_k\ , k = 1,2,\ldots,K\ , \\ Q_2(u, v_k, x_k) \le N(x_k)\, w_k\ , k = 1,2,\ldots K, \\ w_k \in \{0,1\},\ k = 1,2,\ldots K\ . \end{cases} \tag{30}$$

Основной результат статьи представляет следующая теорема о возможности сведения задачи двухэтапной квантильной оптимизации (25) к задаче частично целочисленного программирования (30).

Теорема 2. *При условиях А, В, Г1, Г2 задачи* (27) – (30) *эквивалентны в смысле определения 3. Если решение задачи* (25) *существует и выполнены предположения А, Б, Г1', Г2', то задачи* (27) – (30) *также эквивалентны. Если* $(\varphi^*, u^*, v_1^*, \ldots, v_K^*, w_1^*, \ldots, w_K^*)$ *– оптимальное решение задачи* (30)*, то* u^* *– оптимальное решение задачи* (27)*,* (φ^*, u^*) *оптимальное решение задачи* (28)*, а* $(\varphi^*, u^*, v_1^*, \ldots, v_K^*)$ *– оптимальное решение задачи* (29)*.*

Доказательство. Эквивалентность задач (27) –(29) доказана в лемме 3. Докажем эквивалентность задач (29) и (30) с помощью леммы 1.

В предположениях А, В, Г1, Г2 докажем первое утверждение теоремы. Пусть $(\varphi', u', v_1', \ldots, v_K')$ – допустимое решение задачи (29) со значением целевой функции $f' = f_1(u') + \varphi'$. Вычислим булевы значения

$$w_k' = \begin{cases} 0, f_2(u', v_k', x_k) - \varphi' \le 0,\ Q_2(u', v_k', x_k) \le 0; \\ 1, \textit{в противном случае.} \end{cases}$$

Покажем, что $(\varphi',u',v'_1,...,v'_K,w'_1,...,w'_K)$ – допустимое решение задачи (30). Обозначим I'_0 и I'_1 множества индексов k таких, что $w_k=0$ и $w_k=1$ соответственно. Тогда из ограничения-неравенства (29) следует

$$\sum_{k=1}^{K} p_k w'_k = 1-\sum_{k=1}^{K} p_k (1-w'_k) = 1-\sum_{k\in I'_0} p_k =$$

$$1-\sum_{k=1}^{K} p_k \, I\{f_2(u',v'_k,x_k)-\varphi'\le 0, Q_2(u',v'_k,x_k)-\varphi'\le 0\} \le 1-\alpha.$$

Таким образом, ограничение $\sum_{k=1}^{K} p_k w'_k \le 1-\alpha$ в (30) выполнено. Проверим две следующие группы ограничений в (30). Для $k\in I'_0$ эти ограничения принимают вид $f_2(u',v'_k,x_k)-\varphi'\le 0$, $Q_2(u',v'_k,x_k)-\varphi'\le 0$ и выполнены в силу определения множества индексов I'_0. В силу предположения допустимости $(\varphi',u',v'_1,...,v'_K)$ из (29) следует, что существуют v'_k, x_k такие, что $f_2(u',v'_k,x_k)-\varphi'\le 0$, $Q_2(u',v'_k,x_k)-\varphi'\le 0$. В силу предположения В выполнено

$$\varphi' \ge f_2(u',v'_k,x_k) \ge \inf_{u,v,x}\{f_2(u,v,x): u\in U, v\in V(x), x\in X(\Omega), Q_2(u,v,x)\le 0\} \ge \mu$$

Для $k\in I'_1$ в силу предположений Г1, Г2 и доказанного неравенства $\varphi'\ge\mu$ имеет место

$$f_2(u',v'_k,x_k)-\varphi' \le \sup_{u\in U, v\in V(x_k)} f_2(u',v'_k,x_k)-\mu \le M(x_k)-\mu,$$

$$Q_2(u',v'_k,x_k) \le \sup_{u\in U, v\in V(x_k)} Q_2(u',v'_k,x_k) \le N(x_k).$$

что и требовалось доказать. Таким образом, $(\varphi',u',v'_1,...,v'_K,w'_1,...,w'_K)$ – допустимое решение задачи (30) с тем же самым значением $f'=f_1(u')+\varphi'$ целевой функции.

Пусть теперь $(\varphi',u',v'_1,...,v'_K,w'_1,...,w'_K)$ – допустимое решение задачи (30) со значением целевой функции $f'=f_1(u')+\varphi'$. Проверим, что $(\varphi',u',v'_1,...,v'_K)$ – допустимое решение задачи (29). Обозначим I'_0 множестве индексов k таких,

что $w_k = 0$. В силу допустимости имеет место $\sum_{k=1}^{K} p_k w'_k \le 1-\alpha$, что эквивалентно неравенству $\sum_{k\in I'_0} p_k \ge \alpha$. При $k \in I'_0$ выполнено $w_k = 0$ и, следовательно, $f_2(u', v'_k, x_k) - \varphi' \le 0,\ Q_2(u', v'_k, x_k) - \varphi' \le 0$. Теперь проверим выполнение вероятностного ограничения в (29),

$$\sum_{k=1}^{K} p_k\, I\{f_2(u', v'_k, x_k) - \varphi' \le 0,\ Q_2(u', v'_k, x_k) - \varphi' \le 0\} \ge$$

$$\ge \sum_{k\in I'_0} p_k\, I\{f_2(u', v'_k, x_k) - \varphi' \le 0,\ Q_2(u', v'_k, x_k) - \varphi' \le 0\} = \sum_{k\in I'_0} p_k \ge \alpha$$

Таким образом, $(\varphi', u', v'_1, ..., v'_K)$ допустимое решение задачи (29) с тем же самым значением $f' = f_1(u') + \varphi'$ целевой функции. Теперь справедливость первого утверждения теоремы следует из леммы 1. Второе утверждение теоремы доказывается аналогично, но на основе леммы 2. Теорема доказана.

Отметим, что если функции $f_2(u, v, x),\ Q_2(u, v, x)$ кусочно-линейны по (u, v) например, являются функциями максимума из линейных по (u, v) функций, а множества $U,\ V(x)$ задаются линейными ограничениями, то задача (30) сводится к задаче линейного частично целочисленного программирования.

Замечание 5. Рассмотрим задачу (28) с фиксированным φ например, $\varphi \equiv 0$ которая в этом случае называется задачей двухэтапного стохастического программирования с вероятностным ограничением, а также рассмотрим задачи (29), (30) с тем же самым фиксированным φ. При этом лемма 4 и теорема 2 остаются справедливыми. Таким образом, задача двухэтапного стохастического программирования с вероятностным ограничением (28) с $\varphi \equiv 0$ при дискретном распределении случайных данных эквивалентно сводится к задаче частично целочисленного программирования (30) с $\varphi \equiv 0$.

Численные эксперименты

Линейная двухэтапная задача квантильной оптимизации. В работе [23] рассмотрен численный пример двухэтапной линейной задачи квантильной оптимизации вида (27), где на первом этапе решается задача

$$c_1^T u + \varphi_\alpha(u) \to \min_{u \in U},\ U = \{u : A_1 u \geq b_1, u \geq 0\},$$

а задача второго этапа имеет вид:

$$\Phi(u,x) = \min_{y \in Y(u,x)} c_2^T v,\ Y(u,x) = \{v : v \geq 0, B_2 v \geq x - A_2 u\}.$$

Параметры задач имеют следующие числовые значения $\alpha = 0.5$,

$$c_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix},\ c_2 = \begin{pmatrix} 0.1 \\ 0.36 \end{pmatrix},\ B_2 = \begin{pmatrix} -0.875 & 1.8 \\ 1 & -1 \end{pmatrix},\ A_2 = \begin{pmatrix} 1.1 & 1.7 \\ 2.8 & 2.4 \end{pmatrix},\ A_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, b_1 = \begin{pmatrix} -10 \\ -10 \end{pmatrix}.$$

Случайный вектор X (правых частей) задан рядом распределения в следующей таблице:

x_k	$\begin{pmatrix} 2.5 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 1.5 \\ 2.1 \end{pmatrix}$	$\begin{pmatrix} 2.3 \\ 3 \end{pmatrix}$	$\begin{pmatrix} 3.2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 7 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 5 \end{pmatrix}$	$\begin{pmatrix} 8 \\ 9 \end{pmatrix}$
p_k	0.05	0.10	0.10	0.25	0.15	0.15	0.15	0.05

Приближенное решение $u' = (u'_1, u'_2)$, найденное в (Наумов, Бобылев (2012) [23]) путем перебора имеет вид:

u'_1	u'_2	$\Phi_\alpha(u')$	$c_1^T u^* + \Phi_\alpha(u')$
0	1.56	0.1	1.67

Решим эту задачу путем сведения к задаче частично целочисленного линейного программирования вида (30):

$$c_1^T u + \varphi \to \min_{0 \leq u_1, u_2 \leq \overline{U}, \varphi \geq \mu, 0 \leq v_1, \ldots, v_K \leq V, w_1 \in \{0,1\}, \ldots, w_K \in \{0,1\}},$$

$$\sum_{k=1}^{K} p_k w_k \leq 1 - \alpha\ ,$$

$$A_1 u \geq b_1,$$

$$c_2^T v_k - \varphi \leq (M\ (x_k) - \mu) w_k,\ k = 1, 2, ..., K,$$

$$-A_{2i} u - B_{2i} v_k + x_{ki} \leq N(x_k) w_k,\ i = 1, 2,\ k = 1, 2, ..., K;$$

где $A_{2i}, B_{2i} - i$-е строки матриц $A_2, B_2; x_k = \{x_{ki}, i = 1, 2\},\ \mu = 0, \overline{U} = \overline{V} = 100;$

$$M\ (x_k) \equiv M = \overline{V}\ (c_{21} + c_{22}) = 46;$$

$$N(x_k) \equiv N = \max_{i=1,2} \left(\overline{U} \sum_{j=1,2} \left| (A_2)_{ij} \right| + \overline{V} \sum_{j=1,2} \left| (B_2)_{ij} \right| + \max_{k=1,...,K} \left| x_{ki} \right| \right) = 303.5$$

Данная задача содержит 11 непрерывных переменных, 8 булевых переменных, одно булево ограничение-неравенство, два непрерывных ограничений-неравенств, 24 смешанных ограничений-неравенств. Точное решение этой задачи $u^* = (u_1^*, u_2^*)$ полученное программой IBM ILOG CPLEX V12.1 [24] (с параметрами по умолчанию) за 0.05 секунды на том же персональном компьютере равно:

$$u_1^* = 0,\ u_2^* = 1.5331,\ \varphi^* = 0.1188,\ c_1^T u^* + \varphi^* = 1.6518.$$

Очевидно, что найденное оптимальное решение u^* лучше, чем приближенное решение u'.

Заключение

Функции квантили могут использоваться вместо функций среднего значения в задачах стохастического программирования. Они описывают функционирование стохастической системы в экстремальных условиях. Это популярный подход в современной финансовой оптимизации. Однако, функции квантили, как правило, не выпуклы. Поэтому для решения задач квантильной оптимизации применяют либо эвристические приближенные методы, либо методы глобальной оптимизации. В настоящей статье изучаются задачи двухэтапной квантильной оптимизации с дискретным распределением случайных данных. Оказывается, что как и в случае одноэтапных задач они могут быть эквивалентно

сведены к детерминированным частично целочисленным задачам оптимизации. При этом число дополнительных булевых переменных равно числу возможных значений вектора случайных параметров исходной задачи. Аналогичное преобразование возможно и для двухэтапных задач с вероятностным ограничением на целевую функцию второго этапа. Полученные смешанные задачи предполагается решать стандартными программными средствами дискретной оптимизации. Результаты проиллюстрированы численным примером небольшой размерности. Таким образом, обоснован практический метод решения реальных задач квантильного двухэтапного стохастического программирования. Дальнейшая работа будет направлена на разработку специальных методов дискретной оптимизации, учитывающих структуру возникающих частично целочисленных задач.

Список литературы

[1] Ермольев Ю.М. Методы стохастического программирования. – М.: Наука, 1976.

[2] Юдин Д.Б. Задачи и методы стохастического программирования. – М.: Сов. Радио, 1979.

[3] Малышев В.В., Кибзун А.И. Анализ и синтез высокоточного управления летательными аппаратами. – М.: Машиностроение, 1987.

[4] Kibzun, A.I., Kan, Y.S. Stochastic programming problems with probability and quantile functions. – Chichester, New York, Brisbane, Toronto, Singapore: John Wiley & Sons, 1996.

[5] Кибзун А.И., Кан Ю.С. Задачи стохастического программирования с вероятностными критериями. – М.: ФИЗМАТЛИТ, 2009.

[6] Кибзун А.И., Наумов А.В., Норкин В.И. О сведении задачи квантильной оптимизации с дискретным распределением к задаче смешанного целочисленного программирования // Автоматика и телемеханика. (В печати, 2012).

[7] Кибзун А.И., Наумов А.В. Гарантирующий алгоритм решения задачи квантильной оптимизации // Космические исследования, 1995. – Т. 33. – № 2. – С. 160 – 165.

[8] Богданов А.Б., Наумов А.В. Исследование двухэтапной целочисленной задачи квантильной оптимизации // Изв. РАН. Теория и системы управления, 2003. – № 5. – С. 62 – 69.

[9] Богданов А.Б., Наумов А.В. Решение двухэтапной задачи логистики в квантильной постановке // Автоматика и телемеханика, 2006. –№ 12. –С.36 – 42.

[10] Наумов А.В. Двухэтапная задача квантильной оптимизации инвестиционного проекта // Изв. РАН. Теория и системы управления, 2010. – № 2. – С. 40 – 47.

[11] Larsen N., Mausser H., Uryasev S. Algorithms for optimization of value-at-risk. In P. Pardalos and V.K. Tsitsiringos, editors, Financial Engineering, e-Commerce and Supply Chain, Kluwer Academic Publishers, 2002. – Pp. 129 – 157.

[12] Wozabal D., Hochreiter R., Pflug G.Ch. A D.C. Formulation of Value-at-Risk constrained optimization. Optimization, 2010. – V. 59(3). – Pp. 377 – 400.

[13] Norkin V. On mixed integer reformulations of monotonic probabilistic programming problems with discrete distributions // http://www.optimization-online.org/DB_HTML/2010/05/2619.html. 2010.

[14] Иванов С.В., Наумов А.В. Алгоритм оптимизации квантильного критерия для полиэдральной функции потерь и дискретного распределения случайных параметров // Автоматика и телемеханика, 2012. – № 1. – С. 95 – 108.

[15] Sen S. Relaxation for probabilistically constrained programs with discrete random variables // Operations Research Letters, 1992. –Vol. 11. – Pp. 81 – 86.

[16] Ruszczyński A. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra // Math. Program, 2002. – Vol. 93. – Pp. 195 – 215.

[17] Benati S., Rizzi R. A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem // European Journal of Operational Research, 2007. – Vol. 176. – Pp. 423 – 434.

[18] Luedtke J., Ahmed S., Nemhauser G. An integer programming approach for linear programs with probabilistic constraints // Math. Program, 2010. – Vol. 122(2). – Pp. 247 – 272.

[19] Корбут А.А., Финкельштейн Ю.Ю. Дискретное программирование. – М.: Наука, 1969.

[20] Норкин В.И., Бойко С.В. Оптимизация финансового портфеля на основе принципа безопасности // Кибернетика и системный анализ, 2012. – № 2. – С. 29 – 41.

[21] Birge J., Luveaux F. Introduction to Stochastic Programming. – New York: Springer-Verlag, 1997.

[22] Shapiro A., Dentcheva D., Ruszczyński A. Lectures on stochastic programming: Modeling and theory. – SIAM, Philadelphia, 2009.

[23] Наумов А.В., Бобылев И.М. О двухэтапной задаче стохастического линейного программирования с квантильным критерием и дискретным распределением случайных параметров // Автоматика и телемеханика, 2012. – № 2. – С. 61 – 72.

[24] IBM ILOG CPLEX V12.1. User's Manual for CPLEX. − International Business Machines Corporation, 2009. − 952 p.

[25] Райк Э. Качественные исследования в задачах стохастического нелинейного программирования // Изв. АН СССР, физ.-мат, 1971. – Т.20. – №1. – С.8 – 14.

[26] Prekopa A. Stochastic Programming. – Kluwer Academic Publishers, 1995.

[27] Райк Э. О функции квантиля в стохастическом нелинейном программировании // Изв. АН СССР, физ.-мат, 1971. – Т.20. – № 2. – С. 229 – 231.

[28] Райк Э. О задачах стохастического программирования с решающими функциями // Изв. АН СССР, 1972. – Т. 21. – С. 258 – 263.

[29] Обен Ж-П., Экланд И. Прикладной нелинейный анализ. – М.: Мир, 1988.

[30] Kataoka S. A stochastic programming model // Econometrica, 1963. – Vol.31. – Pp. 181 – 196.

[31] Михалевич В.С., Гупал А.М., Норкин В.И. Методы невыпуклой оптимизации. – М.: Наука, 1987.

[32] Pagnoncelli B.K., Ahmed S., Shapiro A. Sample Average Approximation Method for Chance Constrained Programming: Theory and Applications // J. Optim. Theory Appl., 2009.

УДК 519.856

МЕРЫ РИСКА В ЗАДАЧАХ СТОХАСТИЧЕСКОЙ ОПТИМИЗАЦИИ ДЛЯ ПОЛУЧЕНИЯ РОБАСТНЫХ РЕШЕНИЙ[1]

Кирилюк В. С.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: vlad00@ukr.net

Аннотация. Обсуждается понятие риска, модели и методы стохастического программирования. Анализируются традиционные подходы сведения задач стохастической оптимизации к детерминированным эквивалентам. Рассматривается универсальный подход с использованием мер риска, позволяющий оценивать риск, как при полной, так и неполной информации о распределениях случайных величин. На примерах задач оптимизации портфеля описано применение аппарата полиэдральных когерентных мер риска, позволяющее свести эти задачи к проблемам линейного программирования. Рассмотрена связь между используемыми мерами риска и робастностью получаемых оптимальных решений.

Ключевые слова. Стохастическое программирование, когерентные меры риска, полиэдральные когерентные меры риска, CVaR, оптимизация портфеля, робастное решение.

1. Риск и неопределенность

Неопределенность, как правило, определяется как отрицание определенности (детерминированности) результата. Одним из первых исторических примеров, описывающих неопределенность, был Буриданов осел (XIV век), который находился точно между двумя охапками сена [1]. Вопрос заключался в том, куда пойдет осел?

Не вдаваясь в детальный анализ понятия и природы неопределенности, отметим, что будущие события и процессы всегда несут в себе неопределенность. Мы не знаем того, что случится в будущем, но такое незнание может быть разным. Степень, в которой мы можем использовать свое представление о будущем, различается от случая к случаю. В контексте данной статьи будем рассматривать лишь случаи, когда можно: 1) либо идентифицировать вероятности различных событий, либо 2) как-то их оценивать (интервалами, множествами ограничений). Первые случаи естественно назвать случаями полной информации о распределениях случайных величин (с.в.), вторые –

[1] Работа выполнена при частичной поддержке Государственного фонда фундаментальных исследований Украины в рамках совместного российско-украинского проекта Ф40.1/016 (2011-2012)

случаями неполной информации. Заметим, что в соответствии с [2], первые обстоятельства называются условиями риска, а вторые – условиями неопределенности.

1.1. Об этимологии слова «риск»

Существуют разные версии относительно этимологии англоязычного слова «risk» (риск) [1]. В 16-ом столетии слово «risk» уже активно использовалось в романских языках. Первые корни слова «risk» можно найти в итальянских словах «risco» и «risicare» (опасность и риск). Ряд публикаций упоминают греческое слово «rhizia», что дословно означает корень дерева. Позже на Крите оно было расширено до слова «rhizicon», которое означало скалы у подножия гор на море, представляющие угрозу для мореплавателей. Некоторые публикации ссылаются на древнеперсидское слово «rozi(k)» (дневной заработок, судьба), трансформировавшееся в «rizq», что означало жизнь, зависящая от Бога и судьбы.

1.2. Определение понятия риск

Известно много определений слова «риск». Например, можно найти подобное определение с точки зрения: 1) статистики страхования; 2) токсикологии-эпидемиологии; 3) инженерно-технической; 4) экономической; 5) психологической; 6) социально-теоретической; 7) культурно-теоретической. Наиболее абстрактное его содержание означает потенциальную возможность негативных последствий (убытков, потерь, прочего), обусловленных неопределенностью результата изучаемого процесса.

Риски обычно распределяют на некоторые классы. К примеру, это могут быть: 1) природные; 2) технические; 3) по здоровью; 4) социальные, прочие.

Такая классификация, вообще говоря, условна. Риски зачастую накладываются и мультиплицируются. К примеру, критическая ситуация 2011 года на АЭС Фукусима возникла после землетрясения и вызванного ним цунами. Для оценки разнообразных рисков известно значительное количество разра-

ботанных и применяемых на практике как объективных, так и субъективных оценочных мер риска [1].

1.3. Меры риска, построенные по потенциальных ущербах

В дальнейшем, ограничимся мерами риска, построенными по потенциальных ущербах. Возможный ущерб является интегральным показателем, который позволяет оценивать многие типы и классы рисков унифицированным образом.

Предположим, что можно формально (численно) трансформировать негативные последствия будущих событий изучаемых процессов в форму потенциальных ущербов в денежной виде. Более точное определение меры риска представлено в разделе 3.

Сами же риски, по сути, описывают лишь «одну сторону медали» изучаемых процессов. Как правило, при принятии решений или некотором выборе лица, их принимающие, (ЛПР) делают это, сравнивая риски (в виде потенциальных потерь) и преимущества (в виде потенциальных выигрышей) соответствующих альтернатив.

Отметим, что, как правило, потенциальные выигрыши обычно легко описываются средней прибыльностью, ожидаемой полезностью, или прочими прозрачными критериями, вытекающими из содержания задачи. Однако адекватный выбор меры для оценки риска остается нетривиальной проблемой. Неудачный выбор такой меры может привести в дальнейшем к противоречивым и неэффективным решениям.

2. Модели стохастического программирования

Традиционно, задачи детерминированной оптимизации ставятся в следующем стандартном виде. Для точно определенных функций цели и ограничений $f^0(x)$ и $f^i(x), i = 1,...,m$ соответственно необходимо решить следующую задачу

$$\min\{f^0(x) : f^i(x) \le 0,\ i = 1,...,m;\ x \in X\}$$

для некоторого вектора решений $x=(x_1,...,x_n)\in X$, где множество X имеет «простую» структуру. Но любая детерминированная оптимизационная модель может содержать разнообразные параметры, которые могут быть случайными. В таком случае модель трансформируется в следующую стохастическую постановку

$$\text{"min"}\{f^0(x,\omega): f^i(x,\omega)\le 0,\ i=1,...,m;\ x\in X\}, \qquad (1)$$

где ω – соответствующие стохастические параметры. Тогда при любом фиксированном x ограничения $f^i(x,\omega)\le 0$ могут не выполняться для некоторых реализаций случайного параметра ω, а целевая функция $f^0(x,\omega)$ может принимать разные значения. Проблема заключается в том, что нужно понимать под решением такой задачи. Набор функций $f^0(x,\omega)=\{f^0(x,\omega_1),f^0(x,\omega_2),...,f^0(x,\omega_n),...\},\omega_n\in\Omega$ можно рассматривать как векторную характеристику решения $x\in X$ и трактовать задачу выбора оптимального x как проблему векторной оптимизации, вообще говоря, с бесконечным числом критериев. Более точно, в этом случае «min» в соотношении (1) трактуется как минимальная по распределению случайная величина. Такое отношение порядка называется стохастическим доминированием первого порядка (First-Order Stochastic Dominance), которое будем обозначать как FSD.

Формализацией подобных проблем занимается теория стохастического программирования (SP). Например, вместо проблемы (1) рассматривают проблему одноэтапного SP, в которой функции $f^i(x,\omega)$ заменяют их средними значениями

$$F^0(x)=Ef^0(x,\omega)\to\min \qquad 2)$$

при ограничениях

$$F^i(x)=Ef^i(x,\omega)\le 0, x\in X, \qquad (3)$$

где ω – вектор случайных параметров, E – математическое ожидание по ω на некотором вероятностном пространстве (Ω,Σ,P). Отметим, что иногда в ка-

честве функций $F^i(x)$ учитываются ожидаемые полезности вида $(-Eu(f^i(x,\omega)))$ для функции полезности $u(.)$.

Для такой задачи в качестве основного метода решения можно упомянуть метод стохастического квазиградиента (SQG), предложенный Ю.М. Ермольевым и развиваемый в его научной школе (например, [3–6]).

2.1. SQG как прямой метод решения SP проблем

Идея метода заключается в использовании статистических оценок функций, а также их градиентов (квазиградиентов). Другими словами, последовательность приближений к решению $\{x^k\}, k = 0,1,...$ строится с помощью таких с.в. $\eta^i(k)$ и случайных векторов $\xi^i(k)$ для $i = 0,...,m$, что условное математическое ожидание для «предыстории» $x^0,...,x^k$ имеет вид

$$E[\eta^i(k)/x^0,...,x^k] = F^i(x^k) + a^i(k), \tag{4}$$

$$E[\xi^i(k)/x^0,...,x^k] = F_x^i(x^k) + b^i(k), \tag{5}$$

где $a^i(k), b^i(k)$ – смещения оценок $\eta^i(k), \xi^i(k)$. Для точной сходимости последовательности $\{x^k\}, k = 0,1,...$ к оптимальному решению задачи $a^i(k), b^i(k)$ должны стремиться к 0 при $k \to \infty$. $\xi^i(k)$ называются стохастическими квазиградиентами.

Для негладких функций F^i условие (5) заменяется следующим

$$E[\xi^i(k)/x^0,...,x^k] \in a_k \partial F_x^i(x^k) + b^i(k).$$

SQG в простейшем случае реализуется итеративно с помощью операции проектирования на множество ограничений X как

$$x^{k=1} = \Pi_X[x^k - \rho_k \xi(k)], k = 0,1,..., \tag{6}$$

где ρ_k – позитивный размер шага, x^0 – начальное приближение.

Метод SQG генерирует последовательность аппроксимаций, сходящихся к решению с вероятностью 1 (локальных, или глобальных в выпуклом случае). Для этого размер шага ρ_k выбирается следующим образом:

$$\rho_k \geq 0, \sum_{k=0}^{\infty} \rho_k = \infty,$$

$$\sum_{k=0}^{\infty} E\left[\rho_k \| b(k) \| + \rho_k^2 \| \xi(k) \|^2\right] < \infty.$$

SQG эффективно работает и при использовании метода Монте-Карло, когда доступна только информация о текущем решении x^k и текущее наблюдение ω. Он был развит и распространен на широкий класс задач (например, минимаксных проблем, поиска глобальных оптимумов, нестационарной оптимизации, негладких и даже разрывных проблем).

2.2. Традиционные подходы для построения детерминированных эквивалентов

Непрямые методы решения задач SP сводятся к тому, что вместо исходных проблем решают их детерминированные эквиваленты.

Обратимся к проблемам SP (1). Традиционные подходы относительно решения (1) заключаются в замене ее более простыми моделями. Рассмотрим некоторые из них в соответствии с [7].

Подход* 1**: ***по наиболее вероятному сценарию (по моде). Заключается в идентификации наиболее вероятного сценария $\overline{\omega}$ и решении простой детерминированной задачи

$$\min\left\{f^0(x, \overline{\omega}) : f^i(x, \overline{\omega}) \leq 0, i = 1,..., m; x \in X\right\}. \quad (8)$$

Подход является простым, но не надежным. Любое отклонение от сценария может привести к неприятным последствиям, поскольку такие возможности в постановке просто не учитываются.

Подход* 2**: ***на наихудший случай. В таком случае проблему (1) заменяют следующей

$$\min\{\sup_{\omega \in \Omega} f^0(x, \omega) : \sup_{\omega \in \Omega} f^i(x, \omega) \leq 0, i = 1,..., m; x \in X\} \quad (9)$$

Понятно, что такой подход генерирует робастные решение для исходной задачи, однако он достаточно трудоемок и генерирует консервативные решения.

Подход* 3**: ***по средним значениям. Фактически, это сведение к проблеме (2)-(3), т.е.

$$\min\{Ef^0(x,\omega) : Ef^i(x,\omega) \le 0, i = 1,...,m; x \in X\}. \qquad (10)$$

Может быть пригодным для целевой функции, однако выглядит достаточно бессмысленным для ограничений. Решение по средним не является надежным.

Подход* 4**: ***по средним, но с учетом стандартных отклонений. Отклонения $\sigma(.)$ учитываются с помощью некоторых коэффициентов $\lambda_i>0$, $i=0,1,\ldots,m$ следующим образом

$$\min\{Ef^0(x,\omega)+\lambda_0\sigma(f^0(x,\omega)) : Ef^i(x,\omega)+\lambda_i\sigma(f^i(x,\omega)) \le 0, i = 1,...,m; x \in X\}. \quad (11)$$

Позволяет учесть отклонение $\sigma(.)$. Однако с теоретической точки зрения такой критерий нарушает упорядочение с.в. в смысле FSD. Например, может существовать решение x_1, доминирующее некоторое x_2 по распределению, $f^0(x_1,\omega) \le f^0(x_2,\omega), \omega \in \Omega$, которое может быть отброшено в пользу x_2.

Подход* 5**: ***по уровням вероятности. Выбираем некоторые уровни вероятностей $\alpha_i \in (0,1)$, $i=0,1,\ldots,m$, тогда исходная проблема трансформируется в следующую

$$\min\{y : P\{\omega : f^0(x,\omega) \le y\} \ge \alpha_0, P\{\omega : f^i(x,\omega) \le 0\} \ge \alpha_i, i = 1,...,m; x \in X^n\}. \quad (12)$$

Проблема решается с определенной вероятностью. Подход не трансформирует задачу в детерминированный эквивалент, он приводится для полноты изложения. Подход популярен, но достаточно противоречив. Не учитывается, насколько серьезны нарушения ограничений (только их вероятности), и нарушается аксиома субаддитивности (см. ниже). Не является безупречным с теоретической точки зрения.

3. Меры риска как аппарат для построения детерминированных аналогов

Универсальный подход к формулировке детерминированных эквивалентов для задачи (1) может состоять в следующем [7]. Оценим риск с.в. $X(\omega)$ некоторой детерминированной величиной, которая обычно называется мерой

риска, скажем функционалом $\rho: X \to R^+$ в пространстве с.в. Пусть определены некоторые $\rho_i(.)$, i=0,1,…,m. Тогда вместо проблемы (1) рассмотрим следующую

$$\min\{\rho^0(f^0(x,\omega)) : \rho^i(f^i(x,\omega)) \le 0, i = 1,..., m; x \in X\}. \qquad (13)$$

Замечание 1. Проблема (13) требует решения некоторых минимаксных задач, однако получаемые решения будут менее консервативны по сравнению с (9). Грубо говоря, надо перестраховываться, но не всегда на наихудший случай.

Важным методологическим вопросом при такой трансформации является выбор адекватных мер риска, который зависит от смысла исходной задачи и ее свойств. Особенное значение такой выбор приобретает при потенциально катастрофических больших ущербах. От свойств таких мер зависит и корректность получаемых в дальнейшем решений.

Например, в высоконадежных технических системах в качестве такой меры традиционно выступает вероятность отказа (аварии), в страховании – вероятность банкротства, основоположных работах по портфельной теории – дисперсия и полудисперсия [8]. Достаточно универсальной представляется мера, основанная на потенциальных ущербах. В финансах наиболее популярной мерой риска является VaR_α (Value-at-Risk) [9] (α-квантиль)

$$q_\alpha(X) = \min\{z / F_X(z) \ge \alpha\} = VaR_\alpha(X), \qquad (14)$$

которая пренебрегает малыми вероятностями на хвосте распределения. Она, к тому же, не субаддитивна, что может привести к некоторым парадоксальным эффектам. Так, может оказаться, что диверсификация портфеля может увеличить его риск (в смысле VaR).

3.1. Когерентные меры риска

В [10] были сформулированы 4 аксиомы, которым с теоретической точки зрения должна удовлетворять функция, чтобы претендовать на роль меры риска для финансовых потоков, а соответствующий класс функций был

назван классом когерентных мер риска (CRM). Напомним, что согласно аксиомам такие функции должны быть:

A1) $\rho(X+c)=\rho(X)-c$ трансляционно инвариантны;

A2) $\rho(X_1+X_2)\le\rho(X_1)+\rho(X_2)$ субаддитивны;

A3) $\rho(\lambda X)=\lambda\rho(X),\lambda\ge 0$ положительно однородны;

A4) $\rho(X_1)\le\rho(X_2),X_1\ge X_2$ (по FSD) монотонны.

Подобные меры риска описывают потенциальные убытки финансовых потоков, поэтому добавление к потоку некоторой сумму денег уменьшает его риск (потенциальные убытки) на эту сумму (A1); больший (по распределению) финансовый поток имеет меньший риск (A4).

Для случая (1), когда с.в. имеют некоторый отрицательный смысл (ущербы, затраты, пр.), такие с.в. берутся со знаком «–».

Основной результат в [10] состоял в том, что тогда CRM $\rho(.)$ с аксиомами A1)–A4) для с.в. X имеет вид:

$$\rho(X)=\sup\{E_P[-X]\,/\,P\in Q\}, \qquad (15)$$

где Q – некоторое выпуклое замкнутое множество вероятностных мер. Точное описание Q однозначно определяет CRM. Нетрудно видеть, что выпуклая комбинация и взятие max являются замкнутыми операциями на классе CRM. Для случая (1), когда с.в. X имеют некоторый отрицательный смысл, они берутся со знаком «–». Поэтому подобная мера для них определяется как:

$$\rho(X)=\sup\{E_P[X]\,/\,P\in Q\}. \qquad (16)$$

Возвращаясь к разделу 2.2, отметим, что меры в подходах 1–3 являются CRM, в 4–5 – нет. При переходе к проблеме (11) нарушается монотонность, к (12) – субаддитивность.

3.2. CVaR

В [11] была предложена мера *CVaR* (conditional VAR), которая интерпретируется как интегральный VAR (интеграл по соответствующему хвосту распределения) и попадает в класс CRM

$$CVaR_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_\beta(X)d\beta$$ (матожидание X на α-хвосте).

Понятно, что $CVaR_\alpha \to \sup X$ при $\alpha \to 1$, и $CVaR_\alpha \to E[X]$ при $\alpha \to 0$.

Основной результат для вычисления *CVaR* имеет вид [11–12]:

$$CVaR_\alpha(X) = \min_{\zeta \in R}\left\{\zeta + \frac{1}{1-\alpha}E\left[-X-\zeta\right]^+\right\}, \qquad (17)$$

Можно рассматривать различные выпуклые комбинации *CVaR* с весами. Например, взять неотрицательную меру взвешивания λ на (0,1), тогда когерентной будет мера

$$\int_0^1 CVaR_\alpha(X)d\lambda(\alpha).$$

3.3 Полиэдральные когерентные меры риска

3.3.1 Случай известных распределений с.в.

Полиэдральной когерентной мерой риска (PCRM) были названы функции вида (15), где замкнутое выпуклое множество вероятностных мер Q имеет вид выпуклой оболочки конечного числа точек [13]. Для случая (1), когда с.в. имеют некоторый отрицательный смысл, представление меры (15) заменяется соотношением (16).

Рассмотрим случай дискретно распределенных с.в. X, которые представляются в виде вектора значений $x = (x_1,...,x_n)$ и вектора сценарных вероятностей $p = (p_1,...,p_n)$. Тогда такое множество таких вероятностных мер Q имеет вид

$$Q = \mathrm{co}\{p_i\colon i=1,\ldots,k\},$$

или, эквивалентно,

$$Q = \{p\colon B\,p \le c,\, p \ge 0\}, \qquad (18)$$

где B и c – матрица и вектор соответствующих размерностей.

Замечание 2. Поскольку Q является множеством вероятностных мер, поэтому ее описание в (18) должно включать стандартное условие $\sum_{i=1}^{n} p_i = 1$, которое выписывается в (18) в виде двух следующих неравенств: $\sum_{i=1}^{n} p_i \le 1$ и $-\sum_{i=1}^{n} p_i \le -1$.

Разделим теперь описание Q в (18) на стандартную (обязательную) и содержательную части. Представим матрицу B и вектор c как

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \tag{19}$$

где B_1 и c_1, описывающие упомянутые выше неравенства, являются стандартными:

$$B_1 = \begin{pmatrix} 1 \dots 1 \\ -1 \dots -1 \end{pmatrix}, c_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{20}$$

а B_2 и c_2, собственно, описывают содержательную часть в соотношении (19), которая определяет саму меру риска из (15), (18) (или (16), (18)).

Примеры PCRM:

1) Наихудший случай **WCR**

$$Q_{WCR} = \{p = (p_1, \dots, p_n) : \sum_1^n p_i \le 1, -\sum_1^n p_i \le -1, p_i \ge 0, i = 1, \dots, n\},$$

т.е. в (19) содержится только стандартная часть B_1 и c_1.

2) **CVaR$_\alpha$**

$$Q_{CVaR_\alpha} = \{p = (p_1, \dots, p_n) : \sum_1^n p_i \le 1, -\sum_1^n p_i \le -1, Ip \le \frac{p_0}{(1-\alpha)}, p_i \ge 0, i = 1, \dots, n\},$$

где I – единичная матрица, а p_0 – вектор сценарных вероятностей;

3) Наихудшее условное среднее **WCE$_\alpha$** [10];

4) Спектральная КМР **SCRM** [14], представленная как выпуклая комбинация мер *CVaR*;

5) Мера по полуотклонению от среднего с.в. [15] $\delta_S(x; r) = -E[x] + r\,E[(E[x] - x)^+]$;

6) Мера по абсолютному отклонению от среднего с.в. [15] $\delta_A(x; r) = -E[x] + r\,E[\,|x-E[x]|\,]$.

Кроме того, класс PCRM является инвариантным относительно операций: 1) выпуклой комбинации; 2) функции максимума; 3) инфимальной конволюции.

Нетрудно видеть, что

$$\mathrm{CRM} \supset \mathrm{PCRM} \supset \mathrm{CVaR}_{\alpha}\,.$$

3.3.2. Случай неполной информации о сценарных вероятностях

Хотя аппарат PCRM не предлагает использовать меру конкретного вида**,** перекладывая такой выбор на ЛПР, он позволяет рассматривать проблемы достаточно широкого класса, например, учитывать частичную неопределенность проблемы, т.е. условия, в которых распределения с.в. не известны точно.

Рассмотрим более точное определение PCRM из [16]:

$$\rho(X) = \sup\{E_P[-X] : P \in a(P_0)\}, \tag{21}$$

где многозначное отображение (м.о.) $a(.)$ с выпуклыми многогранными образами определяет меру риска $\rho(.)$ в зависимости от вероятностной меры P_0. Понятно, что если $a(P_0) \equiv P_0$, то $\rho(.)$ сводится к среднему значению.

В случае, когда распределение, т.е. вероятностная мера P_0 известна и фиксирована, отличия между (15) и (21) нет, поскольку $a(P_0) = Q$. Но ситуация меняется в случае неопределенности, когда P_0 – не известна точно, а доступна лишь информация

$$P_0 \in P_U\,, \tag{22}$$

где P_U – некоторое множество вероятностных мер, характеризующих неопределенность. Тогда соответствующая мера риска определяется [16] как

$$\rho_U(X) = \sup\{E_P[-X] : P \in \overline{co}\; a(P_U)\}, \tag{23}$$

где $\overline{co}$ означает выпуклую замкнутую оболочку.

Пример 1: неточные вероятности, риск оценивается средним

Пусть имеются лишь покоординатные оценки сверху и снизу для вектора p_0 сценарных вероятностей, а именно

$$P_U = \{p : p_l \le p \le p_u\} \Leftrightarrow \begin{pmatrix} -Ip \le -p_{lw} \\ Ip \le p_{up} \end{pmatrix}. \tag{24}$$

Рассмотрим простой случай, когда для известных распределений риск оценивается средними значениями, т.е. $a(P_0) \equiv \{P_0\}$. В таком случае $Q = P_U$. Представим эти множества в форме (18), из технических соображений изменив обозначение матрицы B и вектора c на A и b соответственно:

$$P_U = \{p : Ap \le b, p \ge 0\}, \tag{25}$$

где

$$A = \begin{pmatrix} B_1 \\ -I \\ I \end{pmatrix}, b = \begin{pmatrix} c_1 \\ -p_{lw} \\ p_{up} \end{pmatrix}, \tag{26}$$

а стандартная часть в виде B_1 и c_1 описывается соотношением (20).

Следовательно, это PCRM вида (15), (18), в которой (18) описывается с помощью (25), (20) и (26).

Пример 2: неточные вероятности, риск оценивается CVaR

Как и ранее, имеются лишь оценки вектора p_0 сценарных вероятностей (24). Оценим риск с помощью CVaR при неточных вероятностях.

В соответствии с (21) и описанным ранее множеством Q_{CVaR_α}, изучаемое м.о. $a(.)$ имеет вид

$$a(p_0) = \left\{ p : \sum_1^n p_i \le 1, -\sum_1^n p_i \le -1, Ip \le \frac{p_0}{(1-\alpha)}, p \ge 0 \right\}.$$

Следовательно, учитывая (24), имеем

$$a(P_U) = \left\{ p : \sum_1^n p_i \le 1, -\sum_1^n p_i \le -1, Ip \le \frac{p_{up}}{(1-\alpha)}, p \ge 0 \right\}. \tag{27}$$

Это выпуклое замкнутое множество, следовательно, знак $\overline{co}$ можно опустить. Переписывая множество в стандартном для PCRM виде, получим форму (25), где B_1 и c_1 описываются посредством (20), а B_2 и c_2 как

$$B_2 = I, c_2 = \frac{p_{up}}{(1-\alpha)}. \tag{28}$$

Замечание 3. Сценарии-события могут быть альтернативными, но не независимыми. Например, они рассчитываются как вероятности в модели некоторой последовательности событий. Тогда между ними могут возникать некоторые соотношения, например, в форме линейных равенств-неравенств. В этом случае они формируют в PCRM описание соответствующих м.о. $a(.)$. Множества неопределенностей P_U также могут быть более сложными, чем оценки вероятностей сверху и снизу, например, в виде линейных неравенств.

4. Оптимизация портфеля с PCRM

Вообще говоря, аппарат мер риска описывался достаточно подробно для того, чтобы обосновать содержательность трансформации исходной проблемы (1) к ее детерминированному аналогу (13) с помощью соответствующих мер риска.

Например, для класса PCRM тогда имеем

$$\min\{\rho^0(f^0(x,\omega)) : \rho^i(f^i(x,\omega)) \le 0, i = 1,...,m; x \in X\}, \tag{29}$$

где

$$\rho^i(f^i(x,\omega)) = \sup\{E_P[f^i(x)] : P \in \overline{co}\, a(P_U)\}, i = 0,...,m. \tag{30}$$

Напомним, что в случае известных распределений с.в. множество неопределенности P_U состоит из единственной вероятностной меры, т.е. $P_U = \{P_0\}$.

Как нетрудно видеть, проблема (29) – (30) является некоторым вариантом минимаксной задачи. Она может быть сведена к проблеме линейного программирования (ПЛП), если функции $f^i(x)$ являются линейными (кусочно-линейными). Такой, к примеру, является задача оптимизации портфеля.

Пусть распределение прибыльности компонент портфеля z_j, j=1,…, k описывается матрицей H размерности $n \times k$, где j-й столбец описывает распре-

деление j – ой компоненты. Вектор $u = (u_1, \ldots, u_k)$, который описывает структуру портфеля рассматривается как переменная, где $\sum_{i=1}^{k} u_i = 1$, $u_i \geq 0$, i=1,…, k.

Нужно найти такую структуру портфеля u, которая оптимизирует совокупный результат портфеля по соотношению доходность-риск. В случае известных распределений компонент (в такой постановке это описывается матрицей H распределений компонент согласно сценариям и известным вектором $p_0 = (p_1^0, \ldots, p_n^0)$ вероятностей сценариев), в качестве доходности рассматривается средняя прибыльность, а риска – некоторая PCRM. Рассмотрим две взаимосвязанные постановки задачи.

4.1 Оптимальные портфели при известных распределениях

Минимизация меры риска при гарантированной средней доходности. Зафиксируем нижнее допустимое значение средней эффективности $E_p[Hu]$, которую должен гарантировать исследуемый портфель, величиной μ в виде ограничений и минимизируем его меру риска $\rho(Hu)$:

$$\begin{aligned} &\min \rho(Hu), \\ &\sum_{i=1}^{n} u_i = 1,\ u_i \geq 0, \\ &E_{p_0}[Hu] \geq \mu. \end{aligned} \tag{31}$$

Максимизация средней доходности при ограничении на меру риска

Зафиксируем определенный уровень меры риска $\rho(Hu)$, которую не должен превышать исследуемый портфель, величиной σ в виде ограничений и максимизируем его среднюю прибыльность $E_p[Hu]$:

$$\begin{aligned} &\max E_{p_0}[Hu], \\ &\sum_{i=1}^{n} u_i = 1,\ u_i \geq 0, \\ &\rho(Hu) \leq \sigma. \end{aligned} \tag{32}$$

Сравнивая с общей постановкой (29) – (30), нетрудно заметить, что функции $f^i(.)$, i=0,1 здесь линейны вида Hu, причем в качестве второй меры риска

здесь выбрана обычная средняя величина, а PCRM строится по ущербам в виде (15), (18).

Теорема 1 [13]. *Решением проблем* (31), (15), (18) *и* (32), (15), (18) *есть соответственно компоненты и решений* (*v, u*) *следующих ПЛП*

$$\begin{gathered}\min\limits_{(v,u)} < c, v >, \\ -B^T v - Hu \le 0, \\ Au \le b, \\ v \ge 0,\ u \ge 0.\end{gathered} \tag{33}$$

$$\begin{gathered}\max\limits_{(v,u)} < H^T p_0, u >, \\ -B^T v - Hu \le 0, \\ < c, v > \le \sigma \\ \sum u_i = 1, \\ v \ge 0,\ u \ge 0\end{gathered} \tag{34}$$

где $A = \begin{pmatrix} B_1 \\ -p_0^T H \end{pmatrix}$, $b = \begin{pmatrix} c_1 \\ -\mu \end{pmatrix}$, $B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$, $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, а B_1 и c_1 являются стандартными из (20).

Замечание 4. Ранее близкие результаты были получены для CVaR [11, 12], используя его представление из (17).

Можно рассмотреть задачу вида (33) при нескольких ограничениях на соответствующие меры риска, например, в следующей форме

$$\begin{gathered}\max E_{p_0}[Hu], \\ \sum_1^n u_i = 1,\ u_i \ge 0, \\ \rho_j(Hu) \le \sigma_j,\ j = 1, \ldots, m,\end{gathered} \tag{35}$$

где каждая из ρ_j есть PCRM вида (15), (18) со своими B_j и c_j в (18).

Эта проблема также сводится к ПЛП. Так, решением проблемы (35) есть компонента *u* решения ($v_1, \ldots, v_m, u$) следующей ПЛП

$$\begin{array}{l} \max\limits_{(v_1,\ldots,v_m,u)} < H^T p_0, u >, \\ -B_1^T v_1 - Hu \le 0, \\ < c_1, v_1 > \le \sigma_1, \\ -B_m^T v_m - Hu \le 0, \\ < c_m, v_m > \le \sigma_m, \\ \sum_1^n u_i = 1,\ u \ge 0, \\ v_1 \ge 0, \ldots, v_m \ge 0. \end{array} \tag{36}$$

4.2. Оптимальные портфели при неточных вероятностях

Портфельные проблемы (31), (32), (35) легко переформулируются в условиях неполной информации о сценарных вероятностях. Приведем в качестве примера их переформулировку при неточных вероятностях.

При неточных вероятностях множество P_U, как известно, описывается соотношениями (25), (26), (20). Пусть, как и ранее, некоторые матрица B и вектор c описывают меру риска, полученную с помощью конструкции (23), (25), (26). Хотя, вообще говоря, представление ее как PCRM возможно не всегда. Например, в том случае, когда в качестве исходной мерой риска был CVaR, они представлены как

$$B = \begin{pmatrix} B_1 \\ I \end{pmatrix}, c = \begin{pmatrix} c_1 \\ \dfrac{p_{up}}{(1-\alpha)} \end{pmatrix}, \tag{37}$$

где стандартная часть в виде B_1 и c_1 представлена в (20). Обозначим как $\rho_U(.)$ меру риска, построенную по исходной $\rho(.)$ с помощью конструкции (21), (22), (23), тогда соответствующие проблемы имеют вид

$$\begin{array}{l} \min\limits_{\sum\limits_{i=1}^n u_i = 1,\ u_i \ge 0} \rho_U(Hu), \\ \min\limits_{Ap \le b,\ p \ge 0} < p, Hu > \ge \mu, \end{array} \tag{38}$$

$$\begin{array}{l} \max\limits_{\sum\limits_{i=1}^n u_i = 1,\ u_i \ge 0} \min\limits_{Ap \le b, p \ge 0} < Hu, p >, \\ \rho_U(Hu) \le \sigma, \end{array} \tag{39}$$

где матрица A и вектор b описаны в соотношении (26).

Если мера $\rho_U(.)$ может быть представлена в виде ПКМР в форме (15), (18), проблемы (38) и (39) можно свести к соответствующим ПЛП. Напомним, что для случая, когда исходной мерой был CVaR, это имеет место, а матрица B и c из (18) описаны соотношением (37).

Теорема 2. *Решением проблем* (38) *и* (39) *с учетом обозначений* (26) *и* (18) *есть соответственно компоненты и решений (v, u, w) следующих ПЛП*

$$\begin{gathered}\min_{(v,u,w)} < c, w >, \\ -A^T v - Hu \le 0, \\ < b, v > \le -\mu, \\ -B^T w - Hu \le 0, \\ \sum_1^n u_i = 1, u \ge 0, \\ v \ge 0, w \ge 0.\end{gathered}$$

$$\begin{gathered}\max_{(v,u,w)} < -b, v >, \\ -A^T v - Hu \le 0, \\ -B^T w - Hu \le 0, \\ < c, w > \le \sigma, \\ \sum_1^n u_i = 1, u \ge 0, \\ v \ge 0, w \ge 0.\end{gathered}$$

Нетрудно также в условиях неточных вероятностей переформулировать проблему (35) и выписать соответствующую ей ПЛП.

Замечание 5. Зачастую в различных портфельных постановках появляются некоторые технические ограничения, накладываемые на потенциальные возможности выбора портфельных компонент u_i. Они формулируются в виде линейных неравенств, поэтому их добавление в постановки проблем осуществляется абсолютно естественно.

5. Понятие робастного решения и мера риска

Робастность решения означает его устойчивость к всевозможным возмущениям и неопределенностям. Смысл такого решения заключается в том, что не стоит искать оптимальные решения для каждого из возможных сценариев развития будущих событий, важно найти решение, которое является хо-

рошим по сравнению с альтернативами на широком диапазоне вероятных будущих сценариев (см., например, [17]).

Такое понятие крайне полезно и в разнообразных оптимизационных постановках. Например, зачем искать точные оптимальные решения, которые возникающие на практике возмущения, флуктуации или отказы могут сделать бессмысленными? Понятно, что не стоит сначала искать оптимальные решения, которые затем тестировать на устойчивость (робастность). Нужно искать эффективные решения среди устойчивых (робастных). Для этого, естественно, необходима некоторая мера, описывающая свойство робастности решений при потенциальных возмущениях, на всем множестве вероятных сценариев, прочее. В случае, когда решения принимаются в условиях риска и неопределенности, в этом качестве вполне адекватно может выступать мера риска.

Закладывая в выбранную меру риска все потенциальные неопределенности и решая задачу вида (13), мы можем гарантировать робастность относительно них получаемых оптимальных (эффективных) решений.

Отметим, что выбор мер риска и исходных функций в оптимизируемых проблемах существенно зависит от специфики таких задач. Для разнообразных финансовых приложений основным индикатором риска является потенциальный ущерб, при этом доходность (убыточность) финансового портфеля линейно зависит от доходности его компонент. Поэтому для таких проблем вполне адекватным является использование аппарата и результатов разделов 3 – 4.

В экономических приложениях, когда в качестве критериев служат некоторые функции полезности (эффективности), меры риска следует строить по таким функциям. Например, если имеется некоторая функция полезности $UF(.)$, то соответствующая CRM будет иметь вид

$$\rho(UF(x,\omega)) = \sup\{E_p[-UF(X)] : p \in Q\}. \qquad (40)$$

Как показано в [18], в случае, когда функция *UF*(.) из (40) представляется в кусочно-линейном виде, а мера риска является PCRM, соответствующие задачи оптимизации портфеля могут быть сведены к ПЛП.

Список литературы

[1] Proske D. Catalogue of Risks. Natural, Technical, Social and Health Risks. – Berlin: Springer, 2008. – 509 p.

[2] Knight F.H. Risk, Uncertainty and Profit. – Houghton Miffin: Boston, 1921 (Risk, Uncertainty and Profit**,** By Frank H. Knight, 2002/04 – Beard Books 1587981262 – Paperback – Reprint – 447 p.).

[3] Ермольев Ю.М. Методы стохастического программирования. – М.: Наука, 1976. – 240с.

[4] Ermoliev Yu.M., Wets R. (eds.) Techniques for Stochastic Optimization. – Springer: Berlin, 1988. – 571 p.

[5] Кнопов П.С., Сергиенко И.В. О некоторых научных результатах Ю.М. Ермольева и его школы в области современной теории оптимизации // Кибернетика и системный анализ. – 2011. – № 6. – С. 3 – 27.

[6] Ермольев Ю.М., Норкин В.И. Методы решения невыпуклых негладких задач стохастической оптимизации // Кибернетика и системный анализ. – 2003. – № 5. – С. 89 – 106.

[7] Rockafellar R.T. The Fundamental Quadrangle of Risk in Optimization and Estimation, Lecture at Risk Workshop, University of Florida, April 2010. www.ise.ufl.edu/rmfe/seminar/rockafellar/rock_2010_spring_workshop.html.

[8] Markowitz H.M. Portfolio Selection, Efficient Diversification of Investment. – Wiley: NY. – 1959. – 344 p.

[9] Jorion P.H. Value at Risk: A New Benchmark for Measuring Derivatives. – New York: Irwin Professional Publishers, 1996. – 284 p.

[10] Artzner P., Delbaen F., Eber J.-M., Heath D. Coherent Measures of Risk // Mathematical Finance. – 1999. – 9/3. – Pp. 203 – 228.

[11] Rockafellar R.T., Uryasev S. Optimization of Conditional Value-at-Risk // The Journal of Risk. – 2000. – 2. – Pp. 21–41.

[12] Rockafellar R.T., Uryasev S. Conditional Value-at-Risk for General Loss Distribution // Journal of Banking and Finance. – 2002. – 26. – Pp. 1443 – 1471.

[13] Кирилюк В.С. О классе полиэдральных когерентных мер риска // Кибернетика и системный анализ. – 2004. – № 4. – С. 155 – 167.

[14] Acerbi C. Spectral Measures of Risk: a Coherent Representation of Subjective Risk Aversion // Journal of Banking and Finance. – 2002. – 26(7). – Pp. 1505 – 1518.

[15] Ogryczak W., Ruszczynski A. From Stochastic Dominance to Mean-Risk Models: Semideviation as Risk Measures // European Journal of Operation Research. – 1999. – 116. – Pp. 33 – 50.

[16] Кирилюк В.С. Полиэдральные когерентные меры риска и оптимизация инвестиционного портфеля // Кибернетика и системный анализ. – 2008. – № 2. – С. 120 – 133.

[17] Marti K., Ermoliev Yu., Makowski M. Coping with Uncertainty. Robust Solutions. – Berlin: Springer, 2010. – 277 p.

[18] Кирилюк В.С. Некоторые робастные решения в условиях риска и неопределенности // Теорія оптимальних рішень. – 9/2010. – С. 54 – 61.

УДК 519.21

МЕТОД ЭМПИРИЧЕСКИХ СРЕДНИХ В ЗАДАЧАХ СТОХАСТИЧЕСКОЙ ОПТИМИЗАЦИИ И ОЦЕНИВАНИЯ

Кнопов П.С.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: Knopov@gmail.com

Аннотация. При решении задач стохастической оптимизации и идентификации не всегда имеется возможность найти точный экстремум математического ожидания некоторой случайной функции. Одним из способов решения возникающей проблемы является метод эмпирических средних, состоящий в аппроксимации имеющейся критериальной функции ее эмпирической оценкой, для которой существует возможность решения соответствующей оптимизационной задачи. Естественно, что условия сходимости существенно зависят от функции критерия, вероятностных свойств случайных наблюдений, метрики пространств, для которых будет исследоваться сходимость, априорных ограничений на неизвестные параметры и т.д. В терминологии теории статистических решений эти вопросы тесно связаны с асимптотическими свойствами оценок неизвестных параметров, а именно, состоятельностью, асимптотическим распределением, скоростью сходимости оценок и т.д. Следует отметить, что методу эмпирических средних посвящено довольно большое число публикаций, среди которых отметим работы [1 – 14]. Одним из наиболее важных является подход, основанный на так называемой epi-сходимости, базой которого являются условия сходимости, основанные на понятии epi-расстояния [2 – 7]. В предлагаемой работе рассмотрим другие подходы к доказательству утверждений о сходимости, приведенные в работах [15 – 20]. Предлагаемая статья носит, в основном, обзорный характер, поэтому доказательства мы будем опускать, отсылая читателя к соответствующим первоисточникам.

Ключевые слова. Оценивание, оптимизация, сходимость, состоятельность, распределение, большие уклонения.

Сходимость метода эмпирических средних в случае независимых и слабо зависимых наблюдений

Начнем с простейшей, но важной задачи стохастического программирования, на которой проиллюстрируем основные подходы для решения задачи. Затем будем рассматривать более сложные модели, для исследования которых будут использованы аналогичные методы.

Итак, рассмотрим модель с независимыми наблюдениями. Пусть $\{\xi_i, i \in N\}$ – независимые наблюдения случайной величины, заданной на вероятностном пространстве $(\Omega, \Im, P)$ со значениями в некотором метрическом пространстве $(Y, L(Y))$, $L(Y)$ – минимальная σ-алгебра на Y. Предположим, что I – замкнутое подмножество в R^l, $l \geq 1$, возможно $I = R^l$. Предпо-

ложим, что $f: I \times Y \to R^+$ – неотрицательная функция, удовлетворяющая условиям:

1) $f(\vec{u}, z)$, $\vec{u} \in I$, непрерывна для всех $z \in Y$;

2) для любого $\vec{u} \in I$ отображение $f(\vec{u}, z)$, $z \in Y$, является $L(Y)$-измеримым.

Проблема состоит в нахождении точки минимума функции

$$F(\vec{u}) = E\big(f(\vec{u}, \xi)\big),\ \vec{u} \in I \tag{1}$$

и ее минимального значения.

Эта задача аппроксимируется следующей: найти точки минимума функции

$$F_n(\vec{u}) = \frac{1}{n}\sum_{i=1}^{n} f(\vec{u}, \xi_i) \tag{2}$$

и ее минимального значения. Имеет место следующее утверждение.

Теорема 1. *Пусть выполнены следующие условия:*

1) для любого $c > 0$ $E\left(\max\limits_{\|\vec{u}\| \le c} f(\vec{u}, \xi)\right) < \infty$, *где* $\|\cdot\|$ *– норма в* R^l;

2) для всех $z \in Y'$, $P\{\xi \in Y'\} = 1$, $f(\vec{u}, z) \to \infty$ *при* $\|z\| \to \infty$;

3) существует единственная точка $\vec{u}_0$, *для которой функция* $F(\vec{u})$ *достигает минимума.*

Тогда для каждого n *и* $\omega \in \Omega'$, $P(\Omega') = 1$, *существует по крайней мере один вектор* $\vec{u}_n = \vec{u}_n(\omega) \in I$, *при котором достигается минимальное значение* $F_n(\vec{u})$, *и для любого* $n \ge 1$ $\vec{u}_n$ *можно выбрать* G_n'*-измеримым, где* $G_n' = G_n \cap \Omega'$, $G_n = \sigma\left\{\xi_i, i = \overline{1,n}\right\}$. *При этом с вероятностью 1* $\vec{u}_n \to \vec{u}_0$, $F(\vec{u}_n) \to F(\vec{u}_0)$.

Доказательство этого утверждения приведено в [8, 9, 11, 13] и опирается на фундаментальные результаты, полученные в [15 – 20], где предложен универсальный подход для доказательства свойств состоятельности и сильной состоятельности оценок неизвестных параметров случайных величин со

значениями в произвольном метрическом пространстве. Приведем формулировку этого замечательного результата, придерживаясь работы [19].

Теорема 2. *Пусть* $(\Omega, \Im, P)$ *– некоторое заданное вероятностное пространство и* $\{\Im_n, n \geq 1\}$ *– неубывающий поток* σ*-алгебр такой, что* $\Im_n \subset \Im_{n+1}$, $\Im_n \subset \Im$, $n \geq 1$. *Предположим, что* K *– компактное подмножество некоторого банахова пространства с нормой* $\|\cdot\|$. *Пусть* $Q_n(s, \omega)$ *– последовательность действительных функций, удовлетворяющих следующим условиям:*

1) для всех n *и* $s \in K$ *функция* $Q_n(s, \omega)$, $\omega \in \Omega$, *является* $\Im_n$*-измеримой;*

2) для всех n *и* ω *функция* $Q_n(s, \omega)$, $s \in K$, *является непрерывной на* K*;*

3) для некоторого фиксированного $s_0 \in K$ *и всех* $s \in K$ *имеет место*

$$P\{\lim_{n \to \infty} Q_n(s, \omega) = \Phi(s, s_0)\} = 1,$$

где $\Phi(s, s_0)$, $s \in K$, *– непрерывная на* K *функция, причем* $\Phi(s, s_0) > \Phi(s_0, s_0)$, $s \neq s_0$*;*

4) существуют величина $\gamma_0 > 0$ *и функция* $c(\gamma)$, $\gamma > 0$, $c(\gamma) \to 0$ *при* $\gamma \to 0$, *такая, что для любых* $s' \in K$, $\delta > 0$ *и* $0 < \gamma < \gamma_0$ *имеет место соотношение*

$$P\left\{\lim_{n \to \infty} \sup_{\substack{\|s - s'\| < \gamma \\ \|s - s_0\| > \delta}} \left|Q_n(s) - Q_n(s')\right| < c(\gamma)\right\} = 1.$$

Пусть также для любых $n \geq 1$ *и* $\omega \in \Omega$ *величина* $s_n = s_n(\omega) \in K$ *определена соотношением* $Q_n(s_n) = \min_{s \in K} Q_n(s)$.

Тогда элемент s_n *можно выбрать таким образом, что для любой действительной непрерывной функции* h *на* K *функцию* $h(s_n(\omega))$, $\omega \in \Omega$, *можно выбрать* $\Im_n$*-измеримой, и справедливы следующие соотношения*

$$P\left\{\lim_{n \to \infty} \|s_n - s_0\| = 0\right\} = 1.$$

Если условия 3) и 4) теоремы 2 выполняются по вероятности, имеет место утверждение о слабой или просто состоятельности оценок $s_n(\omega)$.

Теорема 3 [19]**.** *Пусть выполнены условия 1) и 2) теоремы 2, а сходимость в условиях 3) и 4) имеет место по вероятности. Тогда элемент* $s_n(\omega)$ *можно выбрать таким образом, что для любой действительной непрерывной функции* h *на* K *функцию* $h\left(s_n(\omega)\right)$, $\omega \in \Omega$, *можно выбрать* $\Im_n$*-измеримой, и справедливы следующие соотношения*

$$\lim_{n\to\infty} P\left\{\left\|s_n - s_0\right\| > \varepsilon\right\} = 0 \text{ для любого } \varepsilon > 0,$$

$$\lim_{n\to\infty} P\left\{\left|Q_n(s_n) - \Phi(s_0, s_0)\right| > \varepsilon\right\} = 0 \text{ для любого } \varepsilon > 0.$$

Замечание 1. Теоремы 2, 3 имеют место в случае, когда функция $\Phi(s, s_0)$ непрерывна на K. Исключая, возможно, некоторую окрестность точки s_0.

Замечание 2. Теоремы 2, 3 имеют место и в случае, когда вместо n рассматривается некоторый непрерывный параметр.

Теорема 2 указывает путь к доказательству теоремы 1 кратко остановимся на его основных моментах.

1. Принимая во внимание условие 2) теоремы 1 можно показать, что существует $c > 0$ такая, что с вероятностью 1 все $\vec{u}_n$ начиная с некоторого $n \geq 1$ принадлежат компактному множеству $K = \{\vec{u} \in I : \|\vec{u}\| \leq c\}$.

2. Исходя из эргодической теоремы, нетрудно видеть, что $P\{\lim_{n\to\infty} F_n(\vec{u}) = F(\vec{u})\} = 1$.

3. Выбирая $c(\gamma) = 2E\left\{\sup_{\{\vec{u},\vec{u}':\|\vec{u}-\vec{u}'\|<\gamma\}} \left|f(\vec{u}, s) - f(\vec{u}', s)\right|\right\}$ нетрудно показать, что $P\left\{\lim_{n\to\infty} \sup_{\|\vec{u}-\vec{u}'\|<\gamma} \left|F_n(\vec{u}) - F_n(\vec{u}')\right| < c(\gamma)\right\} = 1$. Поэтому условие 4) теоремы 2 тоже выполнено. Применяя теорему 2 немедленно получаем утверждение теоремы 1.

Замечание 3. Используя эргодическую теорему, можно показать, что утверждение теоремы 1 имеет место и в случае, если последовательность $\{\xi_i, i \in N\}$ является стационарной в узком смысле эргодической случайной последовательностью. Иными словами, задача

$$F(\vec{u}) = E\big(f(\vec{u},\xi)\big) \to \min_{\vec{u} \in I}$$

аппроксимируется в этом случае задачей

$$F_n(\vec{u}) = \frac{1}{n}\sum_{i=1}^{n} f(\vec{u},\xi_i) \to \min_{\vec{u} \in I},$$

и, используя эргодическую теорему, получаем результат, аналогичный теореме 1.

В случае непрерывного времени также имеет место утверждение, аналогичное теореме 1. А именно.

Теорема 4 [8, 9]. *Пусть* $\{\xi(t), t \in R\}$ *– стационарный в узком смысле случайный процесс, определенный на вероятностном пространстве* $(\Omega, \Im, P)$*, и выполнены следующие условия:*

1) *для любого* $c > 0$ $E\left\{\max_{\|\vec{u}\| \le c} f(\vec{u},\xi(0))\right\} < \infty$*;*

2) *если* I *–* *неограниченное множество для любого* $z \in Y'$*,* $P\{\xi(t) \in Y', \forall t \ge 0\} = 1$*, то* $f(\vec{u},z) \to \infty$ *при* $\|\vec{u}\| \to \infty$*;*

3) *существует единственный элемент* $\vec{u}_0 \in I$*, для которого достигается минимальное значение функции* $F(\vec{u}) = Ef(\vec{u},\xi(0))$.

Тогда для всех $T > 0$ *и* $\omega \in \Omega'$*,* $P(\Omega') = 1$*, существует по крайней мере один вектор* $\vec{u}(T) \in I$*, при котором достигается минимальное значение функции*

$$F_T(\vec{u}) = \frac{1}{T}\int_0^T f(\vec{u},\xi(t))dt,$$

причем для каждого $T>0$ *величина* $\vec{u}(T)$ *будет* G'_T*-измеримой, где* $G'_T = G_T \bigcap \Omega'$, $G_T = \sigma\{\xi(t), 0 \le t \le T\}$.

Пусть $\vec{u}_0 = \operatorname{argmin} F(\vec{u})$, *где* $F(\vec{u}) = Ef(\vec{u}, \xi(0))$. *Тогда*

$$P\{\lim_{T\to\infty} \vec{u}(T) = \vec{u}\} = 1$$

$$P\{\lim_{T\to\infty} F_T(\vec{u}_T) = F(\vec{u}_0)\} = 1.$$

Доказательство проводится аналогично случаю дискретного времени, принимая во внимание замечание 1.

Дальнейшее обобщение утверждений о сильной состоятельность или сходимости с вероятностью 1 аппроксимирующей задачи стохастического программирования к первоначальной состоит в рассмотрении вместо $f(u,\xi)$ функций более общего вида. А именно будем предполагать, что имеется функция вида

$$\widetilde{F}_n(\vec{u}) = \frac{1}{n}\sum_{i=1}^{n} f(i, \vec{u}, \xi_i),$$

где ξ_i – некоторые, вообще говоря, зависимые наблюдения, на которые накладываются определенные ограничения на характер зависимости. Сформулируем утверждение более точно.

Теорема 5 [9, 13, 14]. *Пусть случайная функция* $f(i, \bar{u}, \zeta_i)$ *удовлетворяет следующим условиям:*

1) для любого $\vec{u} \in I$ *существует функция* $F(x)$ *такая, что* $\widetilde{F}(\vec{u}) = \lim\limits_{\|u\|\to\infty} \widetilde{F}_n(\vec{u})$ *и точка* $\bar{u}_0 \in I$ *такая, что* $\widetilde{F}(\vec{u}_0) < \widetilde{F}(\vec{u})$ *при* $\bar{u} \ne \bar{u}_0$;

2) функция $f(i,\vec{u},z)$ *является неотрицательной, непрерывной по второму аргументу равномерно по* i *и* z;

3) если множество I *неограниченно, то* $f(i,\vec{u},z) \to \infty$ *при* $\|u\| \to \infty$, *при фиксированных* i *и* z;

4) существует функция $c(\gamma) \to 0$ *при* $\gamma \to 0$ *и для любого* $\delta > 0$ *существует* γ_0 *такое, что для любого элемента* $u \in I$ *и* $0 < \gamma < \gamma_0$ *имеет место соотношение* $\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} E \sup_{\substack{\|\bar{u}-\bar{u}'\|<\gamma \\ \|\bar{u}-\bar{u}_0\|<\delta}} \left| f(i,\bar{u},\xi_i) - f(i,\bar{u}',\xi_i) \right| < c(\gamma)$;

5) функция $f(i,\bar{u},\xi_i)$ *удовлетворяет условию сильного перемешивания*

$$\alpha(j) = \sup_i \sup_{\substack{A \in \sigma_{-\infty}^{i} \\ B \in \sigma_{i+j}^{\infty}}} \left| P(AB) - P(A)P(B) \right| \le \frac{c}{1 + j^{1+\varepsilon}}, \varepsilon > 0,$$

$$\sigma_n^m = \sigma\left\{ f(i,\bar{u},\xi_i), n \le i \le m, \bar{u} \in I \right\};$$

6) $E\left(f(i,\bar{u},\xi_i)\right)^{2+\delta} < \infty,\ \varepsilon\delta > 2.$

Пусть $\bar{u}_n = \underset{\bar{u}\in I}{\operatorname{argmin}} F_n(\bar{u})$. Тогда

$$P\{ \lim_{n\to\infty} \left\| \bar{u}_n - \bar{u}_0 \right\| = 0 \} = 1,$$

$$P\{ \lim_{n\to\infty} \tilde{F}_n(\bar{u}_n) = \tilde{F}(\bar{u}_0) \} = 1.$$

Аналогичное утверждение справедливо и для случая, когда имеется непрерывная реализация наблюдения случайной функции $f(t,\bar{u},\xi(t))$ на интервале $[0,T]$, т.е. рассматривается функционал $F_T(\bar{u}) = \frac{1}{T}\int_0^T f(t,\bar{u},\xi(t))dt$, где $\xi(t)$ стационарный в узком смысле случайный процесс, необходимо найти $\min_{\bar{u}\in I} F_T(\bar{u})$ и исследовать асимптотическое поведение $\bar{u}_T = \underset{\bar{u}\in I}{\operatorname{argmin}} F_T(\bar{u})$ и $F_T(\bar{u}_T)$ при $T \to \infty$.

Особо хотелось бы остановиться на случае, когда неизвестный параметр u является элементом некоторого компактного множества K из некоторого функционального пространства. Сформулируем более подробно задачу и полученные результаты. Сначала рассмотрим случай, когда $K \subset C_{[0,1]}$ – пространства непрерывных на $[0,1]$ функций с равномерной метрикой.

Пусть ξ – случайная величина, заданная на вероятностном пространстве $(\Omega, \Im, P)$, $f(t,y,z):[0,1]\times R\times R\to R^+$ – функция, непрерывная на $[0,1]\times R$ при фиксированном z и измеримая по z при фиксированных t и y. Проблема состоит в нахождении

$$\min_{u\in K} E\left\{\int_0^1 f(t,u(t),\xi)dt\right\} = \min_{u\in K}\widetilde{F}(u).$$

Эта задача аппроксимируется следующей:

Найти

$$\min_{u\in K}\frac{1}{n+1}\sum_{i=1}^{n} f\left(\frac{i}{n}, u\left(\frac{i}{n}\right), \xi_i^n\right) = \min_{u\in K}\widetilde{F}_n(u),$$

где $\left\{\xi_i^n, 0\le i\le n, n\ge 1\right\}$ – последовательность серий независимых наблюдений случайной величины ξ.

Пусть $u_n = \operatorname{argmin}_{u\in K}\widetilde{F}_n(u)$, $u_0 = \operatorname{argmin}_{u\in K}\widetilde{F}(u)$. Задача состоит в нахождении условий, при которых $u_n\to u_0$, $\widetilde{F}_n(u_n)\to\widetilde{F}(u_0)$ при $n\to\infty$ в некотором вероятностном смысле. Имеет место следующее утверждение.

Теорема 6 [8, 9]. *Пусть выполнены условия:*

1) $E\left\{\max_{u\in K} f^4(t,u(t),\xi)\right\}\le L$, $t\in[0,1]$, *с некоторой постоянной* $L>0$;

2) существует единственная точка $u_0\in K$ *такая, что* $F(u)>F(u_0)$, $u\ne u_0$.

Тогда для любого $n\ge 1$ *функцию* u_n *можно выбрать такой, чтобы для каждого* $t\in[0,1]$ *функция* $u_n(t,\omega)$, $\omega\in\Omega$, *была* $\Im_n$*-измеримой, где* $\Im_n$ – σ*-алгебра, порожденная случайными величинами* $\left\{\xi_i^n, 0\le i\le n\right\}$. *Кроме того,*

$$P\{\lim_{n\to\infty}\| u_n - u_0\| = 0\} = 1,$$

$$P\{\lim_{n\to\infty}\widetilde{F}_n(\vec{u}_n) = \widetilde{F}(\vec{u}_0)\} = 1.$$

Как и ранее, доказательство состоит в проверке условий теоремы 2.

Замечание 4. Выбор $C_{[0,1]}$ в качестве функционального пространства, на котором задано компактное множество K, безусловно, не является единственно возможным. Например, не менее интересным представляется случай, когда множество K принадлежит некоторому гильбертовому пространству с соответствующей метрикой, пространству непрерывных функций с многомерным аргументом и т.д.

Замечание 5. Существенным моментом при доказательстве измеримости оценок неизвестных параметров является следующее утверждение об измеримом выборе, которое используется как в явном виде, так и частично, применяя идеи его доказательства. Приведем его для функции двух переменных.

Теорема 7 [21]. *Пусть T – произвольное замкнутое или открытое подмножество в R^l, $l \geq 1$, $(X, \mathcal{X})$ – некоторое измеримое пространство. Предположим, что $f : T \times X \to [-\infty, \infty]$ – функция, удовлетворяющая условиям:*

1) $f(t,x)$, $t \in T$, *непрерывна для всех* $x \in X$;

2) $f(t,x)$, $x \in X$, $\mathcal{X}$ *– измерима для любого* $t \in T$;

3) для всех $x \in X$ *существует точка* $t^* \in T$, *для которой* $f(t^*,x) = \inf_{t \in T} f(t,x)$.

Тогда существует измеримое отображение $\varphi : X \to T$ *такое, что* $f(\varphi(x),x) = \inf_{t \in T} f(t,x)$, $x \in X$.

Асимптотическое распределение оценок

Сначала остановимся на случае, когда в модели (2) ξ_i – независимые, одинаково распределенные случайные величины. В дальнейшем, там , где это не вызывает недоразумений, индекс *1* мы будем опускать. Справедливо утверждение.

Теорема 8[9]. *Пусть выполнены условия теоремы 1 и следующие:*

1) u_0 *является внутренней точкой множества* I;

2) существует такая замкнутая окрестность S точки u_0, что для любых $z \in X$ функция $f(u, y)$, $u \in S$, является дважды непрерывно дифференцируемой на S;

3) $E\left\{\max\limits_{u \in S} \left\| \nabla f(u, \xi) \right\|\right\} < \infty$, $\nabla f(u, z) = \left(\dfrac{\partial f}{\partial u_j}(u, z) \right)_{j=1}^{l}$, $u = \left(u_j\right)_{j=1}^{l} \in S$,

$z \in Y$ (благодаря свойствам измеримости функций для всех $u \in S$ отображение $\{\nabla f(u, \xi), z \in Y\}$ является $L(Y)$-измеримым);

4) $E\{\max\limits_{u \in S} \left\| \Phi(u, \xi) \right\|\} < \infty$, *где* $\Phi(u, z) = \left(\dfrac{\partial^2 f}{\partial u_j \partial u_k}(u, z) \right)_{j,k=1}^{l}$;

5) $\det A_0 \neq 0$, *где* $A_0 = E\Phi(u_0, \xi)$;

6) $E\{\left\| \nabla f(u, \xi) \right\|^2\} < \infty$ *и* $\det C \neq 0$, *где* $C = E\{\nabla f(u_0, \xi)\left(\nabla f(u_0, \xi)\right)'\} < \infty$.

Тогда $\sqrt{n}\left(u_n - u_0\right)$ сходится по распределению при $n \to \infty$ к $N\left(0, (A_0)^{-1} C (A_0)^{-1}\right)$, где $N(a, B)$ – гауссовское распределение со средним a и корреляционной матрицей B.

Если, кроме того, выполнены условия

7) $Ef^2(u_0, \xi) < \infty$;

8) $\sigma^2 = E\left\{\left(f(u_0, \xi) - Ef(u_0, \xi)\right)^2\right\} > 0$,

то $\sqrt{n}\left(F(u_n) - F(u_0)\right)$ сходится по распределению при $n \to \infty$ к $N\left(0, \sigma^2\right)$

Если в модели (2) $\{\xi_i, i \geq 1\}$ – стационарная в узком смысле случайная последовательность, то можно получить следующие общие утверждения.

Теорема 9 [9] . *Пусть $\{\xi_i, i \geq 1\}$ – стационарная в узком смысле случайная последовательность, для которой выполнены условия 1),2),4),5) теоремы 8 и следующие:*

1) функция $f(u, \xi_i)$ удовлетворяет условию сильного перемешивания

$$\alpha(j) = \sup_{\substack{A\in\sigma_{-\infty}^{i} \\ B\in\sigma_{i+j}^{\infty}}} \left|P(AB) - P(A)P(B)\right| \le \frac{c}{1+j^{1+\varepsilon}}$$

с коэффициентом перемешивания $\alpha(k) = O\left(k^{-1-\varepsilon}\right)$, $k \to \infty$, $\varepsilon > 0$;

2) для некоторого $\delta > \frac{4}{\varepsilon}$ *выполняется* $E\left\{\left\|\nabla f(u,\xi_i)\right\|^{2+\delta}\right\} < \infty$;

3) для спектральной плотности $g(\lambda)$ *процесса* $\left\{\nabla f(u_0,\xi_i), i \in Z\right\}$, *которая существует в силу условий 1) и 2), выполнено условие* $\det g(0) \neq 0$.

Тогда $\sqrt{n}\left(u_n - u_0\right)$ *сходится по распределению при* $n \to \infty$ *к гауссовскому вектору* $N\left(0, 2\pi(A_0)^{-1} g(0)(A_0)^{-1}\right)$, *где* A_0 *определена в теореме 8. Если, кроме того, выполнены условия*

4) для некоторого $\delta_1 > \frac{4}{\varepsilon}$ *выполняется* $E\left\{f(u_0,\xi_i)^{2+\delta_1}\right\} < \infty$;

5) для спектральной плотности $g_1(\lambda)$ *случайного процесса* $\left\{f(u_0,\xi_i), i \in Z\right\}$ *справедливо* $g_1(0) \neq 0$.

Тогда $\sqrt{n}\left(F(u_n) - F(u_0)\right)$ *сходится по распределению при* $n \to \infty$ *к* $N\left(0, 2\pi g_1(0)\right)$.

Аналогичные результаты имеют место и для случая непрерывного временного пространственного параметра, а также в случае, когда функция f также зависит от этих параметров.

Дальнейшее обобщение состоит в рассмотрении случая, когда неизвестный параметр удовлетворяет ограничениям, заданных в виде системы неравенств. Остановимся более детально на полученных результатах.

Рассмотрим задачу минимизации функционала

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} f(x,\xi_i), \quad x \in J,$$

где

$$J = \{x : g(x) = (g_1(x), ..., g_m(x)) \le 0\}.$$

и $\{\xi_t, i = 1,2,...\}$ – строго стационарный в узком смысле эргодический процесс.

Предположим, что выполнены следующие условия.

A1. Функция $f(x,y)$ является дважды непрерывно дифференцируемой по первому аргументу и $E \max\limits_{\|u\| \le c} f(x,\xi) < \infty$ для любого $c > 0$.

A2. Для точки u_0 такой, что $F(u_0) < F(u)$, $u \in I$, $u \ne u_0$ выполнено

$$P\{\lim_{n \to \infty} \|u_n - u_0\| = 0\} = 1.$$

A3. Существует $\gamma_0 > 0$ такое, что $E\left|\dfrac{\partial^2 f(u,\xi_i)}{\partial u_j \partial u_k}\right| \le c$ при $\|u - u_0\| < \gamma$, $0 < \gamma < \gamma_0$.

A4. Пусть N_1 и N_2 – множества индексов $\{i\}$ таких, что $g_i(u_0) = 0$, $i \in N_1$ и что $g_i(u_0) < 0$, $i \in N_2$.

A5. Функции $g_i(u)$ дважды непрерывно дифференцируемы и для $\|u - u_0\| \le \gamma$, $0 < \gamma < \gamma_0$, $\left|\dfrac{\partial g_i(u)}{\partial u_k}\right| \le c$ и $\left|\dfrac{\partial^2 g_i(u)}{\partial u_j \partial u_k}\right| \le c$.

A6. Существует точка u^* такая, что $g(u^*) < 0$.

A7. Функции $g_i(u)$ выпуклы.

A8. Случайный вектор $f'(u_0,\xi_i)$ удовлетворяет условию сильного перемешивания, определенного в условии 1) теоремы 9, с коэффициентом перемешивания $\alpha(n) \le \dfrac{c}{n^{1+\varepsilon}}$, $\varepsilon > 0$.

A9. $E\left\{\left\|f'(u_0,\xi_i)\right\|^{2+\delta}\right\} < \infty$, $\delta\varepsilon > 2$.

A10. Спектральная плотность $h(\lambda)$ вектора $f'(u_0,\xi_i)$ является невырожденной матрицей в точке $\lambda = 0$.

Теорема 10 [13]. *Пусть выполнены условия A1 – A10. Тогда вектор* $\zeta_n = \sqrt{n}\left(u_n - u_0\right)$ *слабо сходится при* $n \to \infty$ *к случайному вектору* ζ*, который является решением задачи квадратического интегрирования вида*

$$\frac{1}{2} n^T \Phi(u_0) u + \zeta u \to \min,$$

$$\nabla g^T(u_0) u \le 0,$$

где T *– знак транспонирования, а* ζ *– нормально распределенная случайная величина с параметрами* $\left(0, 2\pi h(0)\right)$.

Замечание 6. При рассмотрении случая непрерывного временного параметра в стационарной или нестационарной моделях также имеет место утверждение об асимптотическом распределении оценок, полученных методом эмпирических средних.

Стохастические системы с долгой памятью

Выше мы рассматривали случаи зависимых наблюдений, когда соответствующие случайные последовательности или процессы удовлетворяют условию сильного перемешивания. Оно накладывает довольно сильные ограничения на скорость убывания зависимости между наблюдениями при удалении расстояния между ними. В последние годы проведено достаточно много исследований, направленных на ослабление условий сильного перемешивания. Было введено понятие сильной зависимости, а стохастические системы, удовлетворяющие этим условиям, получили название систем с долгой памятью. Приведем один из характерных результатов для таких систем, касающийся состоятельности оценок неизвестных параметров.

Пусть

$$S_T(\beta) = \frac{1}{T} \int_0^T f\left(t, \beta, \eta(t)\right) dt,$$

где $f\left(t,\beta,\eta(t)\right):[0,\infty)\times J\times R^s\to[0,\infty), s\geq 1,$ – известная функция; J – замкнутое подмножество в R^p, $p\geq 1$; $\|\cdot\|_p$ – норма в R^p; $\eta(t)$, $t\in R$, – случайный шум.

Пусть выполнены условия:

1. $f\left(t,\beta,y\right)$ непрерывная по совокупности аргументов функция.

2. $\eta(t)$, $t\in R$ – стационарный процесс второго порядка с нулевым математическим ожиданием, который представлен в виде $\eta(y)=G\left(\xi(t)\right)$, где $\xi(t)$, $t\in R$, – действительный, измеримый, непрерывный в среднем квадратичном, стационарный гауссовский случайный процесс с нулевым математическим ожиданием, единичной дисперсией и корреляционной функцией

$$B(t)=\operatorname{cov}\left(\xi(0),\xi(t)\right)=L(t)\left|t\right|^{-\alpha},\ 0<\alpha<1,$$

где, $L(t)=L'(\left|t\right|)$ $t>0$, – неотрицательная функция, медленно меняющаяся на бесконечности, т.е. $\forall s>0\ \lim\limits_{t\to\infty}\dfrac{L(ts)}{L(t)}=1$, и ограничена на каждом конечном интервале, кроме того

$$\lim_{t\to\infty}\sup_{t\in[0,T)}\frac{L(ts)}{L(t)}<\infty,$$

$G(u)$, $u\in R$ – действительная, измеримая, неслучайная функция, для которой $EG^2(\xi(0))<\infty$.

Пусть существует $t_0>0$ такое, что при $t>t_0$ функция $B(t)$, $t\in R$ убывает. Справедливо утверждение.

Теорема 11 [22, 23]. *Пусть выполнены условия 1, 2, а также*

1) для любого $\beta\in J$ существует функция $S(\beta)$ такая, что

$$S(\beta)=\lim_{T\to\infty}ES_T(\beta)$$

и точка $\beta^*\in J$, в которой эта функция имеет единственный минимум;

2) если множество J неограничено, то

$$f(t,\beta,y)\to\infty,\ \textit{при}\ \|\beta\|_p\to\infty$$

при фиксированных t, y;

3) *существует функция* $c(\gamma)>0$ *такая, что* $c(\gamma)\to 0$, $\gamma\to 0$ *и для любого* $\varepsilon>0$ *существует* $\gamma_0>0$ *такое, что при* $0<\gamma<\gamma_0$ *для любого* $\beta'\in J$:

$$\overline{\lim_{T\to\infty}}\frac{1}{T}\int_0^T E\sup_{\substack{\|\beta-\beta'\|_p<\gamma\\ \|\beta-\beta^*\|_p>\varepsilon}}\left|f(t,\beta,\eta(t))-f(t,\beta',\eta(t))\right|dt<c(\gamma)$$

4) $Ef(t,\beta,G(\xi(0)))<\infty$;

5) $\sup\limits_{\substack{t\ge 0\\ \beta\in J}}\sum\limits_{i\ge 1}\dfrac{C_i^2(t,\beta)}{i!}<\infty$, *где* $C_i(t,\beta)$ *– коэффициенты в разложении функции* f *по полиномам Чебышева – Эрмита;*

6) $\sup\limits_{T\ge 0}\int_0^1\int_0^1\dfrac{L(T|t-s|)}{L(T)|t-s|^{\alpha}}dtds<\infty$.

Тогда оценка параметра β^*, *которая определяется как точка минимума функционала* $S_T(\beta)$, *будет сильно состоятельной.*

В работах [24, 25] получен ряд фундаментальных результатов, касающихся асимптотического поведения оценок нелинейной регрессии для стохастических систем с сильно зависимыми наблюдениями. Приведем один из этих результатов, касающихся сильной состоятельности оценки наименьших квадратов [24, 25] .

Пусть $(\Omega,\Im,P)$ – полное вероятностное пространство, $\xi(t)=\xi(\omega,t):\Omega\times\Re\to\Re$ – случайный процесс, удовлетворяющий условиям.

А. $\xi(t), t\in\Re$, – случайный измеримый непрерывный в среднем квадратическом стационарный гауссовский процесс с $E\xi(t)=0$ и функцией ковариации

$$B(t)=\operatorname{cov}\big(\xi(0),\xi(t)\big)=\frac{1}{\left(1+t^2\right)^{\alpha/2}},\quad 0<\alpha<1,\ t\in\Re.$$

При условии А, процесс $\xi(t), t\in\Re$, допускает представление

$$\xi(t)=\int\limits_{R^1} e^{i\lambda t}\sqrt{f_\alpha(\lambda)}W(d\lambda),$$

где $W(\cdot)$ – гауссовский белый шум на измеримом пространстве $\big(\Re,B(\Re)\big)$.

В. Нелинейная борелевская функция $G:\Re\to\Re$ такова, что

$$\int\limits_{-\infty}^{\infty} G^2(u)\varphi(u)du<\infty,$$

где

$$\varphi(u)=\frac{1}{\sqrt{2\pi}}e^{-u^2/2},\ u\in\Re.$$

При условии В функцию $G(u), u\in\Re$, можно разложить в ряд

$$G(u)=\sum_{k=0}^{\infty}\frac{C_k}{k!}H_k(u),\quad C_k=\int\limits_{-\infty}^{\infty}G(u)H_k(u)\varphi(u)du,\quad k=0,1,...,$$

по ортогональным полиномам Чебышева-Эрмита

$$H_k(u)=(-1)^k e^{u^2/2}\frac{d^k e^{-u^2/k}}{du^k},\quad k=0,1,....$$

в гильбертовом пространстве $L_2\big(\Re,\varphi(u)du\big)$.

Предположим также, что функция *G* удовлетворяет условию

В′. Существует целое $m\ge 1$, такое что $C_1=...=C_{m-1}=0$ и $C_m\neq 0$.

Целое $m\ge 1$ называется порядком Эрмита функции *G* Пусть $\Theta\subset\Re$ - ограниченный открытый интервал. При условиях А и В рассмотрим регрессионную модель

$$y(t)=y_\theta(t)=g(t,\theta)+G\big(\xi(t)\big),\ t\in\Re,$$

где $g(t,\theta):\Re\times\Theta^c\to\Re$ - измеримая функция, зависящая от неизвестного $\theta\in\Theta$, а случайный шум $\varsigma(t)=g\big(\xi(t)\big),\ t\in\Re$, такой что $E\varsigma(0)=0$ (or

$C_0 = 0$) и $E\varsigma^2(0) = 1$. Необходимо оценить параметр θ из наблюдений случайного процесса $y_\theta(t),\ t \in [0,T]$, при $T \to \infty$.

Любая случайная переменная $\hat{\theta}_T \in \Theta^c$ удовлетворяющая соотношению

$$Q_T\left(\hat{\theta}_T\right) = \inf_{\tau \in \Theta^c} Q_T(\tau),\ \ Q_T(\tau) = \int_0^T \left[y_\theta(t) - g(t,\tau) \right]^2 dt$$

называется оценкой наименьших квадратов неизвестного параметра $\theta \in \Theta$, полученной из наблюдений $y_\theta(t), t \in [0,T]$, где Θ^c - замыкание Θ. Положим

$$b_T(\tau_1,\tau_2) = \int_0^T \left(g(t,\tau_1) - g(t,\tau_2) \right)^2 dt,$$

$$\eta_T(\tau) = \int_0^T G\left(\xi(t)\right)\left(g(t,\tau) - g(t,\theta) \right) dt,$$

$$\gamma(T) = \int_0^T G^2\left(\xi(t)\right) dt.$$

По определению оценки наименьших квадратов, почти наверное

$$Q_T(\theta) = \gamma(T) \ge Q_T\left(\hat{\theta}_T\right) = \gamma(T) - 2\eta_T\left(\hat{\theta}_T\right) + b_T\left(\hat{\theta}_T,\theta\right),$$

или

$$b_T\left(\hat{\theta}_T,\theta\right) \le 2\eta_T\left(\hat{\theta}_T\right).$$

Если $g(t,\theta)$ дифференцируема по θ, можно записать

$$\frac{\partial^r g(t,\theta)}{\partial \theta^r} = g_r(t,\theta),\ \ d_{rT}^2(\tau) = \int_0^T g_r^2(t,\theta)\, dt,\ \ r = 1,2,3.$$

Предположим функцию $g(t,\tau)$ можно продолжить на некоторый интервал $\Theta^* \supset \Theta$ таким образом, что $g(t,\cdot) \in C^3\left(\Theta^*\right)$, функция $g_r(t,\tau),\ r = 1,2$, ограничена на $[0,T] \times \Theta^c$ для любого $T > 0$, и $g_3(t,\tau)$ - локально квадратически интегрируема по *t*.

Заметим, что функция

$$A_\theta^2(T)=\frac{1}{T}d_{1T}^2(\theta)=\int_0^1 g_1^2(Ts,\theta)ds$$

используется в предельных теоремах и входит в нормализующие коэффициенты.

Через $k_i, i=0,1,2,...$ обозначены положительные константы, точные значения которых несущественны в рассматриваемой задаче.

Для широкого класса функций регрессии выполнено следующее условие.

C. Предположим что

$$\sup_{0\le s\le 1}\frac{\left|g_1(Ts,\tau)\right|}{A_\theta(T)}\le k_1<\infty$$

равномерно по $T>0$ и $\tau\in\Theta^c$.

Следующее условие называется условием контрастности или условием различимости параметров.

D. ля любого $\rho>0$ существует величина $r=r(\rho)>0$ такая что

$$\inf_{|u|>\rho\left(T^{-1}d_{1T}^2(\theta)\right)^{-1/4}}\frac{b_T(u+\theta,\theta)}{T\left[T^{-1}d_{1T}^2(\theta)\right]^{1/2}}>r.$$

Для метода наименьших квадратов получим

Теорема 12[24]. *Пусть условия A, B,* B′, *C и D выполнены для* $\alpha\in\left(0,\frac{1}{m}\right)$, *где m - порядок Эрмита функции G. Тогда*

$$T^{-1/4}d_{1T}^{1/2}(\theta)\left(\hat\theta_T-\theta\right)\xrightarrow{P}0 \text{ при } T\to\infty.$$

Сходимость по функционалу в задачах стохастического программирования

В задачах стохастического программирования часто возникают случаи, когда не всегда представляется возможным доказать состоятельность оценки параметра, на которой достигается минимум некоторого функционала. Достаточно, чтобы последовательность функционалов в точке минимума сходи-

лась в том или ином вероятностном смысле. В этом случае имеет место сходимость по функционалу. Ниже представлены результаты, полученные в работах [7, 26].

Рассмотрим задачу

$$F(x)=\int_{\Omega} f(x,\xi)P(d\xi)\to\min_{x\in K} \tag{3}$$

где $x\in K\subset X$, K – компактное множество в некотором топологическом пространстве X, $\xi\in\Omega$, $(\Omega,\Im,P)$ – вероятностное пространство, функция $f: X\times\Omega\to\Re$ интегрируема для любого фиксированного x, $F(x)$ – полунепрерывна снизу.

Метод решения (3) состоит в замене исходной задачи последовательностью вида:

$$F_n(x)=\frac{1}{n}\sum_{i=1}^{n} f(x,\xi_i)\to\min_{x\in K}$$

где ξ_i – независимые одинаково распределенные случайные величины, Естественным образом возникает вопрос о сходимости оптимальных величин $F_n^*(\omega)=\min_{x\in K}F_n(x)$ и $F^*(\omega)=\min_{x\in K}F(x)$ Запишем $F_n(x)$ как

$$F_n(x)=F(x)+\frac{1}{n}\sum_{i=1}^{n}\left(f(x,\xi_i)-F(x)\right)=F(x)+y_n(x,\omega).$$

Рассмотрим функционал

$$\Phi(y)=\min_{x\in K}\varphi\left(x,y(x)\right)=\min_{x\in K}\left(F(x)+y(x)\right),$$

где $F(x)$ полунепрерывная снизу на компакте K функция, $y(x)\in C(K;\Re)$, $C(K;\Re)$– банахово пространство непрерывных на K функций с нормой $\|y\|=\max_{x\in K}|y(x)|$.

Справедливы следующие утверждения [7,26].

Теорема 13. *Пусть* $(\Omega,\Im,P)$ *- вероятностное пространство, D – относительно открытое выпуклое множество в* $\Re^p$. *Предположим, функция*

$f : D \times \Omega \to \Re$ выпукла по **x** *в D для почти всех ξ, и интегрируема по ξ для всех* $\mathbf{x} \in D$. *Тогда для любого компакта* $K \subset D$

$$\left\| y_n(\cdot, \omega) \right\| = \max_{\mathbf{x} \in K} \left| y_n(\mathbf{x}, \omega) \right| \to 0 \;\; P_\omega\text{-п. н.}$$

Теорема 14. *Пусть* $(\Omega, \Im, P)$- *вероятностное пространство, K – компактное множество в полном сепарабельном метрическом пространстве. Предположим что функция* $f : K \times \Omega \to \Re$ *интегрируема в* $\xi \in \Omega$ *для всех* $x \in K$, *и является функцией Липшица в* $x \in K$ *равномерно по* ξ:

$$\left| f(x, \xi) - f(y, \xi) \right| \le L \left\| y - x \right\|, \;\; x, y \in K.$$

Тогда P_ω*-почти наверное*

$$\left\| y_n(\cdot, \omega) \right\| = \max_{x \in K} \left| \frac{1}{n} \sum_{i=1}^{n} f(x, \xi_i) - F(x) \right| \to 0, \;\; n \to \infty \;.$$

Отметим, что условия выполнения равномерного закона больших чисел для эмпирического риска изучались многими авторами, но здесь мы мы не имеем возможности останавливаться на полученных ими результатах. Отметим лишь, что для непараметрических моделей регрессии рассмотренные проблемы достаточно полно обсуждались в работах [27] .

Большие уклонения для оценок

Перейдем к вопросу о больших уклонениях метода эмпирических средних, существенному для анализа влияния редких экстремальных событий Одной из важных проблем стохастической оптимизации является исследование скорости сходимости приближенного решения к истинному решению задачи. Среди работ в этом направлении хотелось бы отметить [28]. Существенное влияние на эти исследования оказала монография [29]. Основываясь на развитом в ней аппарате, были получены экспотенциальные оценки больших уклонений метода эмпирических средних при достаточно общих условиях на множество неизвестных параметров и независимых наблюдениях ξ_i.

В дальнейшем нам понадобятся следующие результаты из функционального анализа.

Определение 1 [28]. Пусть $(V,\|\circ\|)$ - линейное нормированное пространство, $B(x,\rho)$ – замкнутый шар в V радиуса ρ с центром x, $f:V\to[-\infty,+\infty]$ – некоторая функция, $f(x_f)=\min\{f(x),x\in V\}$. *Улучшающей* функцией ψ для f в точке x_f называется монотонно неубывающая функция $\psi:[0,+\infty)\to[0,+\infty]$ с $\psi(0)=0$, такая, что существует $\rho>0$, для которого при всех $x\in B(x_f,\rho)$ имеем

$$f(x)\ge f(x_f)+\psi\left(\left\|x-x_f\right\|\right).$$

Пусть $V_0\subset V$. Обозначим δ_{V_0} индикаторную функцию V_0:

$$\delta_{V_0}(x)=0,\ x\in V_0;$$

$$\delta_{V_0}(x)=+\infty,\ x\notin V_0.$$

Определение 2 [29]. Пусть $\sum-$ сепарабельное банахово пространство, $\{\xi_i,i\in Z\}$ – стационарная в узком смысле случайная последовательность, определенная на вероятностном пространстве (Ω,F,P) и принимающая значения из Σ. Обозначим B_{mk} σ-алгебру подмножеств Ω, порожденную случайными элементами $\{\xi_i,m\le i\le k\}$. Для данного $l\in N$ действительные случайные величины $\eta_1,\ldots,\eta_p$, $p\ge 2$ называются *l-измеримо отделенными*, если

$$-\infty\le m_1\le k_1<m_2\le k_2<\ldots<m_p\le k_p\le+\infty;$$

$$m_j-k_{j-1}\ge l,\quad j=2,\ldots,p$$

и при каждом $j\in\{1,\ldots,p\}$ случайная величина η_j является $B_{m_jk_j}$-измеримой.

Определение 3 [29]. Случайная последовательность $\{\xi_i\}$ из определения 2 называется последовательностью с *гиперперемешиванием*, если существуют число $l_0\in N\bigcup\{0\}$ и невозрастающие функции $\alpha,\beta:\{l>l_0\}\to[1,+\infty)$ и $\gamma:\{l>l_0\}\to[0,1]$, удовлетворяющие соотношениям

$$\lim_{l\to\infty}\alpha(l)=1,\quad \limsup_{l\to\infty} l(\beta(l)-1)<\infty,\quad \lim_{l\to\infty}\gamma(l)=0,$$

и для которых

$$\left\|\eta_1\dots\eta_p\right\|_{L^1(P)}\le\prod_{j=1}^{p}\left\|\eta_j\right\|_{L^{\alpha(l)}(P)}$$

при любых $p\ge 2, l>l_0, \eta_1,\dots,\eta_p$ l-измеримо отделенных, где

$$\left\|\eta\right\|_{L^r(P)}=\left(\int_\Omega\left|\eta(\omega)\right|^r dP\right)^{1/r},$$

и

$$\left|\int_\Omega\left(\xi(\omega)-\int_\Omega\xi(\omega)dP\right)\eta(\omega)dP\right|\le\gamma(l)\left\|\xi\right\|_{L^{\beta(l)}(P)}\left\|\eta\right\|_{L^{\beta(l)}(P)}$$

для всех $l>l_0, \xi,\eta\in L^1(P)$ l-измеримо отделенных.

Справедливо следующее утверждение.

Теорема 15 [30, 31]. *При выполнении условия гиперперемешивания имеет место равенство*

$$\limsup_{n\to\infty}\frac{1}{n}\ln P\left\{\left|\min\left\{Ef(u,\xi_i),u\in K\right\}-\min\left\{F_n(u),u\in K\right\}\right|\ge\varepsilon\right\}\le-\inf\left\{I(z),z\in A_\varepsilon\right\},$$

где

$$A_\varepsilon=\left\{z\in K:\left\|z\right\|>\varepsilon\right\},$$

$$I(z)=\sup_{Q\in M(X)}\int_X z(x)Q(dx)-\Lambda(Q),\ Q\in M(X),$$

$$\Lambda(Q)=\lim_{n\to\infty}\frac{1}{n}\left(\int_\Omega\exp\left\{\sum_{i=1}^{n}\int_X\xi_i(\omega)(u)Q(du)\right\}dP\right).$$

Предположим, что существует улучшающая функция ψ для Ef в точке x^* с некоторой постоянной ρ. Пусть x_n – точка минимума (2) по множеству $B(x^*,\rho)$. Если ε достаточно мало, так что выполнено условие

$$\psi\left(\left|x-x^*\right|\right)\le 2\varepsilon\ \Rightarrow\left|x-x^*\right|\le\rho,$$

где $\Rightarrow$ – знак логического следования, то

$$\limsup_{n\to\infty}\frac{1}{n}\ln P\left\{\psi\left(\left|x_n - x^*\right|\right)\geq 2\varepsilon\right\}\leq -\inf\left\{I(z), z\in A_\varepsilon\right\}.$$

Замечание 7. Если, кроме условий теоремы 17, ψ выпукла и строго возрастает на $[0,\rho]$, то

$$\limsup_{n\to\infty}\frac{1}{n}\ln P\left\{\left|x_n - x^*\right|\geq \psi^{-1}(2\varepsilon)\right\}\leq -\inf\left\{I(z), z\in A_\varepsilon\right\}.$$

Список литературы

[1] Юби Э. Статистическое исследование и метод решения задач стохастического программирования // Изв. АН ЭССР, Физ.-мат, 1977. – Т. 26. – №4. – С. 369 – 375.

[2] Dupacova J. Experiance in stochastic programming models // Servey Math. Programming: Symp. – Budapest: Akademia Kiado, 1979. – Pp. 99 – 105.

[3] Wets R.J.-B. Statistical approach to the solution of stochastic programs with (convex) simple recourse. Lexington, 1979. – 30 p., (Working paper/Univ. of Kentucky).

[4] Salinetti G., Wets R.J.B. On the convergence in distribution of measurable multifunctions and stochastic intima // Math. Oper. Res. – №11. – №3. – Pp. 385 – 419

[5] Dupacova J., Wets R.J.-B. Asymptotic behavior of statistical estimators and optimal solutions for stochastic optimization problems //Ann. Statist., 1988. – №16. – Pp. 1517 – 1549.

[6] Shapiro A. Asymptotic properties of statistical estimators in stochastic programming // Ann. Statist., 1989. – Vol. 17. – №2. – Pp. 841 – 858.

[7] Норкин В.И. Об условиях и скорости сходимости метода эмпирических средних в математической статистике и стохастического программировании. // Киб. и СА, 1992. – №2. С. 107 – 120.

[8] Knopov P.S., Kasitskaya E.J. Properties of Empirical Estimates in Stochastic Optimization and Identification Problems // Ann. Res. – №56 – Pp. 225 – 239.

[9] Knopov P.S., Kasitskaya E.J. Empirical estimates in Stochastic optimization and identification. – Dordrecht-Boston-London, Kluwer Academic Publishers, 2005. – 250 p.

[10] Кнопов П.С. Оптимальные оценки параметров стохастических систем. – Киев: Наукова думка, 1981. – 152 с.

[11] Кнопов П.С. Об одном подходе к решению задач стохастической оптимизации // Кибернетика, 1988. – №4. – С. 126 – 127.

[12] Kankova V. An approximate solution of a stochastic optimization problem // Trans. of the 8-th Prague conf. – Prague: Academia, 1978. – Pp. 349 – 353.

[13] Knopov P., Korkhin A. Regression Analysis Under A Priori Parameter Restrictions. – Springer, 2012. – 234 p.

[14] Ермольев Ю.М., Кнопов П.С. Метод эмпирических средних в задачах стохастического программирования // КиСА, 2006. – № 6. – С. 3 – 18.

[15] Хьюбер Дж. Робастность в статистике. – М.: Мир, 1984. – 304 с.

[16] Pfanzagl J. On the Measurability and Consistency of Minimum Contrast Estimates // Metrica-14. – Pp. 249 – 272.

[17] Вапник В.Н. Восстановление зависимостей по экспериментальным данным. – М.: Наука, 1979. – 488 с.

[18] Jenrich R.I. Asymptotic properties of non-linear least squares estimators // Ann. Math. Statist., 1969. – Pp. 633 – 643.

[19] Дороговцев А.Я. Теория оценивания параметров случайных процессов. – Киев: Вища школа, 1982. – 190 с.

[20] Le Cam L. On some asymptotic properties of maximum likehood estimates and related Bayes estimates // Univ. California Publ.Statist., 1953. – Vol. 1. – Pp. 277 – 330.

[21] Schmetterer Introduction to Mathematical Statistics. – Berlin: Springer–Verlag, 1966. – 520 p.

[22] Leonenko N.N., Moldavskaya E.M. Non-Gaussian scenarios in long-memory regression models with non-linear constraints. // Reports of the National Academy of Sciance of Ukraine, 2002. –Vol. 2. – Pp. 44 – 46.

[23] Moldavskaya E.M. Theorem useful in proving of estimates consistency in long-memory regression models // Bulleten of the University of Kiev. Series: Physics, Mathematics, 2002. – Vol. 1. – Pp. 58 – 65.

[24] Ivanov A.V., Leonenko N.N. Asymptotic inferences for a nonlinear Regression with a long-range dependence // Prob Theory Math. Statist., 2001. – № 63. – Pp. 65 – 85.

[25] Ivanov A.V., Leonenko N.N. Asymptotic behavior of M-estimators in continuous-time non-linear regression with long- range dependent errors // Random Oper.and Stoch. Equ., 2002. – Vol. 10, (3). – Pp. 201 – 222.

[26] Норкин В.И. Устойчивость стохастических оптимизационных моделей и статистические методы стохастического программирования. – Киев: Институт кибернетики им. В.М.Глушкова НАНУ, Препринт 89–53, 1989. – 24 с.

[27] Gyorfi L., Kohler M., Krzyak A., Walk H. A distribution-Free Theory of Nonparametric Regression. – Springer, 2002. – 646 p.

[28] Kaniovski Yu.M., King A., Wets R.J.-B. Probabilistic bounds via large deviations for the solutions of stochastic programming problems // Annals of Oper.res., 1995. – № 56. – Pp. 189 – 208.

[29] Deutschel J.D., Strook D.W. Large Deviations. – Boston: Academic Press, 1989. – 310 p.

[30] Кнопов П.С., Касицкая Е.И. О больших уклонениях эмпирических оценок в задачах стохастического программирования // Киб. и СА., 2004. – №4. – С. 52 – 61.

[31] Кнопов П.С., Касицкая Е.И. О сходимости эмпирических оценок в задачах стохастического программирования для процессов с дискретным временем // Киб. и СА, 2005. – № 1. – С. 175 – 178.

УДК 519.85

РЕШЕНИЕ ДВУХЭТАПНЫХ ЗАДАЧ СТОХАСТИЧЕСКОГО ПРОГРАММИРОВАНИЯ БОЛЬШОЙ РАЗМЕРНОСТИ PNK-МЕТОДОМ

Кузьменко В. Н.

Институт кибернетики имени В.М.Глушкова НАН Украины

E-mail: kvnu@mail.ru

Аннотация: Рассматривается применение PNK-метода, использующего переменную кусочно-линейную аппроксимацию целевой функций и ограничений и подбор точного штрафного множителя, для решения двухэтапных задач стохастического программирования большой размерности. Под размерностью понимается количество переменных первого этапа и всех переменных второго этапа с учетом количества сценариев. Описывается PNK-метод, варианты постановки двухэтапной задачи и примеры решенных задач.

Ключевые слова: оптимизация, кусочно-линейная аппроксимация, стохастическое программирование, двухэтапная задача, сценарии.

Введение

В настоящей работе описывается PNK-метод (комбинированного метода выпуклого программирования) [1, 2] и его применение для решения двухэтапных задач стохастического программирования большой размерности. Метод основан на использовании переменной кусочно-линейной аппроксимации целевой функции и ограничений исходной задачи по информации в пройденных точках в виде набора отсекающих плоскостей, а также на подборе точного штрафного множителя для используемой штрафной функции. Опыт применение PNK-метода для решения прикладных задач показал вычислительную эффективность и надежность данного метода [1, 3]. Разработана модификация метода, генерирующая последовательность точек, сходящуюся к решению, по допустимой области исходной задачи [2], а также модификация, позволяющая для гладких задач учитывать информацию об окрестности текущей точки в виде квадратичной регуляризирующей добавки [4]. Для случая невыпуклых задач используется соответствующая модификация PNK-метода [5].

В научной литературе методы, основанные на аналогичных идеях, носят название «bundle method» или «bundle proximal methods». Рассматриваются в основном методы для решения задач без ограничений [6].

Описание PNK-метода

Изначально PNK-метод разработан для решения задач выпуклого программирования

$$\min_x \{f(x): f_j(x) \le 0,\ 1 \le j \le m,\ x \in M\},$$

где $x \in R^n$; f, f_j – выпуклые непрерывные функции, принимающие конечные значения на выпуклом многограннике M. Поиск решения основан на следующей итерационной процедуре.

Пусть сделано $k-1$ итераций. В результате получено множество точек X_k, состоящее из начальной точки $x_1 \in M$ и $k-1$ точек $x_i \in M, i = 2, \ldots, k$, – результатов итераций.

Очередная k – я итерация заключается в поиске решения квадратичной задачи, в которой минимизируется кусочно-линейная аппроксимация штрафной функции исходной задачи, построенная по точкам множества X_k, с учетом регуляризирущей квадратичной добавки, а именно:

$$\min_{x,\ \xi_1, \xi_2} \{\frac{1}{2h_k} \| x - x_r \|^2 + \xi_1 + N_k\, \xi_2\}, \tag{1}$$

$$f(x_i) + (f'(x_i),\ x - x_i) \le \xi_1,\ \forall x_i \in X_k, \tag{2}$$

$$\varphi(x_i) + (\varphi'(x_i),\ x - x_i) \le \xi_2,\ \forall x_i \in X_k, \tag{3}$$

$$0 \le \xi_2,\ x \in M, \tag{4}$$

где $x_r = \arg\min\{F_k(x): x \in X_k\}$ – точка минимума штрафной функции $F_k(x) = f(x) + N_k \max\{0; \varphi(x)\}$ на множестве X_k, r – индекс этой точки; $\varphi(x) = \max_j \{f_j(x):\ 1 \le j \le m\}$ – обобщенное ограничение, $f'(x)$, $\varphi'(x)$ – элементы субдифференциалов $\partial f(x)$, $\partial \varphi(x)$, N_k – значение штрафного множителя на итерации k, h_k – параметр, выполняющий роль шагового множителя,

переменные ξ_1 и ξ_2 аппроксимируют соответственно целевую функцию f и функцию ограничений φ с помощью систем неравенств (2) и (3).

По результатам решения задачи (1) – (4) выполняется регулировка параметров метода и изменение множества точек аппроксимации. А именно:

- вычисляется сделанный шаг $p_k = x_{k+1} - x_r$;
- штрафной множитель N_k увеличивается, если найденное в результате решения квадратичной задачи $\xi_2^* > 0$;
- отсеиваются наиболее неэффективные точки X_k^d аппроксимации задачи: $X_k' = X_k \setminus X_k^d$ (точка считается неэффективной для построения отсечения целевой функции или ограничений, если а) построенное по ней отсечение не есть активным в оптимальном решении квадратичной задачи; б) невязка соответствующего ограничения (2) (или (3)) больше, чем $\delta\|p_k\|^2$, где δ малый параметр);
- точка x_{k+1}, найденная в результате решения задачи (1) – (4), добавляется к множеству точек аппроксимации $X_{k+1} = X_k' \cup \{x_{k+1}\}$;
- регулируется «шаговый» множитель h_k, который отвечает за скорость сходимости метода и точность аппроксимации исходной задачи.

Метод решения квадратичной подзадачи.

Основная вычислительная нагрузка метода приходится, как правило, на процедуру решения квадратичной подзадачи (1) – (4). Поэтому важно уметь быстро решать последовательность близких, но изменяющихся квадратичных подзадач, а также учитывать общую специфику подзадачи (1) – (4). Такой спецификой есть диагональная квадратичная матрица в целевой функции (1) и наличие границ на переменные в множестве M.

Для решения цепочки квадратичных подзадач используется двойственный метод активного набора, учитывающий специфику подзадачи при вычислительной реализации.

Запишем квадратичную подзадачу (1) – (4) в более общем виде

$$\min_{x,\xi} \{0.5 \cdot x^T x + d\,x + p\,\xi \mid Ax + G\xi \geq b\}. \qquad (5)$$

В постановке (5) ограничения (4) учтены как общие, переменные ξ_1, ξ_2 собраны в вектор ξ, а d и p – вектор-строки, собранные по линейной части (1).

Функция Лагранжа для задачи (5) имеет вид

$$L(x,\xi,\lambda) = 0.5 \cdot x^T x + d\,x + p\,\xi + \lambda(b - Ax - G\xi),\ \lambda \geq 0.$$

Необходимые условия оптимальности Каруша-Куна-Таккера имеют вид

$$\frac{\partial L}{\lambda_i} = \begin{cases} b_i - a^i x - g^i \xi = 0,\ \lambda_i \geq 0 \\ b_i - a^i x - g^i \xi < 0,\ \lambda_i = 0 \end{cases},\ i \in I = \{1,\ldots,m\},\ \begin{matrix} \nabla_x L = x + (d - \lambda A)^T = 0 \\ \nabla_\xi L = (p - \lambda G)^T = 0 \end{matrix}, \qquad (6)$$

где – a^i, g^i строки i матриц A и G, m – количество строк ограничений в (5).

Суть метода активного набора можно сформулировать как поиск множества $I_a^* \subset I$ ограничений задачи (5), которые в оптимальном решении выполнены как равенства. Метод называется двойственным потому, что процессе поиска сохраняется допустимость двойственных переменных λ.

Для произвольного активного набора I_a переменные x, ξ, λ_a определяются решением следующей линейной системы, полученной из (6):

$$\begin{pmatrix} E & 0 & -A_a^T \\ 0 & 0 & -G_a^T \\ A_a & G_a & 0 \end{pmatrix} \begin{pmatrix} x \\ \xi \\ \lambda_a^T \end{pmatrix} = \begin{pmatrix} -d^T \\ -p^T \\ b_a \end{pmatrix}, \qquad (7)$$

где A_a, G_a – подматрицы матриц A, G, b_a, λ_a – части векторов b, λ, соответствующие подмножеству I_a. Остальные переменные λ_i полагаются равными 0.

Оптимальность набора I_a проверяется по оставшимся условиям из (6).

Необходимым условием разрешимости системы (7) есть наличие в ней хотя бы одного ограничения, включающего переменные ξ_1, ξ_2, и не более n+2 ограничений из системы (5), где n – длина вектора x.

При поиске оптимального активного набора система (7) должна решаться многократно прямо и косвенно. Поэтому важно рационально изменять систему и пересчитывать ее решение, основываясь на предыдущем наборе и решении. Для этого выполним преобразование пространства переменных x таким образом, чтобы система активных ограничений приняла вид

$$A_a x + G_a \xi = A_a Q y + G_a \xi = (L \;\; 0) y + G_a \xi = b_a ,$$

где Q – ортогональная матрица преобразования пространства $x = Qy$; $(L \;\; 0)$ – матрица, состоящая из нижней треугольной L и нулевой подматриц. Такое преобразование возможно, если ранг A_a равен числу ее строк, т.е. ограничение ξ_2=0 не активно.

В преобразованном пространстве изменяется вид градиентов функции Лагранжа по переменным y, λ, а именно:

$$\nabla_\lambda L = b - \begin{pmatrix} L & 0 \\ S & \end{pmatrix} y - G\xi , \;\; \nabla_y L = y + (dQ)^T - \begin{pmatrix} L^T & S^T \\ 0 & \end{pmatrix} \lambda^T , \qquad (8)$$

где $\begin{pmatrix} L & 0 \\ S & \end{pmatrix} = AQ$.

В таком представлении вектор y разделяется на компоненты, соответствующие активным ограничениям I_a и неактивным: $y = (y_a \;\; y_N)$. Более того можно установить взаимно-однозначное соответствие между переменными y_a и ограничениями I_a.

Система (7) приобретает вид

$$\begin{pmatrix} E_a & 0 & 0 & -L^T \\ 0 & E_N & 0 & 0 \\ 0 & 0 & 0 & -G_a^T \\ L & 0 & G_a & 0 \end{pmatrix} \begin{pmatrix} y_a \\ y_N \\ \xi \\ \lambda_a^T \end{pmatrix} = \begin{pmatrix} -(dQ)_a^T \\ -(dQ)_N^T \\ -p^T \\ b_a \end{pmatrix}, \qquad (9)$$

и решается в так: $y_a = L^{-1}(b_a - G_a\xi)$, $\lambda_a^T = L^{-1T}(y_a + (dQ)_a^T)$, $p^T = G_a^T\lambda_a^T = G_a^T L^{-1T}(y_a + (dQ)_a^T) = G_a^T L^{-1T}\left(L^{-1}(b_a - G_a\xi) + (dQ)_a^T\right)$. Из последнего находим $\xi = \left(G_a^T L^{-1T} L^{-1} G_a\right)^{-1}\left(G_a^T L^{-1T}\left(L^{-1}b_a + (dQ)_a^T\right) - p^T\right)$. Далее, подставляя ξ, находим y_a и λ_a^T. y_N считается независимо $y_N = -(dQ)_N^T$. Несмотря на некоторую громоздкость формул система решается эффективно, так как не используется явное обращение треугольной матрицы L, матрица $L^{-1}G_a$ имеет размер $|I_a| \times 2$, а матрица $\left(G_a^T L^{-1T} L^{-1} G_a\right) - 2 \times 2$.

Если ограничение $\xi_2 = 0$ активно, то оно добавляется к системе (9), а также добавляется столбец соответствующей двойственной переменной λ_{ξ_2}, содержащий единственный ненулевой элемент – 1 в строке, соответствующей ξ_2. В этом случае первыми находятся переменные ξ_1 и λ_{ξ_2} из системы

$$\begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} = \left(G_a^T L^{-1T} L^{-1} G_a\right)^{-1}\left(G_a^T L^{-1T}\left(L^{-1}b_a + (dQ)_a^T\right) - p^T + \begin{pmatrix} 0 \\ \lambda_{\xi_2} \end{pmatrix}\right).$$

Если активный набор не является оптимальным, то алгоритм выполняет его изменение путем добавления нарушенных ограничений из (5) и вывода ограничений, у которых двойственные переменные отрицательны. Порядок ввода, вывода ограничений и пересчета системы (9) зависит от конкретной реализации алгоритма. Но при любой реализации необходимо выполнить соответствующее изменение матриц Q и L. И если матрица L может иметь небольшую размерность, то матрица Q имеет размерность $n \times n$.

Если среди активных ограничений есть границы исходных переменных и соответствующие ограничения записаны первыми, то равенство $A_a Q = (L \;\; 0)$ более детально представляется в виде

$$A_a Q = \begin{pmatrix} P \\ B \end{pmatrix}\begin{pmatrix} P^T & D \end{pmatrix} = \begin{pmatrix} E & 0 \\ BP^T & L' \end{pmatrix}\begin{matrix} 0 \\ 0 \end{matrix} = (L \;\; 0),$$

где P – строки ограничений (5), соответствующие границам переменных, B – строки «обычных» ограничений, D – правая часть столбцов матрицы Q.

Каждая строка в P содержит только один ненулевой элемент +1 или – 1. Отсюда с учетом того, что $PD = 0$, следует, что строки Q, содержащие +1, – 1 подматрицы P^T, не содержат других ненулевых элементов. Следовательно, при соответствующем (Π_x) изменении порядка переменных матрица Q может быть представлена в виде $A_a Q = A_a \Pi_x \Pi_x^T Q = A_a \Pi_x \begin{pmatrix} E' & 0 \\ 0 & Q' \end{pmatrix}$, т.е. $Q = \Pi_x \begin{pmatrix} E' & 0 \\ 0 & Q' \end{pmatrix}$, где E' – диагональная матрица с элементами +1, – 1.

Таким образом наличие границ переменных среди активных ограничений квадратичной подзадачи и диагональность квадратичной матрицы позволяют значительно уменьшить размер обрабатываемой рабочей матрицы Q и ускорить решение подзадачи.

Общая постановка двухэтапной задачи стохастического программирования

Рассмотрим исходную задачу выпуклого программирования в несколько иной постановке, а именно:

$$f_0(x) + F_0(x) \to \min_x, \tag{10}$$

$$f_j(x) + F_j(x) \le 0, \ 1 \le j \le m, \tag{11}$$

$$x \in M, \tag{12}$$

где $F_j(x), \ 0 \le j \le m$ некоторые статистические функции Φ_ω^j от случайных кусочно-линейных функций $Q^j(x, \omega)$, здесь $\omega \in \Omega^j$ – случайная величина. В качестве Φ_ω^j обычно рассматривается среднее. Наряду со средним мы будем рассматривать также статистические функции CVaR, Partial moment, Mean Absolute Penalty [3, 7].

Как правило, в двухэтапных задачах стохастического программирования используется только одна функция статистическая функция $F_j(x)$ либо в

целевой функции, либо в ограничениях. Поэтому, не уточняя, где используется функция $F_j(x)$, индекс j будем опускать.

Функция $Q(x,\omega)$ вычисляется в свою очередь как решение линейной подзадачи второго этапа при фиксированных значениях переменных x и случайной величины ω. Таким образом имеем $F(x) = \Phi_\omega Q(x,\omega),\ \omega \in \Omega,$ где

$$\begin{gathered} Q(x,\omega) = \min_{y \in Y_\omega} \{r_\omega x + q_\omega y\}, \\ T_\omega x + W_\omega y = h_\omega. \end{gathered} \tag{13}$$

Поскольку при некоторых значениях переменных x подзадача (13) может не иметь допустимого решения, что сразу осложняет процесс решения исходной задачи (10) – (12), то введем в задачу (13) дополнительные вектора переменных $s^+, s^- \geq 0$ с большими штрафными коэффициентами, но позволяющие при любом x выполнить ограничения в (13). Тогда задача (13) получает вид

$$\begin{gathered} Q(x,\omega) = \min_{y \in Y_\omega} \{r_\omega x + q_\omega y + eu(s^+ + s^-)\}, \\ T_\omega x + W_\omega y + Is^+ - Is^- = h_\omega, \\ s^+,\ s^- \geq 0, \\ e >> 1, \end{gathered} \tag{14}$$

где I – единичная матрица; u – вектор с компонентами (1, 1, … , 1), e – величина штрафа за использование переменных из векторов s^+, s^-, Y_ω – допустимое множество, заданное границами переменных y, зависящими от случайной величины.

Для организации более эффективного вычислительного процесса задачу (14) лучше представить в двойственном виде. Если границы на переменные y отсутствуют, то двойственная задача имеет такой вид:

$$\begin{gathered} Q(x,\omega) = \max_{\pi} \{r_\omega x + \pi(h_\omega - T_\omega x)\}, \\ \pi W_\omega = q_\omega, \\ -ue \leq \pi \leq ue. \end{gathered} \tag{15}$$

В отличие от постановки (14) в этой постановке от переменных x зависит только целевая функция. Поэтому при изменении переменных x предыдущее оптимальное решение задачи (15) остается хорошим допустимым начальным приближением для поиска нового оптимального решения.

Значение функции $Q(x,\omega)$ и ее градиент $\nabla Q(x,\omega)$ равны

$$Q(x,\omega) = r_\omega x + \pi^*_\omega(x)(h_\omega - T_\omega x),$$
$$\nabla Q(x,\omega) = -\pi^*_\omega(x) T_\omega.$$

Если пространство Ω случайной величины является дискретным и конечным, то мы имеем задачу с дискретным множеством сценариев $\omega = 1, \ldots, J$. В этом случае указанные статистические функции вычисляются как линейные комбинации функций $Q(x,\omega)$ по общему принципу $F(x) = \sum_{\omega=1}^{J} k_\omega Q(x,\omega)$, хотя коэффициенты k_ω постоянны только при вычислении среднего, а при вычислении других статистических функций варьируются.

Если задачи (15) для всех сценариев были решены для некоторой точки x_0, то в PNK-методе порождается линейная оценка соответствующей функции по точке x_0 (левая часть (2) или (3)) в виде

$$L_F(x) = \sum_{\omega=1}^{J} k_\omega (r_\omega x_0 + \pi^*_\omega(x_0)(h_\omega - T_\omega x)).$$

Пример

В качестве примера рассмотрим задачу, описанную в [8]: минимизировать стоимость создания сети, для которой вероятность прохождения случайных потоков не меньше заданной величины. В этой задаче на первом этапе принимаются решения о создании дуг сети с определенной пропускной способностью, а на втором этапе решается ряд потоковых задач, соответствующие сценариям, описывающим потребности в транспортировке. Максимальное количество сценариев, которые рассматриваются равно 700, при этом суммарное количество ограничений равно 66503, а сумма количества пере-

менных первого уровня и переменных второго уровня всех сценариев равна 196014. Время решения этой задачи при дополнительной обработке сценариев, а именно, частичном упорядочивании, составляет порядка 15 сек на компьютере с процессором 2.7 GHz.

Список литературы

[1] Пшеничный Б.Н., Ненахов Э.И., Кузьменко В.Н. Комбинированный метод решения общей задачи выпуклого программирования // Кибернетика и системный анализ, 1998. – № 4. – С. 121 – 134.

[2] Кузьменко В.Н., Бойко В.В. О применении комбинированного метода выпуклого программирования // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2003. – С. 19 – 24.

[3] Бойко В.В., Кузьменко В.Н. Применение комбинированного метода выпуклого программирования в задачах финансовой математики // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2008. – С. 146 – 152.

[4] Бойко В.В., Кузьменко В.Н. Использование квадратичного приближения функций в PNK-методе // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2010. – С. 120 – 125.

[5] Кузьменко В.Н., Бойко В.В. Использование PNK-метода для решения невыпуклых задач оптимизации // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2012. – С. 47 – 52.

[6] Noll D., Prot O., Rondepierre A. A proximity control algorithm to minimize nonsmooth and nonconvex functions // Pac. J. Optim., 2008. – Vol. 4 (4). – Pp. 569 – 602.

[7] Rockafellar R.T., Uryasev S.P. Optimization of Conditional Value-at-Risk // J. of Risk., 2000. – №2. – P. 21 – 41.

[8] Ruszczynski, A. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedral // Math. Programming, Ser. A 93., 2002. – Pp. 195 – 215.

УДК 519.8

УСКОРЕННЫЕ МОДИФИКАЦИИ СУБГРАДИЕНТНОГО МЕТОДА ПОЛЯКА ДЛЯ ОВРАЖНЫХ ВЫПУКЛЫХ ФУНКЦИЙ

Стецюк П. И.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: stetsyukp@gmail.com

Аннотация. Обсуждаются три субградиентных метода с преобразованием пространства для нахождения точки минимума выпуклой функции при известном минимальном значении функции. В методах используется шаг Поляка в направлении нормированного антисубградиента и гарантируется монотонное уменьшение расстояния до точки минимума в последовательно преобразованных пространствах переменных. Преобразование пространства реализуется с помощью однорангового эллипсоидального оператора и направлено на уменьшение степени овражности поверхностей уровня выпуклых функций.

Ключевые слова. Субградиентный метод, шаг Поляка, шаг Агмона-Моцкина-Шенберга, овражная функция, преобразование пространства.

Введение

Данная статья связана с использованием линейных неортогональных преобразований пространства для улучшения обусловленности овражных функций – идеи Н.З. Шора, на которой построены эффективные модификации *r*-алгоритмов. В ней рассмотрены субградиентные методы с шагом Поляка и преобразованием пространства, обеспечивающие монотонное уменьшение расстояния до точки минимума. Эти методы можно рассматривать как расширение предложенного Б.Т. Поляком в 1969 году субградиентного метода для нахождения точки минимума выпуклой функции при известном оптимальном значении функции. В них длина шага в направлении антисубградиента вычисляется (а не подбирается как в *r*-алгоритмах из решения одномерной задачи минимизации функции по направлению), причем для ее вычисления требуется лишь знание минимального значения функции.

Субградиентный метод Поляка

Пусть $f(x)$ – выпуклая функция векторного аргумента $x \in R^n$ и ее субградиент $\partial f(x)$ удовлетворяет условию:

$$(x - x^*, \partial f(x)) \geq m(f(x) - f^*), \quad \forall x \in R^n, \forall x^* \in X^*, \quad m \geq 1. \qquad (1)$$

Здесь R^n – евклидово пространство размерности n со скалярным произведением (x,y); X^* – множество точек минимума функции $f(x)$; f^* – минимальное значение функции $f(x)$: $f^* = f(x^*)$, $x^* \in X^*$. Параметр m считается известным, он введен для учета специальных классов выпуклых функций, например, для квадратичной гладкой функции $m = 2$. Значение $m = 1$ обеспечивает выполнение условия (1) для произвольной выпуклой функции.

Пусть известно f^*. Для нахождения точки $x^* \in X^*$ применяется субградиентный метод Поляка [1,2]:

$$x_{k+1} = x_k - h_k \frac{\partial f(x_k)}{\|\partial f(x_k)\|}, \quad h_k = \frac{m\left(f(x_k) - f^*\right)}{\|\partial f(x_k)\|}, \quad k = 0,1,2,\ldots. \quad (2)$$

При $m = 1$ метод (2) имеет простой геометрический смысл. Функция $f(x)$ аппроксимируется линейной $\tilde{f}(x) = f(x_k) + (f'(x_k), x - x_k)$ и шаг выбирается так, чтобы эта аппроксимирующая функция стала равной f^* (т.е. $\tilde{f}(x_{k+1}) = f^*$). Для одномерного случая метод Поляка совпадает с методом Ньютона для решения уравнения $f(x) = f^*$.

Величину h_k называют шагом Поляка или шагом Агмона-Моцкина-Шенберга. Этот шаг тесно связан с результатами И.И. Еремина о сходимости итерационных методов аппроксимации неподвижных точек с помощью операторов, обладающих свойством квазисжимаемости (фейеровости) [3]. Впервые такой шаг был использован в [4, 5] в релаксационном (субградиентном) методе для нахождения хотя бы одного решения совместной системы линейных неравенств.

Теорема 1. *Для всех точек, генерируемых методом* (2)*, справедливы неравенства*

$$\left\|x_{k+1} - x^*\right\|^2 \le \left\|x_k - x^*\right\|^2 - \left(\frac{m(f(x_k) - f^*)}{\|\partial f(x_k)\|}\right)^2, \quad \forall x^* \in X^*, k = 0,1,\ldots, \quad (3)$$

и неравенства

$$(x^* - x_{k+1}, -\partial f(x_k)) \geq 0, \quad \forall x^* \in X^*, k = 0,1,\ldots, \qquad (4)$$

Доказательство. Для любого $x^* \in X^*$ и произвольного k $(k \geq 0)$ имеем

$$\left\|x_{k+1} - x^*\right\|^2 = \left\|x_k - x^* - h_k \frac{\partial f(x_k)}{\left\|\partial f(x_k)\right\|}\right\|^2 = \left\|x_k - x^*\right\|^2 - 2h_k \frac{(x_k - x^*, \partial f(x_k))}{\left\|\partial f(x_k)\right\|} + h_k^2.$$

Учитывая, что из (1) следует неравенство

$$\frac{(x_k - x^*, \partial f(x_k))}{\left\|\partial f(x_k)\right\|} \geq \frac{m(f(x_k) - f^*)}{\left\|\partial f(x_k)\right\|} = h_k,$$

имеем

$$\left\|x_{k+1} - x^*\right\|^2 \leq \left\|x_k - x^*\right\| - h_k^2 = \left\|x_k - x^*\right\|^2 - \left(\frac{m(f(x_k) - f^*)}{\left\|\partial f(x_k)\right\|}\right)^2,$$

что дает неравенства (3).

Неравенства (4) следуют из того, что на основании (1) имеем

$$(x^* - x_{k+1}, -\partial f(x_k)) = (x_{k+1} - x^*, \partial f(x_k)) = (x_k - x^* - h_k \frac{\partial f(x_k)}{\left\|\partial f(x_k)\right\|}, \partial f(x_k)) =$$

$$= (x_k - x^*, \partial f(x_k)) - h_k \left\|\partial f(x_k)\right\| = (x_k - x^*, \partial f(x_k)) - m(f(x_k) - f^*) \geq 0.$$

Теорема 1 доказана.

В теореме 1 отражены два центральных свойства шага Поляка. Первое свойство следует из неравенств (4). Оно означает, что шаг h_k определяет величину максимального сдвига в направлении нормированного антисубградиента, при котором угол между антисубградиентом и направлением из точки x_{k+1} на произвольную точку из множества минимумов будет нетупым. Второе свойство следует из неравенств (3) и означает, что если множество X^* состоит из единственной точки x^*, то шаг h_k выбирается таким, чтобы расстояние от точки x_{k+1} к точке минимума x^* было минимальным.

Проблема овражности

В [1, 2] показано, что метод (2) сходится со скоростью геометрической прогрессии, и отмечается что во многих практических задачах знаменатель прогрессии может оказаться близким к единице. Для гладкого случая этот факт также хорошо известен (градиентные методы медленно сходятся для функций овражного типа). Тот же эффект имеет место и в негладком случае, например, для функции двух переменных $f(x) = |x_1| + t|x_2|$ метод (2) сходится со скоростью геометрической прогрессии со знаменателем $\left(1 - \frac{1}{t^2}\right)^{\frac{1}{2}}$, т. е. очень медленно для больших t.

Медленную скорость сходимости метода Поляка проиллюстрируем на примере трех выпуклых функций от двух переменных, первые две функции – негладкие (овражная и существенно овражная), а третья – квадратичная овражная функция. В табл. 1 приведено количество итераций метода (2) для нахождения последовательно уточняемых приближений к единственной точке минимума $x^* = (0, 0)$. Эти приближения заданы десятью последовательно уменьшающимися (на порядок) значениями epsf = ε_f (первая колонка в табл. 1). Метод Поляка прекращал работу или на итерации $k = itn$, для которой $f(x_{itn}) - f^* \le \varepsilon_f$, или если превышено максимальное количество итераций, равное 20000.

Колонки itn1, itn2 и itn3 отвечают овражной кусочно-линейной функции $f_1(x_1, x_2) = |x_1| + t|x_2|$ для значений t равных 3, 9 и 27 соответственно. При расчетах использовалась стартовая точка $x_0 = (1, 1)$ и параметр $m = 1$. Из табл. 1 видим, что количество итераций существенно увеличивается с ростом t. Для того чтобы найти точку, где значение функции $f_1(x_1, x_2) = |x_1| + 27|x_2|$ отличается от минимального $f_1^* = 0$ не более, чем на $\varepsilon_f = 10^{-10}$, методу Поляка потребовалось 8634 итерации. Такая

же точность для функции $f_1(x_1,x_2)=|x_1|+t|x_2|$ при $t=100$ не обеспечивается даже при ста тысячах итераций метода Поляка.

Таблица 1

Сходимость метода Поляка для двумерных овражных функций

epsf	itn1	itn2	itn3	itn4	itn5	itn6	itn7
1.0e-01	14	119	1080	16	6	6	6
1.0e-02	24	212	1919	162	10	10	10
1.0e-03	34	305	2759	1604	12	12	12
1.0e-04	45	398	3598	16004	16	16	16
1.0e-05	55	492	4437	20000	20	20	20
1.0e-06	65	585	5277	20000	22(42)	22(42)	22(42)
1.0e-07	76	678	6116	20000	26(48)	26(52)	26(52)
1.0e-08	86	771	6955	20000	28(54)	30(56)	30(56)
1.0e-09	96	865	7795	20000	32(62)	32(62)	32(62)
1.0e-10	107	958	8634	20000	36(68)	36(70)	36(70)

Для $f_2(x_1,x_2)=\max\left\{x_1^2+(2x_2-2)^2-3, x_1^2+(x_2+1)^2\right\}$ – существенно овражной кусочно-квадратичной функции зигзаго-образная траектория последовательных приближений метода Поляка дана на рис. 1. Здесь метод Поляка находит приближение, где значение функции $f_2(x_1,x_2)$ отличается от минимального $f_2^*=1$ не более, чем на 0.0001, только за 16004 итераций (см. колонку itn4 в табл. 1).

Для функций $f_1(x_1,x_2)$ и $f_2(x_1,x_2)$ использовался параметр $m=1$, который можно применять для произвольной выпуклой функции. Для квадратичной функции $f_3(x_1,x_2)=x_1^2+tx_2^2, t>0$ использовался параметр $m=2$. Он в два раза увеличивает длину шага Поляка, соответствующего параметру $m=1$, и обеспечивает более быструю скорость сходимости с различных начальных приближений. Количество итераций метода Поляка со стартовой точки $x_0=(0,0)$ приведено в колонках itn5, itn6 и itn7 и соответствует значениям t равным 100, 10000 и 1000000. Количество

итераций для всех 10 значений epsf одинаково и не зависит от степени вытянутости поверхности квадратичной функции $f_3(x_1,x_2)$. Отличия наблюдаются лишь при достаточно малых $\varepsilon_f = 10^{-12},10^{-14},10^{-16},10^{-18},10^{-20}$, для которых количество итераций приведено в скобках после количеств итераций для шести последних значений epsf. Чтобы найти точку минимума квадратичной функции $f_3(x_1,x_2)$ с такой же точностью, как и точку минимума функции $f_1(x_1,x_2)$, нужно взамен ε_f использовать ε_f^2.

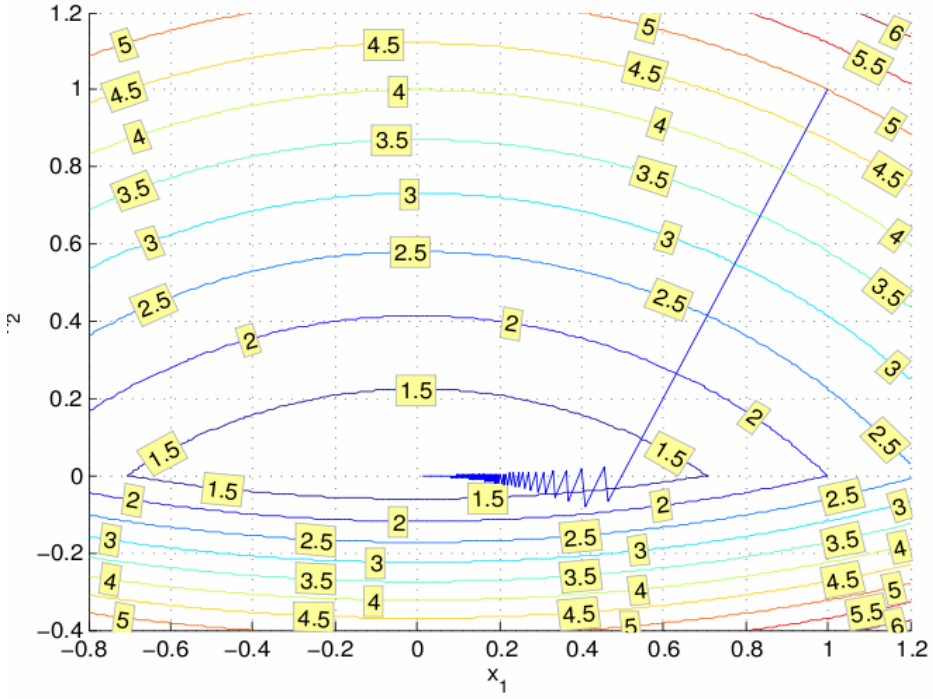


Рис. 1. Траектория метода Поляка для $f_2(x_1,x_2)$ ($x_0 = (1,1)$)

Следовательно, вопрос об ускорении сходимости метода (2) является актуальным и здесь отметим два способа его ускорения. Первый способ состоит в том, чтобы вблизи минимума функцию $f(x)$ аппроксимировать кусочно-линейной функцией, используя информацию из предыдущих итераций. Самый общий метод, реализующий этот способ, предложен в [2] и имеет следующий вид:

$$x_{k+1} = P_{Q_k}(x_k), \quad Q_k = \{x : f(x_i) + (\partial f(x_i), x - x_i) \le f^*, \quad i \in I_k\}, \qquad (2a)$$

где I_k – любое подмножество индексов из $0,1,\ldots,k$, которое обязательно содержит индекс k, P_{Q_k} – оператор проектирования на множество Q_k.

Метод (2а) также сходится со скоростью геометрической прогрессии и не медленнее, чем метод (2), а его частные случаи очень тесно связаны с другими известными методами. Так, если $I_k = \{0, 1, \dots, k\}$, то метод (2а) дает точный минимум для кусочно-линейной функции, причем он идейно близок к методу Келли [6]. Если множество $I_k = \{k-1, k\}$, то проекцию на множество Q_k можно выписать явно. В этом случае метод (2а) близок к методу из [7], который использует всего два вектора и для построения очередного направления спуска использует линейную комбинацию направления субградиента и направления движения на предыдущем шаге (по типу сопряженных градиентов) с тем, чтобы оно составляло более острый угол с направлением на минимум.

Второй способ ускорения метода (2) базируется на идее Н.З. Шора, связанной с использованием линейных неортогональных преобразований пространства для улучшения обусловленности овражных функций. На этой идее построены эффективные модификации *r*-алгоритмов [8, 9]. В статье рассмотрим субградиентные методы с шагом Поляка и преобразованием пространства переменных, которое обеспечивает монотонное уменьшение расстояния до точки минимума и направлено на уменьшение степени овражности поверхностей уровня выпуклых функций подобно тому, как это сделано в *r*-алгоритмах Шора.

Метод Поляка с преобразованием пространства

Пусть произведена замена переменных $x = By$, где B – неособенная $n \times n$-матрица (т.е. существует обратная матрица $A = B^{-1}$). Субградиент выпуклой функции $f(x)$ в точке x_k удовлетворяет неравенству

$$f(x) \geq f(x_k) + (\partial f(x_k), x - x_k) \quad \forall x \in R^n,$$

откуда, осуществляя замену переменных $x = By$, получаем

$$\phi(y) \geq \phi(y_k) + (B^T \partial f(x_k), y - y_k) \quad \forall y \in R^n,$$

где $\phi(y) = f(By)$. Вектор $\partial\phi(y_k) = B^T \partial f(x_k)$ удовлетворяет неравенству

$$\phi(y) \geq \phi(y_k) + (\partial\phi(y_k), y - y_k) \quad \forall y \in R^n$$

и является субградиентом выпуклой функции $\phi(y)$ в точке $y_k = Ax_k$ преобразованного пространства переменных $y = Ax$.

Для нахождения точки $x^* \in X^*$ субградиентный метод Поляка с преобразованием пространства (определяется невырожденной матрицей B) имеет следующий вид:

$$x_{k+1} = x_k - h_k \frac{B^T \partial f(x_k)}{\left\| B^T \partial f(x_k) \right\|}, \quad h_k = \frac{m\left(f(x_k) - f^*\right)}{\left\| B^T \partial f(x_k) \right\|}, \quad k = 0,1,2,\ldots. \qquad (5)$$

Здесь величина h_k – шаг Поляка (шаг Агмона-Моцкина-Шенберга), но в преобразованном пространстве переменных $y = Ax$. Это следует из того, что в преобразованном пространстве переменных метод (5) записывается как субградиентный процесс

$$y_{k+1} = y_k - h_k \frac{\partial\phi(y_k)}{\left\| \partial\phi(y_k) \right\|}, \quad h_k = \frac{m(\phi(y_k) - \phi^*)}{\left\| \partial\phi(y_k) \right\|}, \quad k = 0,1,2,\ldots. \qquad (6)$$

Шаг Поляка в преобразованном пространстве переменных обладает такими же свойствами как и шаг Поляка в исходном пространстве. Это следует из априорного знания минимального значения функции и связанного с ним неравенства (1).

Теорема 2. *Пусть* $A = B^{-1}$. *Для всех точек, генерируемых методом* (5)*, справедливы неравенства*

$$\left\| A(x_{k+1} - x^*) \right\|^2 \leq \left\| A(x_k - x^*) \right\|^2 - \left(\frac{m(f(x_k) - f^*)}{\left\| B^T \partial f(x_k) \right\|} \right)^2, \forall x^* \in X^*, k = 0,1,\ldots, \qquad (7)$$

и неравенства

$$(A(x^* - x_{k+1}), -B^T \partial f(x_k)) \geq 0, \quad \forall x^* \in X^*, k = 0,1,\ldots. \qquad (8)$$

Доказательство теоремы 2 аналогично доказательству теоремы 1.

Очевидно, что если матрицу B выбрать такой, чтобы вытянутые поверхности уровня овражной функции в преобразованном пространстве

переменных становились менее вытянутыми, то субградиентный метод Поляка с преобразованием пространства (5) окажется эффективнее, чем субградиентный метод Поляка без преобразования пространства (2). Это согласуется с количеством итераций метода (5) для нахождения десяти последовательно уточняемых приближений к точке минимума функции $f_1(x_1, x_2) = |x_1| + 10|x_2|$ для шести различных матриц B, каждой из которых соответствует свой столбец itn в табл. 2. Матрицы B получены в результате растяжения пространства переменных в направлении x_2 с коэффициентами растяжения $\alpha = 1; 1.5; 2; 3; 4; 5$. Матрица B имеет вид

$$B = \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{1}{\alpha} \end{pmatrix}$$

и если $\alpha = 1$, то она совпадает с единичной матрицей. Это сооветствует случаю, когда метод Поляка с преобразованием пространства тождественно равен методу (2) – методу Поляка без преобразования пространства. В табл. 2 ему отвечает столбец itn1, где приведено число итераций метода Поляка (2) для нахождения последовательно уточняемых приближений к единственной точке минимума $x^* = (0, 0)$.

Из табл. 2 видим, что количество итераций метода Поляка с преобразованием пространства монотонно уменьшается по мере того, как уменьшается степень овражности функции $\phi_1(y_1, y_2) = |y_1| + \dfrac{10}{\alpha}|y_2|$ в преобразованном пространстве переменных (соответствует увеличению коэффициента растяжения α).

Если для овражной негладкой функции можно подобрать такую матрицу B, чтобы ускорить сходимость метода Поляка, то для существенно овражной функции это сделать практически невозможно. Об этом говорят результаты о числе итераций метода Поляка с преобразованием пространства для тех же шести матриц B, что и раньше, но для существенно овражной функции $f_2(x_1, x_2) = \max\left\{x_1^2 + (2x_2 - 2)^2 - 3, x_1^2 + (x_2 + 1)^2\right\}$. Из табл. 3 видим,

что монотонного уменьшения количества итераций по мере увеличения α – коэффициента растяжения пространства в направлении x_2 здесь не наблюдается. Зато наблюдается некоторый разрыв, который происходит при коэффициенте $\alpha = 2$ (соответствует столбцу itn3). При всех остальных коэффициентах α метод требует максимального количества итераций для достижения точности $\varepsilon_f = 10^{-6}$.

Таблица 2

Сходимость метода Поляка с преобразованием пространства для кусочно-линейной функции $f_1(x_1, x_2) = |x_1| + 10|x_2|$

epsf	itn1	itn2	itn3	itn4	itn5	itn6
1.0e-01	147	63	33	6	10	9
1.0e-02	262	114	62	19	17	13
1.0e-03	377	165	91	31	24	18
1.0e-04	492	216	119	44	31	22
1.0e-05	607	268	148	57	38	27
1.0e-06	722	319	177	70	45	31
1.0e-07	837	370	206	82	53	36
1.0e-08	952	421	234	95	60	40
1.0e-09	1000	472	263	108	67	45
1.0e-10	1000	523	292	121	74	49

Таблица 3

Сходимость метода Поляка с преобразованием пространства для кусочно-квадратичной $f_2(x_1, x_2) = \max\left\{x_1^2 + (2x_2 - 2)^2 - 3, x_1^2 + (x_2 + 1)^2\right\}$

epsf	itn1	itn2	itn3	itn4	itn5	itn6
1.0e-01	16	4	4	5	7	8
1.0e-02	162	37	4	6	7	9
1.0e-03	1604	679	5	6	8	9
1.0e-04	10000	7079	5	6	9	446
1.0e-05	10000	10000	6	8	8061	6206
1.0e-06	10000	10000	6	10000	10000	10000
1.0e-07	10000	10000	6	10000	10000	10000
1.0e-08	10000	10000	6	10000	10000	10000
1.0e-09	10000	10000	6	10000	10000	10000
1.0e-10	10000	10000	6	10000	10000	10000

Ниже рассмотрим две модификации метода Поляка с преобразованием пространства, где матрица B будет изменяться, если обнаружена возможность зигзагообразного движения вдоль русла оврага. В этих

модификациях будет использована одноранговая коррекция несимметричной матрицы B только на тех итерациях, когда угол между двумя последовательными субградиентами тупой.

Одноранговый эллипсоидальный оператор

Тупой угол между двумя нормированными векторами ξ и η из R^n можно преобразовать в прямой с помощью линейного оператора из R^n в R^n, который в матричной форме представим

$$T_1(\xi,\eta) = I - \frac{1}{1-(\xi,\eta)^2}\left(\left(1-\sqrt{1-(\xi,\eta)^2}\right)\eta - (\xi,\eta)\xi\right)\eta^T. \qquad (9)$$

Здесь $\xi,\eta \in R^n$ – векторы такие, что $\|\xi\| = 1$, $\|\eta\| = 1$ и их скалярное произведение удовлетворяет условию $(\xi,\eta)^2 \neq 1$, I – единичная матрица размера $n \times n$. Для оператора $T_1(\xi,\eta)$ существует обратный $T_1^{-1}(\xi,\eta)$

$$T_1^{-1}(\xi,\eta) = I + \frac{1}{\sqrt{1-(\xi,\eta)^2}}\left(\left(1-\sqrt{1-(\xi,\eta)^2}\right)\eta - (\xi,\eta)\xi\right)\eta^T \qquad (10)$$

Оператор $T_1(\xi,\eta)$ введен в [10] под названием "одноранговый эллипсоидальный оператор". Он связан с преобразованием в шар специального эллипсоида, описанного вокруг тела W (см. рис. 2), которое получено в результате пересечения шара и двух полупространств, проходящих через центр шара. В случае тупого угла между нормалями полупространств этот эллипсоид содержит тело W и является минимальным по объему в рамках семейства эллипсоидов, центр которых совпадает с центром шара (рис. 3).

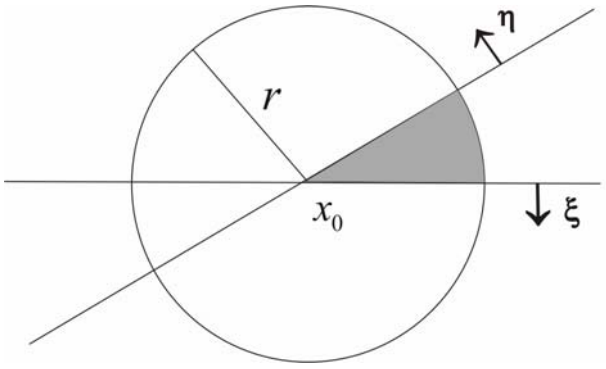


Рис. 2. Тело W (проекция на плоскость)

Минимальный обьем эллипсоида меньше, чем обьем шара, и это уменьшение определяется величиной $\sqrt{1-(\xi,\eta)^2}$. Преобразование минимального по объему эллипсоида в шар требует растяжения пространства в направлении $\frac{\xi-\eta}{\|\xi-\eta\|}$ с коэффициентом $\alpha_1=\frac{1}{\sqrt{1+(\xi,\eta)}}$ и "сжатия" пространства в ортогональном направлении $\frac{\xi+\eta}{\|\xi+\eta\|}$ с коэффициентом $\alpha_2=\frac{1}{\sqrt{1-(\xi,\eta)}}$. В преобразованном пространстве эллипсоид станет шаром, а образы векторов ξ и η будут ортогональными (см. рис. 4).

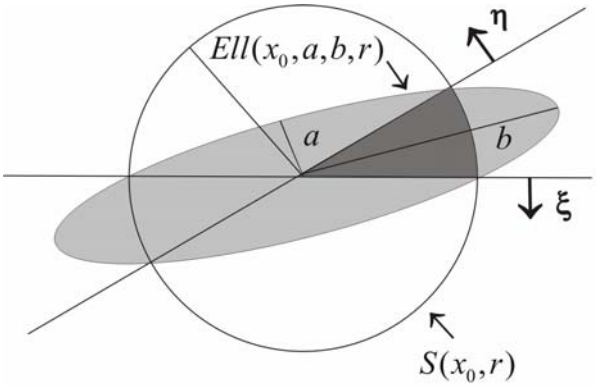


Рис. 3. Специальный эллипсоид, содержащий тело W

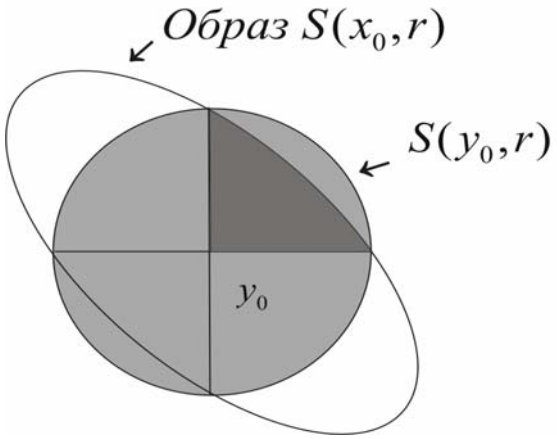


Рис. 4. Оптимальный эллипсоид после преобразования

Ускоренный субградиентный метод Поляка

Медленную сходимость метода Поляка для овражных функций определяет угол между двумя последовательными субградиентами: $\partial f(x_k)$ и

$\partial f(x_{k+1})$. Чем ближе этот угол к 180 градусам, тем более медленной будет скорость сходимости. Тупой угол между векторами $\partial f(x_k)$ и $\partial f(x_{k+1})$ можно преобразовать в прямой с помощью "однорангового эллипсоидального оператора" [10]. Если на каждой итерации избавляться от тупого угла между последовательными субградиентами в преобразованном пространстве переменных, то для овражных функций следует ожидать повышения скорости сходимости метода. Этот принцип реализован в изложенной ниже модификации метода Поляка.

Ускоренный субградиентный метод Поляка имеет следующий вид:

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T \partial f(x_k)}{\left\| B_k^T \partial f(x_k) \right\|}, \quad h_k = \frac{m(f(x_k) - f^*)}{\left\| B_k^T \partial f(x_k) \right\|}, \quad k = 0,1,2,\ldots, \quad (11)$$

где матрица $B_0 = I$, а матрица B_{k+1} размера $n \times n$ вычисляется по правилу:

$$B_{k+1} = \begin{cases} B_k, & \text{если } \mu_k \geq 0 \\ B_k + (B_k \eta)\xi_{k+1}^T, & \text{иначе} \end{cases} \quad (12)$$

где

$$\mu_k = (\xi_k, \xi_{k+1}), \quad \xi_k = \frac{B_k^T \partial f(x_k)}{\left\| B_k^T \partial f(x_k) \right\|}, \quad \xi_{k+1} = \frac{B_k^T \partial f(x_{k+1})}{\left\| B_k^T \partial f(x_{k+1}) \right\|},$$

$$\eta = \left(\frac{1}{\sqrt{1-\mu_k^2}} - 1 \right) \xi_{k+1} - \frac{\mu_k}{\sqrt{1-\mu_k^2}} \xi_k.$$

Метод (11), (12) естественно назвать ускоренным методом Поляка за счет антиовражного приема, подобного тому, который использован в r-алгоритмах Шора [8]. Действительно, на k-й итерации растяжение пространства производится в направлении разности нормированных последовательных субградиентов в преобразованном пространстве переменных: $y = A_k x = B_k^{-1} x$, где B_k – невырожденная матрица размерности $n \times n$. Если нормы субградиентов одинаковы, то это направление совпадает с разностью двух последовательных субградиентов, по которой реализуется растяжение пространства в r-алгоритмах. Отличие состоит в том, что в

преобразованном пространстве переменных для r-алгоритмов второй субградиент определяется согласно шагу наискорейшего спуска в направлении антисубградиента, а в методе (9), (10) – согласно шагу Поляка.

Теорема 3. *Пусть* $A_k = B_k^{-1}$, $A_{k+1} = B_{k+1}^{-1}$. *Для всех точек, генерируемых методом (11), (12), справедливы неравенства*

$$\left\|A_{k+1}(x_{k+1} - x^*)\right\|^2 \le \left\|A_k(x_k - x^*)\right\|^2 - \left(\frac{m(f(x_k) - f^*)}{\left\|B_k^T \partial f(x_k)\right\|}\right)^2, \forall x^* \in X^*, k = 0,1,\ldots \quad (13)$$

и неравенства

$$(A_k(x^* - x_{k+1}), -B_k^T \partial f(x_k)) \ge 0, \quad \forall x^* \in X^*, k = 0,1,\ldots. \quad (14)$$

Доказательство теоремы 3 аналогично доказательству теоремы 2 [10].

Преобразование пространства в методе (11), (12) направлено на уменьшение степени вытянутости поверхностей уровня овражных выпуклых функций. Неравенства (13) означают, что в каждом очередном преобразованном пространстве переменных гарантируется уменьшение расстояния до множества точек минимума. Поэтому для каждой итерации $k > 1$ имеет место неравенство

$$\left\|A_k(x_k - x^*)\right\|^2 \le \left\|x_0 - x^*\right\|^2 - \sum_{i=0}^{k-1}\left(\frac{m(f(x_i) - f^*)}{\left\|B_i^T \partial f(x_i)\right\|}\right)^2.$$

Для овражных функций детерминант матрицы B_k уменьшается, а следовательно, уменьшается обьем эллипсоида, локализующего точку x^*. Действительно, если на k-м шаге реализуется преобразование пространства, то

$$\det(B_{k+1}) = \det(B_k)\det(I + \eta\xi_{k+1}^T) = \det(B_k)\sqrt{1 - \mu_k^2} = \det(B_k)\sqrt{1 - \cos^2\varphi_k},$$

где φ_k – тупой угол между двумя последовательными субградиентами.

Для овражных негладких функций метод (11), (12) это обеспечивает ускоренную сходимость по отношению к методу (2). Так, например, для

кусочно-линейной функции двух переменных $f_1(x_1,x_2)=|x_1|+t|x_2|$ при любом значении параметра $t>1$ и произвольной стартовой точке x_0 метод (11), (12) находит точку минимума $x^*=(0,0)$ не более, чем за три итерации. Метод (2) сходится к x^* со скоростью геометрической прогрессии со знаменателем $\sqrt{1-1/t^2}$ и требует существенного количества итераций даже при сравнительно небольших значениях t. Траектория ускоренного метода Поляка для кусочно-квадратичной функции $f_2(x_1,x_2)$ дана на рис. 5, здесь метод находит точку, где $f_2(x)<1+10^{-5}$ всего за 16 итераций, а где $f_2(x)<1+10^{-10}$ – за 31 итерацию.

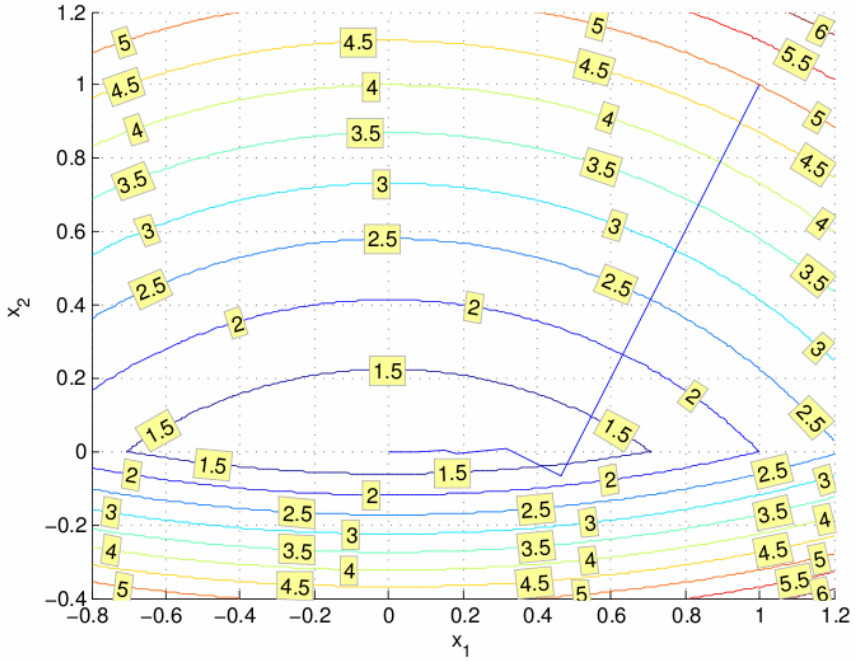


Рис. 5. Траектория ускоренного метода Поляка для $f_2(x_1,x_2)$ ($x_0=(1,1)$)

Метод amsg2p

Метод (11), (12) можно усилить более радикальным уменьшением обьема эллипсоида, локализующего точку x^*. Если на k-итерации была реализована операция преобразования пространства, то в преобразованном пространстве $Y=A_{k+1}X=B_{k+1}^{-1}X$ для функции $\phi_{k+1}(y)=f(B_{k+1}y)$ субградиенты будут ортогональны и в силу шага Поляка (шага Агмона-Моцкина-Шенберга) полупространства, определяемые ими в точке

$y_{k+2} = A_{k+1}x_{k+2}$, не отсекают точку $y^* = A_{k+1}x^*$. Очевидно, что если вычисленный в точке y_{k+2} субградиент образует тупой угол с двумя предыдущими, то можно уменьшить объем области локализации x^*. Для этого достаточно выбрать в качестве векторов, определяющих оператор $T_1(\xi,\eta)$, субградиент в точке y_{k+2} и вектор, являющийся выпуклой комбинацией первых двух, так, чтобы угол между ними был максимально тупым. Для очередной итерации (если реализуется преобразование пространства) эта ситуация повторяется, только в качестве первого вектора уже будет использоваться не субградиент, а вектор, являющийся выпуклой комбинацией двух предыдущих субградиентов в очередном преобразованном пространстве. Это позволяет замкнуть цикл вычислений и построить конструктивное правило использования в качестве одного из векторов, определяющих оператор $T_1(\xi,\eta)$, агрегированного вектора.

Метод, построенный на этом принципе условимся называть amsg2p [11, 12]. В его названии "ams" указывает на способ регулировки шага в направлении нормированного антисубградиента, а «g2p» указывает, что ams-шаг используется в пространстве переменных, преобразованном с помощью двух последних субградиентов (g2) и агрегатного вектора (p). Преобразование пространства реализуется с помощью однорангового эллипсоидального оператора и только на тех итерациях метода, когда тупым является хотя бы один из углов – угол между двумя последовательными субградиентами, либо угол между последним субградиентом и агрегатным вектором, который является выпуклой комбинацией вычисленных на предыдущих итерациях субградиентов.

В [13] метод amsg2p расширен на случай произвольного значения f_{min} и позволяет либо найти такую точку, где значение выпуклой функции $f(x)$ меньше или равно $f_{min} + \varepsilon$, либо гарантирует достаточное условие того, что точки, где значение $f(x)$ равно f_{min}, в шаре заданного радиуса не существует. В первом случае метод amsg2p находит точку

$x_\varepsilon^* \in \{x : f(x) - f_{min} \le \varepsilon\}$ и соответствующий ей номер итерации k_ε^*, а во втором – останавливается с сообщением «точки не существует». Метод amsg2p состоит в следующем.

На итерации $k = 0$ заданы: начальное приближение $x_0 \in R^n$; начальный радиус r_0 такой, что $\|x_0 - x^*\| \le r_0$; достаточно малое $\varepsilon > 0$. Вычислим $f(x_0)$ и $\partial f(x_0)$. Если $f(x_0) - f_{min} \le \varepsilon$, то $x_\varepsilon^* = x_0$, $k_\varepsilon^* = 0$ и окончание алгоритма. Иначе положим $h_0 = \dfrac{m(f(x_0) - f_{min})}{\|\partial f(x_0)\|}$, $\xi_0 = \dfrac{\partial f(x_0)}{\|\partial f(x_0)\|} \in R^n$, $p_0 = 0 \in R^n$, $B_0 = I$ – единичная $n \times n$ –матрица. Перейдем к следующей итерации.

Пусть на k-й итерации получены $x_k \in R^n$, h_k, r_k, $\xi_k \in R^n$, $p_k \in R^n$, B_k – матрица $n \times n$. Для $(k+1)$-й итерации выполним пп. 1 - 5.

1. Вычислим $t_k = h_k / r_k$. Если $t_k > 1$, то "точки не существует" и окончание алгоритма. Иначе положим $r_{k+1} = r_k \sqrt{1 - t_k^2}$ и вычислим очередное приближение

$$x_{k+1} = x_k - h_k B_k \xi_k .$$

2. Вычислим $f(x_{k+1})$ и $\partial f(x_{k+1})$. Если $f(x_{k+1}) - f_{min} \le \varepsilon$, то $x_\varepsilon^* = x_{k+1}$, $k_\varepsilon^* = k + 1$ и окончание алгоритма. Иначе положим

$$\xi_{k+1} = \frac{B_k^T \partial f(x_{k+1})}{\|B_k^T \partial f(x_{k+1})\|}, \quad h_{k+1} = \frac{m(f(x_{k+1}) - f_{min})}{\|B_k^T \partial f(x_{k+1})\|}.$$

3. Вычислим $\lambda_1 = -p_k^T \xi_{k+1}$ и $\lambda_2 = -\xi_k^T \xi_{k+1}$. Положим

$$p_{k+1} = \begin{cases} \dfrac{\lambda_1}{\sqrt{\lambda_1^2 + \lambda_2^2}} p_k + \dfrac{\lambda_2}{\sqrt{\lambda_1^2 + \lambda_2^2}} \xi_k, & \text{если } \lambda_1 > 0 \text{ и } \lambda_2 > 0, \\ p_k, & \text{если } \lambda_1 > 0 \text{ и } \lambda_2 \le 0, \\ \xi_k, & \text{если } \lambda_1 \le 0 \text{ и } \lambda_2 > 0, \\ 0, & \text{если } \lambda_1 \le 0 \text{ и } \lambda_2 \le 0. \end{cases}$$

4. Вычислим $\mu_k = p_{k+1}^T \xi_{k+1}$. Если $-1 < \mu_k < 0$, то вычислим

$$B_{k+1} = B_k + (B_k\eta)\xi_{k+1}^T, \text{ где } \eta = \left(\frac{1}{\sqrt{1-\mu_k^2}} - 1\right)\xi_{k+1} - \frac{\mu_k}{\sqrt{1-\mu_k^2}}\, p_{k+1},$$

и пересчитаем

$$h_{k+1} = \frac{h_{k+1}}{\sqrt{1-\mu_k^2}}, \qquad p_{k+1} = \frac{1}{\sqrt{1-\mu_k^2}}(p_{k+1} - \mu_k\xi_{k+1}).$$

Иначе положим $B_{k+1} = B_k$ и $p_{k+1} = 0$.

5. Перейдем к следующей итерации с x_{k+1}, h_{k+1}, r_{k+1}, ξ_{k+1}, p_{k+1}, B_{k+1}.

Теорема 4. *Пусть* $A_k = B_k^{-1}$, $A_{k+1} = B_{k+1}^{-1}$. *Если* $f_{min} \ge f^*$ *и* $X^* = x^*$, *то для всех точек, генерируемых методом amsg2p, справедливы неравенства*

$$\left\|A_{k+1}(x_{k+1} - x^*)\right\|^2 \le \left\|A_k(x_k - x^*)\right\|^2 - \left(\frac{m(f(x_k) - f_{min})}{\left\|B_k^T \partial f(x_k)\right\|}\right)^2, \; k = 0,1,\ldots, \qquad (15)$$

и неравенства

$$(A_k(x^* - x_{k+1}), -B_k^T \partial f(x_k)) \ge 0, \quad \forall x^* \in X^*, k = 0,1,\ldots. \qquad (16)$$

Теорема 4 означает, что в каждом очередном преобразованном пространстве переменных расстояние до точки минимума уменьшается. Благодаря этому для каждой итерации $k > 1$ имеем неравенство

$$\left\|A_k(x_k - x^*)\right\|^2 \le \left\|x_0 - x^*\right\|^2 - \sum_{i=0}^{k-1}\left(\frac{m(f(x_i) - f_{min})}{\left\|B_i^T \partial f(x_i)\right\|}\right)^2 = r_0^2 - \sum_{i=0}^{k-1} h_i^2 = r_k^2,$$

с помощью которого обеспечивается достаточное условие отсутствия точки x_ε^* при $f_{min} < f^* - \varepsilon$ (реализовано в п. 1 метода amsg2p).

Антиовражная техника в методе amsg2p направлена на уменьшение степени овражности поверхностей уровня выпуклых функций подобно тому, как это сделано в r-алгоритмах [8]. Детерминант матрицы B_k уменьшается, так как, если на k-м шаге реализуется преобразование пространства, то $\det(B_{k+1}) = \det(B_k)\det(I + \eta\xi_{k+1}^T) = \det(B_k)\sqrt{1-\mu_k^2}$. Для

овражных функций это обеспечивает ускоренную сходимость метода amsg2p при произвольной начальной стартовой точке x_0 и достаточно малых значениях параметра ε (ε: $10^{-10} - 10^{-14}$).

В табл. 4 приведены результаты вычислительных экспериментов для квадратичных функций от $n = 200$ переменных с различной степенью овражности (рассматривалась функция $f(x) = \sum_{i=1}^{n} q^{i-1} x_i^2$, а степенью овражности считалась величина $Q = q^{n-1}$).

При расчетах использовался параметр $m = 2$, $f_{min} = f^* = 0, x_0 = (1,\ldots,1)^T$. Для ряда стремящихся к нулю ε в таблице даны затраты в числе итераций, которые требуются при степенях овражности $Q1 = 10$, $Q2 = 100$, $Q3 = 1000$, $Q4 = 10^6$, $Q5 = 10^9$.

Таблица 4

Сходимость amsg2p для квадратичных функций, $n = 200$

eps	itn(Q1)	itn(Q2)	itn(Q3)	itn(Q4)	itn(Q5)
1.00E-003	11	36	84	361	773
1.00E-005	15	46	99	405	826
1.00E-007	18	56	113	430	868
1.00E-009	22	65	128	461	916
1.00E-011	25	73	142	493	947
1.00E-013	29	81	154	517	979
1.00E-015	32	89	167	541	1006
1.00E-017	35	96	180	560	1027
1.00E-019	39	102	189	574	1042
1.00E-020	41	105	196	585	1048

Метод amsg2p можно использовать для того, чтобы достаточно точно найти приближение к единственной точке минимума существенно овражных функций. Проиллюстрируем это на примере известной тестовой задачи maxquad [14], которая связана с минимизацией существенно овражной выпуклой кусочно-квадратичной функции $f(x) = \max\limits_{1 \le k \le 5} \phi_k(x)$, $x \in R^{10}$. Здесь $\phi_k(x) = x^T H_k x - b_k^T x$, H_k – симметрические 10×10-матрицы, такие что

$H_{kij} = e^{i/j}\cos(ij)\sin k$, если $i < j$, и $H_{kii} = i|\sin k|/10 + \sum_{j \neq i}|H_{kij}|$, а компоненты векторов b_k определяются $b_{ki} = e^{i/k}\sin(ik)$.

Затраты метода amsg2p для нахождения в задаче maxquad единственного решения с достаточно высокой точностью (до 14-ти значащих цифр) позволяет оценить приведенный ниже фрагмент численных расчетов с одноименной программой amsg2p на языке octave [15].

```
Maxquad: f(x0)   5.3370664293114e+003
     fmin = -8.4140833459641e-001
..epsf..  .....f(itn).........  .itn.
1.0e-001  -7.7355266120112e-001    17
1.0e-003  -8.4084776169123e-001    29
1.0e-004  -8.4132394277880e-001    35
1.0e-005  -8.4140078034524e-001    41
1.0e-006  -8.4140807664455e-001    49
1.0e-011  -8.4140833458913e-001    94
1.0e-012  -8.4140833459555e-001   101
1.0e-013  -8.4140833459633e-001   110
1.0e-014  -8.4140833459640e-001   116
1.0e-015  -8.4140833459641e-001   122
```

Здесь $m = 1$, $x_0 = (1,\ldots,1)^T$ и $\varepsilon = epsf$.

Вычислительные эксперименты

Теорема 4 обеспечивает обоснование сходимости метода amsg2p аналогично тому как теорема 3 – сходимость метода (11), (12). Но более сильное уменьшение обьема эллипсоида, локализующего точку x^*, обеспечивает для овражных функций его ускоренную сходимость по сравнению с методом (11), (12). По количеству итераций метод amsg2p сравним с r-алгоритмом, а в ряде случаев и превосходит его. Это подтверждают результаты тестовых экспериментов из [11], которые приведены на рис. 6. Рассматривались 7 известных тестовых задач безусловной мимимизации гладких и негладких выпуклых функций [16]

(стр. 279–282) и кусочно-квадратичная функция [8] (стр. 176). Количество переменных в тестовых примерах было от 5 до 50. Все примеры решались r-алгоритмом с критериями останова $\varepsilon_x = 10^{-6}$ и $\varepsilon_g = 10^{-6}$ и методом amsg2p при достаточно малых $\varepsilon_f = 10^{-8}$ и $\varepsilon_f = 10^{-12}$. Число затраченных методами итераций дано на рис. 4. Из него видим, что только в одном случае метод amsg2p уступил r-алгоритму (пример № 2).

На рис. 7 приведены количества итераций метода amsg2p и программной реализации r-алгоритма, которую выполнил Д.Л. Крошко, для пяти тестовых примеров из [16] (их номера находятся под столбцами диаграммы). Из диаграммы видим, что для четырех тестовых примеров из пяти выиграш amsg2p по количеству итераций составляет более чем в два раза.

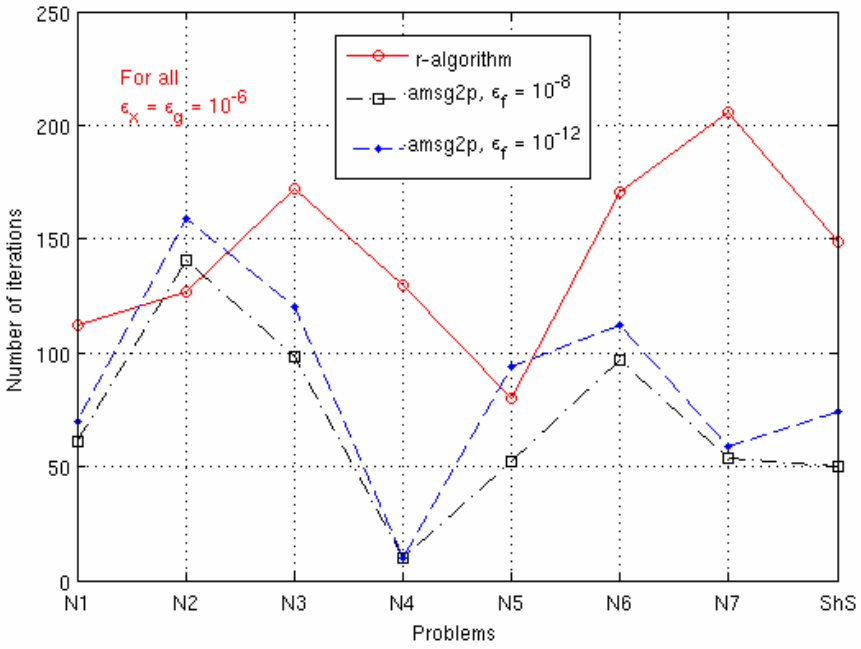


Рис. 6. Сравнение amsg2p и r–алгоритма для восьми функций [11]

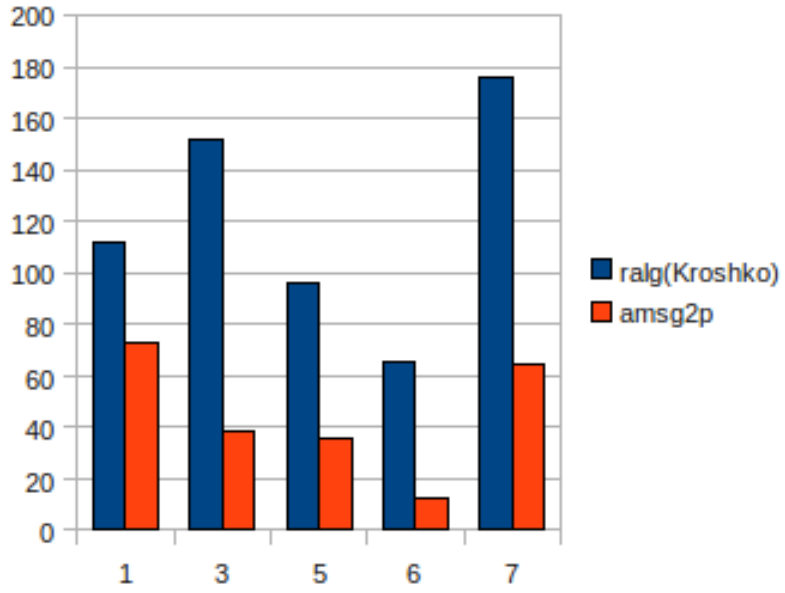


Рис. 7. Сравнение amsg2p и r –алгоритма в реализации Д.Л. Крошко для пяти тестовых примеров из монографии [16]

Вычислительную эффективность метода (11), (12) и метода amsg2p по отношению к методу Поляка приведем на примере кусочно-линейной и квадратичной функций от 20 переменных из работы [17]. Основным показателем эффективности методов будем считать не время вычислений, а число итераций, т.е. число вычислений $f(x)$ и $\partial f(x)$. Рассматривались кусочно-линейная функция $f(x) = \sum_{i=1}^{20}(1+1/4)^{i-1} \mid x_i \mid$ (обозначена *Sabs* (1.25)) и квадратичная функция $f(x) = \sum_{i=1}^{20}(1+1/2)^{i-1} x_i^2$ (обозначена *Squad* (1.5)). Для функции *Sabs*(1.25) использован параметр $m = 1$, а для функции *Squad* (1.5) – параметр $m = 2$. Количество итераций для всех трех методов (polyak, amsg2 и amsg2p) при разных значениях ε приведены в табл. 5, где "прочерк" означает, что метод не решил задачу за 10000 итераций.

Из табл. 5 видим, что методы (11), (12) и amsg2p намного эффективнее, чем метод Поляка без преобразования пространства. Число итераций для них увеличивается слабо при значительном уменьшении ε. Это подтверждает, что преобразования пространства переменных, направленные на уменьшение степени вытянутости поверхностей уровня выпуклых функций, способны для

овражных функций значительно ускорить сходимость субградиентных методов с шагом Поляка.

Таблица 5

Вычислительные эксперименты из [17]

ε	sabs(1.25), n= 20, m=1			squad(1.5), n = 20, m=2		
	polyak	**amsg2**	**amsg2p**	**polyak**	**amsg2**	**amsg2p**
10^{1}	289	48	20	601	32	15
10^{0}	1163	84	31	1205	36	19
10^{-1}	2737	102	42	2047	46	20
10^{-2}	4306	108	48	3079	51	23
10^{-3}	5869	113	55	4237	56	25
10^{-4}	7425	119	68	5463	58	25
10^{-5}	8943	161	78	6719	61	25
10^{-6}	–	197	95	7985	65	26
10^{-8}	–	214	107	–	68	27
10^{-10}	–	228	119	–	71	32
$F(x_0)$	5567.1151			2194649.4419		

Заключение

Несмотря на то, что *r*-алгоритмы используются уже 40 лет, проблема обоснования их сходимости для всего класса выпуклых функций остается открытой и в настоящее время. Еще в 1982 г. Н.З. Шор и В.И. Гершович в работе [18] отметили: «Теория всего класса алгоритмов с растяжением пространства далека от совершенства. Нам кажется достаточно реалистичной целью – построение такого алгоритма, который по своей практической эффективности не уступал бы *r*-алгоритму и был столь же хорошо обоснован, как метод эллипсоидов». Шагом в этом направлении можно считать алгоритм (11), (12), где для преобразования специального эллипсоида в шар используется антиовражный прием, близкий к тому, который имеет место в *r* –алгоритмах. Однако, здесь растяжение пространства реализуется в направлении разности двух нормированных субградиентов, и близким к направлению разности двух субградиентов оно будет только тогда, когда нормы субградиентов близки.

В заключение отметим, что ускоренные варианты субградиентных методов на основе однорангового эллипсоидального оператора можно построить и для других способов регулировки шага. Замечательной чертой

таких методов есть полная определенность в параметрах преобразования пространства.

Список литературы

[1] Поляк Б.Т. Введение в оптимизацию. – М.: Наука, 1983. – 384 с.

[2] Поляк Б.Т. Минимизация негладких функционалов // Журн. вычислит. математики и матем. физики. – 1969. – Т.9. – № 3. – С. 507 – 521.

[3] Васин В.В., Еремин И.И. Операторы и итерационные процессы фейеровского типа (теория и приложения). – Москва; Ижевск: Институт компьютерных исследований, НИЦ «Регулярная и хаотическая динамика», 2005. – 200 с.

[4] Agmon S. The relaxation method for linear inequalities // Canadien Journal of Mathematics, 1954. – №6. – Pp. 382 – 392.

[5] Motzkin T., Schoenberg I.J. The relaxation method for linear inequalities // Canadien Journal of Mathematics, 1954. – №6. – Pp. 393 – 404.

[6] Kelley J. E.. The cutting plane method for solving convex programs. J. Soc. for Industr. and Appl. Math., 1960. – Vol. 8. – № 4. – Pp. 703 – 712.

[7] Camerini P., Fratta L., Maffioli F. On improving relaxation methods by modified gradient techniques // Math. Program., 1975. – Study 3. – P. 26 – 34.

[8] Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.

[9] Шор Н.З., Журбенко Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика, 1971. – № 3. – С. 51 – 59.

[10] Стецюк П.И. Ортогонализующие линейные операторы в выпуклом программировании (Часть I) // Кибернетика и системный анализ, 1997. – № 3. – С. 97 – 119.

[11] Стецюк П. И. Субградиентные методы переменной метрики, использующие шаг Агмона-Моцкина и одноранговый эллипсоидальный опера-

тор // Труды АТИК 2007–2008. – Кишинэу: Эврика, 2009. – Т. I (XII). – С. 16 – 25.

[12] Стецюк П.И. Ускоренные по Шору модификации субградиентного метода Поляка // Математическое моделирование, оптимизация и информационные технологии: материалы 3-й Междунар. науч. конф. (Кишинэу, 19-23 марта 2012 г.): Кишинэу: Эврика, 2012. – С. 509 – 519.

[13] Стецюк П.И. Релаксационный субградиентный метод минимизации овражных выпуклых функций // Проблемы теоретической кибернетики. Материалы XVI Международной конференции (Нижний Новгород, 20-25 июня 2011 г.) / Под ред. Ю.И.Журавлева. – Нижний Новгород: Изд-во Нижегородского госуниверситета, 2011. – С. 449 – 453.

[14] Lemarechal C. Numerical experiments in nonsmooth optimization // In: Progress in nondifferentiable optimization /Ed. E. A. Nurminski. CP-82-58. – Laxenburg: International Institute for Applied System Analysis, 1982. – P. 61 – 68.

[15] Octave [Электронный ресурс]: http://www.octave.org/ – Режим доступа: свободный.

[16] Ржевский С.В. Монотонные методы выпуклого программирования. – Киев: Наукова думка, 1993. – 324 с.

[17] Журбенко М.Г., Стецюк П.І. Субградієнтні методи змінної метрики для розв'язування яружних задач оптимізації. – Київ, 2009. – 27 с. – (Препр. / НАН України. Ін-т кібернетики імені В.М.Глушкова; 2009-3)

[18] Гершович В.И., Шор Н.З. Метод эллипсоидов, его обобщения и приложения // Кибернетика, 1982. – № 5. – С. 61 – 69.

Раздел II

Математическое моделирование

УДК 330.150

МОДЕЛИ НЕСОВЕРШЕННОЙ КОНКУРЕНЦИИ ПРИМЕНИТЕЛЬНО К АНАЛИЗУ ЭЛЕКТРОЭНЕРГЕТИЧЕСКОГО РЫНКА СИБИРИ

Айзенберг Н.И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН, Иркутск

E-mail: zen@isem.sei.irk.ru

Киселёва М.А.

Институт систем энергетики им. Л.А. Мелентьева СО РАН, Иркутск

E-mail: marinee@mail.ru

Аннотация. Рассматриваются механизмы организации аукционов на электроэнергетическом рынке, основанные на подаче заявок оператору рынка со стороны потребителей и производителей энергии. Обсуждаются и сравниваются возможные стратегии поведения генераторов, приводящие к различным равновесным ситуациям и соответствующие моделям Курно, равновесия функций предложения, совершенной конкуренции. Механизмы тестируются на электроэнергетической системе зоны Сибирь.

Ключевые слова. Электроэнергетический рынок, модели несовершенных рынков, олигополия, модель равновесия функций предложения, либерализация.

Введение

Успешное функционирование электроэнергетической отрасли предполагает координирование деятельности генерирующих компаний с целью оптимизации функционирования системы как в техническом, так и экономическом смысле. В 2011 году закончился второй этап реформы электроэнергетики, в результате которого мы перешли к полностью либерализованному рынку, организованному на условиях, совершенно отличных от прежних. Вместо целостной вертикально-интегрированной естественной монополии электроэнергетика теперь представляет собой четыре независимых процесса: производства, транспортировки, распределения и сбыта электроэнергии. Законодательно на всех уровнях, кроме транспортировки, введена конкуренция. Предполагается, что цены формируются посредством рыночной конкуренции, и их уровень должен способствовать как достаточному удовлетворению потребителей, так и эффективному развитию отрасли.

На сегодняшний день в мире известно несколько способов организации функционирования свободного рынка в электроэнергетике, в том числе торговли электроэнергией, мощностью, системными услугами. В сфере оптовой торговли электроэнергией можно выделить механизмы в виде создания организованных рынков, где происходит торговля в реальном времени и на сутки вперёд. Возможна структура с заключением двусторонних краткосрочных и долгосрочных договоров между поставщиком и потребителем, которые могут быть как регулируемыми, так и не регулируемыми государством.

Процессы либерализации, происходящие в мире, имеют непродолжительную историю. Свободные рынки электроэнергии находятся в разной стадии развития и становления. Некоторые из них показали эффективность при внедрении в определённых странах, некоторые - нет. В целом результаты либерализации нельзя оценить однозначно, в том числе из-за краткого периода существования таких рынков. Затруднительно определить наиболее эффективный механизм организации электроэнергетической отрасли. Необходим глубокий всесторонний анализ как форм организации, так и специфических условий, к которым они будут применены.

В ходе бурных процессов реструктуризации электроэнергетики в России проект новых рынков задумывался и внедрялся без всестороннего тестирования. На это были и объективные причины, в том числе необходимость принятия оперативных мер по реформированию из-за кризисного состояния отрасли. При определении и выборе реализованного на сегодняшний момент механизма фактически раздавались заверения в обеспечении "правильных стимулов" в поведении участников и, соответственно, "эффективных результатов" работы рыночной модели в целом. Всё это подкреплялось ссылками на зарубежный опыт успешного внедрения того или иного механизма, либо общим анализом, который мог быть проведён только для некоторых "идеальных" условий функционирования рынка. Глубокого и всестороннего анализа для специфических условий России, по сути, произведено не было.

Ранее в России управление электроэнергетикой осуществлялось, исходя из критериев оптимальной загрузки генерирующей мощности и электрических сетей при фиксированном спросе на электроэнергию. При этом учитывались и другие задачи в управлении функционированием электроэнергетических систем – обеспечение должного (нормативно заданного) уровня надёжности электроснабжения отдельных категорий потребителей, обеспечение необходимого качества электроэнергии, обязательства по покрытию тепловых нагрузок ТЭЦ и др. На сегодняшний момент в условиях либерализации ориентиры меняются. Возникает необходимость вводить другие критерии оптимального функционирования отрасли. На наш взгляд, важнейшими из них являются следующие: максимизация общественного благосостояния; минимизация цен на энергию и мощность; обеспечение устойчивости организации рынка (отсутствие стимулов к смене правил поведения, обеспечение в течение длительного времени стабильного уровня цен). Здесь мы намеренно не говорим о критериях, обеспечивающих надёжность энергоснабжения и создающих условия для успешного развития отрасли, реализации её инфраструктурных задач. Это относятся скорее к задаче регулирования рынка, а не к обеспечению эффективного механизма взаимодействия производителей на рынке электроэнергии.

Оценивать по перечисленным критериям возможные механизмы функционирования отраслевого рынка сложно. Необходимы модели, которые были бы пригодны для анализа качеств правил организации взаимодействий между участниками рынка с точки зрения выдвинутых критериев, в том числе правил торгов, сообразуясь с которыми, экономические агенты будут вести себя тем или иным образом. Речь идёт о математических моделях, описывающих действия генерирующих компаний с одной стороны и потребителей электроэнергии с другой в условиях либерализованного рынка с заданными ограничениями на действия агентов.

Будут рассмотрены несколько моделей. Их различие можно определить по типу подаваемых поставщиком заявок-функций предложения, которые

могут формироваться в зависимости от рыночной цены, эластичности спроса, объёмов выпуска конкурентов, действий конкурентов при изменении объёмов поставок и т.д.

Среди наиболее распространенных в литературе подходов для анализа и прогнозирования ситуации на рынках электроэнергии является применение микроэкономических моделей типа Курно, равновесия функций предложения (SFE-Supply Function Equilibrium, а, точнее, её линейного варианта) и равновесия предполагаемых функций предложения (CSFE-Conjectured Supply Function Equilibrium). Во всех этих подходах учитываются ограничения на мощность.

Модель Курно является одной из наиболее распространенных моделей для анализа функционирования рынков несовершенной конкуренции. Введение в модель Курно ограничений на мощность и адаптация ее к применению на спотовом рынке электроэнергии не представляет значительной сложности, ее можно найти, например, в работах [18, 19]. Там же рассматривается закрытый аукцион, где производители конкурируют своими заявками-функциями предложения, в том числе ступенчатыми. В этом случае может быть реализовано три возможных типа равновесия: без рационирования, с рационированием (когда кривая спроса пересекает функцию суммарного предложения в скачке) и с барьером. В отличие от равновесий с рационированием и с барьером, для равновесия без рационирования существуют условия устойчивости, при этом объемы производства будут соответствовать локальному равновесию Курно.

Модели равновесия функций – предложения сравнительно недавнее изобретение в экономической науке. Впервые они были введены в работе [15], а применительно к рынкам электроэнергии (с учетом ограничений на мощность) рассматривались в работах [7]. Линейный вариант равновесия функций предложения (LSFE) и его применение к анализу рынка Великобритании и Уэльса разрабатывался в ряде зарубежных работ [8].

В российской литературе моделирование отечественного оптового рынка электроэнергии на основе подхода, использующего линейные функции предложения, было проведено в работе [1], в которой авторы рассматривают электроэнергетический рынок Центральной России (Средневолжский район). Целью работы являлась оценка рыночной власти отдельных генерирующих компаний, а также оценка последствий слияния в единый действующий на рынке субъект определённых компаний, введения новых мощностей. Исследования цен в работе проводятся для совокупного предъявляемого спроса, но намечены подходы учёта специфики потребления различных групп, в частности населения и промышленности.

Описывать взаимодействие на электроэнергетическом рынке можно через модель предполагаемых функций предложения (CSFE-подход). В этом подходе каждый агент формирует свою стратегию, исходя из предположений о действиях своих конкурентов. Важно, что информация может быть недостоверной, и это существенно приближает модель к реальности. Активное развитие данный подход получил в середине 90-х годов [11, 12, 16]. С учётом сетевых ограничений подход рассмотрен в [12]. Одним из частных случаев модели предполагаемых функций предложения является конкуренция Курно, который применим при неэластичном спросе (что характерно для спотового рынка электроэнергии). Все вышеперечисленные подходы моделируют одноуровневое взаимодействие фирм, где устойчивым состоянием рынка является равновесие Нэша.

Ряд зарубежных работ посвящен поиску равновесия среди производителей спотового рынка с учетом ограничений на передачу при наличии узлового ценообразования, т.е. определении цены в каждом узле как множителя Лагранжа к балансовому ограничению для соответствующего узла [9, 13, 14]. В этом случае приходится иметь дело с так называемыми MPEC-задачами, т.е. задачами математического программирования с равновесными ограничениями. Модель такого взаимодействия фирм на рынке формулируется задачей двухуровневого программирования, где на нижнем уровне системным

оператором решается задача, определяющая цены, исходя из оптимизации режимов в узлах системы по критерию максимизации функции общественного благосостояния на основе функций предложения, объявленных поставщиками, и функции спроса, предоставленной потребителем (LMP – Location Marginal Pricing). А на верхнем уровне каждый производитель решает задачу максимизации прибыли и формирования параметров собственной функции предложения, исходя из знаний о способе ценообразования на рынке.

В отечественной литературе работ, посвящённых особенностям функционирования современных рынков электроэнергии, представлено немного. Российская электроэнергетика обладает рядом специфических черт, связанных с суровыми климатическими условиями, традициями развития отрасли. Подробно это рассмотрено в [2], где выделено несколько возможных способов организации торговли электроэнергией на рынке и признано, что механизм полной либерализации цен не является наилучшим для России.

В работе [4] исследуется перспективное развитие генерирующих мощностей на долгосрочную перспективу. Для этого рассматривается взаимодействие на рынке электроэнергии России, при моделировании которого используется модель Курно.

Моделирование взаимодействия на рынке электроэнергии является очень сложной задачей, частью которой можно считать реализацию самого механизма оптовой торговли на рынке «на сутки вперёд». Такой механизм формирования свободных цен описан в [3]. Эта модель является действующим техническим инструментом, которым пользуется коммерческий оператор для организации оптовой торговли электроэнергией в России по правилам, закреплённым законодательно. Исходя из сетевых ограничений, заявок, поданных потребителем и производителем, и на основе критерия максимизации общественного благосостояния, определяются оптимальные узловые объёмы и цены с помощью механизма, подобного LMP, т.е. через двойственные переменные оптимизационной задачи. Вопросы формирования заявок производителями электроэнергии в данной постановке не рассматриваются.

В нашем исследовании мы поставили перед собой задачу моделирования функционирования рынка электроэнергии с возможностью анализа степени концентрации компаний в отрасли, определения наиболее влиятельных игроков на рынке, выяснения устойчивости стратегий, выбираемых производителями. В первой части проводится анализ моделей, описывающих взаимодействия на рынке электроэнергии: это модели несовершенной конкуренции для рынка однородного товара без сетевых ограничений.

Во второй части рассмотренные модели расчёта равновесия проигрываются на рынке электроэнергии Сибири в случае реализации различных сценариев: эластичного и неэластичного спроса, маловодного года и др.

Описание моделей

Рассмотрим модели одноуровневого взаимодействия генерирующих компаний в предположении, что сети достаточно развиты и препятствия для передачи любого требуемого объёма электроэнергии отсутствуют, при этом будем учитывать ограничения на величину генерации. Речь идёт об определении цены на спотовом рынке, организованном в виде двойного аукциона, где стратегически взаимодействуют фирмы, генерирующие электроэнергию (электростанции), а цена определяется из равенства совокупного спроса и предложения. Электростанции имеют различия в технологиях и в переменных издержках, поэтому для описания рыночного поведения необходимо использовать модели, разграничивающие фирмы на стратегических производителей, которые могут влиять на рыночную цену, и фирмы конкурентного окружения, которые не участвуют в торгах, а принимают цену как заданную. Стоит отметить, что Закон РФ об электроэнергетике допускает к участию в аукционе участников таких двух типов [6].

Стратегии поведения генераторов электроэнергии – это их функции предложения, которые они предоставляют оператору рынка. Вопрос стоит в том, насколько выгодно производителю отклониться от своих истинных издержек, участвуя в торговле, и как это повлияет на исход торгов.

Нами были рассмотрены несколько стратегий, следование которым может приводить к различным рыночным исходам: соответствующим модели Курно, ценового лидерства и конкурентного окружения, а также варианты модели равновесия функций предложения (LSFE без конкурентного окружения и с наличием последнего, эта модель является расширением модели SFE).

Все изученные модели можно объединить в класс моделей олигополии, где рассматривается зависимость выпуска конкурентных фирм от цены, установившейся на рынке, с учётом того, что на эту цену будут влиять объёмы производства каждого генератора. Функции предложения называются предполагаемыми, так как фирмы могут только догадываться о реакции своих конкурентов. В них присутствуют так называемые коэффициенты влияния каждого участника на ситуацию в целом.

Обозначим через $Q(P)$ функцию отраслевого выпуска, которая будет складываться из функций предложения отдельных фирм. $q_i(P)$ – выпуск фирмы i, $i = \overline{1,n}$, $n \geq 2$ – количество фирм на рынке и

$$\sum_{i=1}^{n} q_i(P) = Q(P).$$

Объёмы конкурентов для фирмы i определим как $q_{-i}(P)$, это общий выпуск за исключением i. Соответственно, остаточный спрос генерирующей компании i:

$$q_i(P) = Q(P) - q_{-i}(P).$$

Здесь $P \in R_+^1$ – цена, формируемая в результате взаимодействия агентов на рынке при условии, что все потребители агрегируются единой невозрастающей функцией спроса $D(P)$ или обратной к ней $D^{-1}(Q)$. Функция издержек $C_i(q_i)$ выпуклая, возрастающая, $q_i \geq 0$, $i = \overline{1,n}$. Генерирующие компании имеют своей целью максимизацию прибыли на остаточном спросе, при условии, что в равновесии спрос будет равен общему выпуску компаний

$$\pi(P, q_i) = D^{-1}(Q(P)) \cdot q_i(P) - C_i(q_i). \quad (1)$$

Функция прибыли фирмы i вогнута по P, а, следовательно, имеет единственный максимум. Запишем условие первого порядка максимизации прибыли:

$$\frac{\partial D^{-1}(Q(P))}{\partial Q(P)} \cdot w_i(P) \cdot q_i(P) + D^{-1}(Q(P)) = \frac{\partial C_i(q_i(P))}{\partial q_i(P)}, \quad (2)$$

где $w_i(P)$ – индексы влияния фирмы i (CV) на состояние рынка. Они имеют тот же смысл, что и влияние изменения выпуска фирмы на выпуск отрасли в целом

$$w_i(P) = \frac{\partial Q(P)}{\partial q_i(P)},$$

определяют возможные реакции конкурентов на изменение выпуска фирмой i. Функция предложения для генерирующей компании i:

$$q_i(P) = \frac{\frac{\partial C_i(q_i(P))}{\partial q_i(P)} - D^{-1}(Q(P))}{\frac{\partial D^{-1}(Q(P))}{\partial Q(P)} w_i(P)}.$$

Каждая фирма определяет функции предложения других фирм и использует эту информацию при максимизации своей прибыли на остаточном спросе. Важно, что тип этих реакций предполагает сама фирма i. Соответственно, ответы могут отличаться от действительных реакций конкурентов. В этом видна взаимосвязь модели с постановкой задачи для конкуренции по Штакельбергу [12]. На практике величину индекса влияния возможно получить только из эконометрических оценок (при этом стоимость получения таких данных может быть достаточно высока) либо возможны оценки в результате включения этой переменной в повторяющиеся игры.

Упрощенным видом этой модели считается модель равновесия функций предложения, где информацией о конкурентах владеют все заинтересованные стороны. Тогда индексы влияния будут однозначно определяться в результате рыночных взаимодействий на рынке. Это предположение значи-

тельно упрощает моделирование. Вариация индексов влияния фирмы приводит к различным моделям равновесия: максимальное будет соответствовать модели Курно, минимальное – модели совершенной конкуренции [17].

Рассмотрим особенности формируемого равновесия на примере со следующими предпосылками: линейная функция спроса, гетерогенные генерирующие компании с квадратичными выпуклыми функциями эксплуатационных издержек и ограничениями на выработку энергии. В этих предположениях для некоторых моделей можно получить аналитическое решение и сравнить результаты.

Пусть

$$D(P) = N - \gamma P \tag{3}$$

– линейная функция совокупного спроса, где γ – положительная величина. Соответственно, обратная функция спроса будет

$$P(Q) = \frac{N - Q}{\gamma},$$

где весь спрос удовлетворяется выпуском n фирм

$$\sum_{i=1}^{n} q_i(P) = D(P).$$

Функция издержек имеет вид:

$$C_i(q_i) = \frac{1}{2} c_i q_i^2 + a_i q_i, \tag{4}$$

$c_i > 0,\ a_i \neq 0,\ i = \overline{1,n}$ – издержки фирмы i. Это квадратичные строго выпуклые функции. Для каждой генерирующей компании i определена максимально вырабатываемая мощность V_i.

Все изложенные ниже модели объединены одной идеей. В них производители определяют стратегию поведения в виде функции предложения, в которую включены возможные реакции конкурентов на изменение объёмов выпуска фирмы.

Для выбранных функций спроса (3) и издержек (4) задачу (1) можно записать следующим образом:

$$\pi_i(P) = P(Q)\cdot q_i(P) - a_i q_i(P) - \frac{1}{2} c_i \; q_i^2(P) \to \max_P,\; i = \overline{1,n}.$$

Условие первого порядка имеет вид

$$\left(\frac{\partial P(Q)}{\partial Q} \cdot \frac{\partial Q}{\partial q_i} - c_i \right) q_i(P) = a_i - P, \;\; i = \overline{1,n},$$

тогда в равновесии, зная, что спрос равен предложению, получим

$$q_i(P) = \frac{P - a_i}{c_i + \dfrac{w_i}{\gamma}}. \qquad (5)$$

Это функция предложения компании без учёта ограничений на выработку, где w_i – индекс влияния i-ой фирмы на выпуск отрасли в целом. В зависимости от того, какие значения будет принимать этот индекс, или насколько существенным окажется влияние одной фирмы на общую функцию предложения, мы получим разные модели функционирования рынка.

Количественная конкуренция (модель Курно)

Стратегией каждого производителя в рамках модели Курно является выбор своего объема производства при цене, складывающейся на рынке олигопольного взаимодействия. Каждая фирма максимизирует прибыль на остаточном спросе. При этом в классической постановке общая функция предложения отрасли является аддитивно сепарабельной относительно объёмов выпуска отдельных компаний, или, что тоже самое, индекс влияния i фирмы $w_i = 1$. При ограничениях на мощность, т.е. в предположении, что объем выпуска генерирующей компании i ограничен величиной V_i, имеем

$$q_i^K(P) = \begin{cases} 0, \;\; P \le a_i, \\ \dfrac{\gamma(P - a_i)}{1 + c_i\gamma}, \; a_i < P \le \dfrac{V_i(1 + c_i\gamma)}{\gamma} + a_i, \\ V_i, \; P > \dfrac{V_i(1 + c_i\gamma)}{\gamma} + a_i, \end{cases} i = \overline{1,n}. \qquad (6)$$

В данном случае можно сказать, что при подаче своей заявки генератор будет ориентироваться на цену и наклон функции спроса в возможной точке

равновесия и на свои издержки. Более сложные взаимосвязи приниматься во внимание не будут.

Равновесная цена p^* определяется аукционистом путем приравнивания спроса и совокупного предложения в каждый рассматриваемый момент времени (например, час, сутки, год и т.д.)

$$\sum_{i=1}^{n} q_i(P) = D(P),$$

откуда

$$P^{K*} = \frac{N + \gamma \sum\limits_i \frac{a_i}{1+c_i\gamma}}{\gamma(1 + \sum\limits_i \frac{1}{1+c_i\gamma})}.$$

Модель олигополии с ценовым лидерством

В рамках данной модели предполагается, что несколько стратегических фирм (назовем их лидерами) конкурируют между собой по ценам, а фирмы из конкурентного окружения выбирают объемы выпуска, считая цену заданной (например, фирмы с электростанциями, предназначенными для покрытия базовой нагрузки).

Каждая фирма конкурентного окружения $k = 1,\ldots,m$ решает задачу максимизации собственной прибыли при заданной цене, т.е.

$$\pi_k(q_k) = P(Q)q_k - C_k(q_k) \to \max_{q_k \geq 0}.$$

Тогда её функция предложения имеет вид

$$q_k(P) = \frac{P - a_k}{c_k}.$$

Соответственно, для стратегических фирм остаточный спрос запишется как

$$q_i(P) = D(P) - \sum_{k=1}^{m} q_k(P) - \sum_{j=1, j\neq i}^{n} q_j(P).$$

Тогда при предположениях этой модели $w_i = 1$, а

$$\frac{\partial P}{\partial Q} = -\frac{1}{\gamma + \sum_{k=1}^{m} \frac{1}{c_k}}.$$

Обозначим

$$\hat{\gamma} = \left(\gamma + \sum_{k=1}^{m} \frac{1}{c_k} \right).$$

Функция предложения стратегического игрока i, $i = \overline{1, n}$, примет вид

$$\hat{q}_i^K(p) = \begin{cases} 0, & P \le a_i, \\ \dfrac{\hat{\gamma}(P - a_i)}{1 + c_i \hat{\gamma}}, & a_i < P \le \dfrac{V_i(1 + c_i \hat{\gamma})}{\hat{\gamma}} + a_i, \\ V_i, & P > \dfrac{V_i(1 + c_i \hat{\gamma})}{\hat{\gamma}} + a_i. \end{cases} \quad (7)$$

Соответственно, равновесная цена равна

$$\hat{P}^{K*} = \frac{N + \hat{\gamma} \sum_i \frac{a_i}{1 + c_i \hat{\gamma}} + \sum_k \frac{a_k}{c_k}}{\hat{\gamma} + \hat{\gamma} \sum_i \frac{1}{1 + c_i \hat{\gamma}} + \sum_k \frac{1}{c_k}}.$$

Теоретическая функция предложения фирмы-участницы аукциона помимо собственных издержек в этом случае зависит от наклона кривой спроса и параметров функций издержек конкурентного окружения. Кроме того, анализируя функции предложения фирм в модели без конкурентного окружения и с его наличием, можно видеть, что объёмы, предлагаемые на рынок фирмой в соответствии со стратегией (7), будут меньше, чем в (6). И в модели с конкурентным окружением на рынке сформируется равновесие с большим объёмом удовлетворения потребителя и более низкими ценами, чем в модели без ценопринимающих фирм.

Модель равновесия линейных функций предложения (модель LSFE)

Как и в модели ценового лидерства, будем рассматривать взаимодействующие на рынке стратегические фирмы и конкурентное окружение. Здесь надо оговориться, что описанная ниже модель является частным случаем модели SFE с произвольными функциями предложения. Для такой модели существует проблема множественности равновесий и сложности их нахождения. Единственное решение достигается, однако, при использовании линей-

ного вида функций предложения фирм-конкурентов [8]. Для электроэнергетики целесообразно применение именно такой модели, так как обычно правила двусторонних аукционов определяют вид подаваемых заявок либо в виде ступенчатых, либо в виде линейных функций.

Предполагается, что правила рынка определяют функцию предложения для каждой фирмы в линейном виде [1]

$$q_i(P) = \beta_i(P - \alpha_i), \tag{8}$$

$\beta_i \geq 0,\ i = 1,\ldots,n$, где параметры α_i и β_i выбираются фирмой i. Цены таковы, что функции предложения не могут быть отрицательными. Рациональный объем выпуска каждой фирмы определяется из задачи максимизации прибыли на остаточном для фирмы i спросе. Остаточный спрос

$$q_i(P) = N - \gamma P - \sum_{j \neq i} \beta_j (P - \alpha_j).$$

Тогда условие первого порядка максимизации прибыли примет вид:

$$\beta_i(P - \alpha_i) = \left(\gamma + \sum_{j \neq i} \beta_j\right)\left((P - a_i) - c_i \beta_i (P - \alpha_i)\right),\ i = \overline{1,n}. \tag{9}$$

Равновесие достигается при $\alpha_i = a_i,\ i = \overline{1,n}$. В [7] доказано, что при этом условии существуют неотрицательные $\beta_i > 0,$ которые дают решение задачи и определяют равновесие в линейных функциях предложения (LSFE).

Для представления (5) параметр функции предложения

$$\beta_i = \frac{1}{c_i + \dfrac{w_i}{\gamma}}.$$

Тогда индексы влияния фирмы w_i находятся из решения системы, следующей из (9):

$$w_i = \frac{1}{\displaystyle\sum_{i \neq j} \frac{1}{w_j + c_j \gamma} + 1}.$$

При этом для достижения равновесия необходимо, чтобы в (8) $\alpha_i = a_i,\ i = 1,\ldots,n$. В [7] доказано, что при этом условии существуют неотри-

цательные $w_i > 0, i = 1,\ldots,n$, которые дают решение задачи и определяют равновесие в линейных функциях предложения (LSFE).

Функция предложения для генераторов с учётом ограничений на выработку энергии для всех i будет иметь вид

$$q_i^{LE}(P) = \begin{cases} 0, & P \le a_i, \\ \dfrac{\gamma(P - a_i)}{w_i + c_i\gamma}, & a_i < P \le \dfrac{V_i(w_i + c_i\gamma)}{\gamma} + a_i, \\ V_i, & P > \dfrac{V_i(w_i + c_i\gamma)}{\gamma} + a_i. \end{cases} \quad (10)$$

Для подсчета равновесных цен в течение заданного промежутка времени приравняем спрос и суммарное предложение всех фирм и получим равновесную цену в модели конкуренции линейных функций предложения:

$$P^{LE*} = \frac{N + \gamma\sum_i \frac{a_i}{w_i + c_i\gamma}}{\gamma(1 + \sum_i \frac{1}{w_i + c_i\gamma})}.$$

Условие равновесия $\alpha_i = a_i$ для каждой фирмы i означает, что есть стимул для участников аукциона сделать значение своих коэффициентов a_i (в функциях предельных издержек) общеизвестными.

Соответственно, данная стратегия будет предполагать, что генератор ориентируется на цену, эластичность спроса в этой точке и некоторую реакцию конкурентов на изменение его цены и объёма предложения.

Модель равновесия линейных функций предложения с конкурентным окружениям

Для случая, когда на рынке наряду со стратегическими фирмами действуют и фирмы конкурентного окружения, равновесные коэффициенты функций предложения $\hat{q}_i^{LE}$ вычисляются по формуле (9), где коэффициент γ меняется на $\hat{\gamma} = \left(\gamma + \sum_{k=1}^{m} \frac{1}{c_k}\right)$. Стоит отметить, что рассмотренная здесь модель равновесия линейных функций предложения верна для диапазона цен, превосходящих значение наибольшего среди взаимодействующих на рынке

стратегических фирм коэффициента a_i (обеспечивается неотрицательность объемов производства). В противном случае необходимо использовать уже кусочно-линейные аппроксимации [7].

Модель совершенно конкурентного взаимодействия генерирующих компаний (модель Вальраса)

В данном случае заявками всех взаимодействующих на рынке фирм будут их предельные издержки, а общая функция предложения будет определяться через прямую сумму объёмов всех генерирующих компаний. Тогда функция предложения каждой фирмы не будет зависеть от эластичности спроса и определится как

$$q_i^W(P) = \begin{cases} 0, & P \le a_i, \\ \dfrac{P - a_i}{c_i}, & a_i < P \le V_i c_i + a_i, \\ V_i, & P > V_i c_i + a_i, \end{cases} \quad i = 1, \ldots, n.$$

В равновесии получим цену Вальраса:

$$P^{W*} = \frac{N + \sum_i \frac{a_i}{c_i}}{\gamma + \sum_i \frac{1}{c_i}}.$$

Формально теоретически сравнить представленные выше модели можно в ситуации, когда цена находится в интервале $P \in \left(\max\limits_i a_i,\ \min\limits_i \left(V_i \left(\dfrac{w_i}{\gamma} + c_i \right) + a_i \right) \right)$, $i = \overline{1,n}$, т.е. когда на рынке задействованы все генерирующие мощности и при этом ни одна из них не выходит на свои ограничения. Это самый интересный случай для исследования, т.к. все агенты являются активными игроками. Для него в табл. 1 представлены выражения для функций предложения.

Зная, что $w_i \in [0,1]$, имеем следующие соотношения объёмов производства фирм в различных моделях друг с другом: $q_i^W \ge q_i^{LE} \ge q_i^K$. При этом цены в равновесии имеют обратное соотношение в силу того, что цена определяется из агрегированной функции спроса. Эти соотношения подтверждает

анализ эластичностей функций предложения фирмы и цены по w_i. Эластичность предложения $\varepsilon_{w_i}^{q_i} < 0$.

Таблица 1

Функции предложения генерирующих компаний в различных моделях взаимодействия в интервале цен, где все игроки являются активными

	Модель Вальраса	Модель LSFE	Модель Курно
Функции предложения	$q_i^W = \frac{P - a_i}{c_i}$	$q_i^{LE} = \frac{\gamma(P - a_i)}{w_i + c_i\gamma}$	$q_i^K = \frac{\gamma(P - a_i)}{1 + c_i\gamma}$

Эластичность цены по изменению (увеличению) индекса влияния фирмы на рынок $\varepsilon_{w_i}^P > 0$. При объективном увеличении рыночной власти у любой фирмы рынок будет уравновешиваться более высокой ценой, и компания может подавать заявки, существенно отличные от её предельных издержек. Максимально отличаться функции предложения от реальных издержек будут в случае, если фирмы будут играть по правилам, заложенным в модели Курно.

Если часть генерирующих компаний выводится из активных игроков и они предоставляют на рынок функции предложения в виде предельных издержек и становятся ценополучателями, то ситуация меняется. Цена равновесия существенно снижается, объёмы растут. Это происходит за счёт того, что для остаточного спроса, на котором играют экономически активные агенты, меняется эластичность. Спрос активнее реагирует на изменение цен – становится эластичнее $\hat{\gamma} > \gamma$ и

$$\frac{\gamma(P - a_i)}{w_i + c_i\gamma} < \frac{\hat{\gamma}(P - a_i)}{w_i + c_i\hat{\gamma}}.$$

В то же время соотношения между $q_i^W \geq q_i^{LE} \geq q_i^K$ остаются прежними, а цена и объёмы Вальраса не изменяются.

Таким образом, теоретический анализ микроэкономических моделей конкурентного взаимодействия экономических агентов на рынке электро-

энергии показал: а) наличие конкурентного окружения увеличивает объем выпуска продукции и снижает равновесную цену по сравнению с одноуровневым взаимодействием стратегических фирм; б) при конкуренции линейных функций предложения производителей электроэнергии рынок приходит к равновесию при меньших ценах и, соответственно, больших объёмах выпуска, чем при конкуренции Курно, что выгодно для потребителя. При этом каждая из фирм получает меньшую прибыль в сравнении с конкуренцией Курно, а потребительский излишек растёт:

$$SD = \frac{1}{2}(N - P)\sum_i \frac{\gamma(P - a_i)}{w_i + c_i\gamma}.$$

Так как эластичность потребительского излишка отрицательная относительно индекса влияния w_i $-\varepsilon_{w_i}^{SD} < 0$, то максимальным SD будет в модели Вальраса ($w_i = 0$), минимальным – в модели Курно ($w_i = 1$). В то же время, эластичность излишка производителя от w_i положительная (чем больше рыночная власть, тем больше у фирмы возможностей получить высокую прибыль), соответственно, он растёт при переходе от условий, когда генерирующие компании предоставляют свои функции предложения в виде q_i^W, к условиям, когда формируются стратегии q_i^K. Из-за этих двух взаимообратных эффектов однозначно оценить функцию общественного благосостояния при увеличении индекса влияния w_i не удаётся.

Для моделирования электроэнергетического рынка зоны Сибирь были выбраны следующие модели: Курно с конкурентным окружением и равновесия линейных функций предложения с конкурентным окружением. С помощью этих моделей, на наш взгляд, максимально близко можно описать ситуацию взаимодействия генерирующих компаний на рынке.

Моделирование взаимодействия на рынке электроэнергии Сибири

Сложность моделирования взаимодействия экономических агентов на рынке электроэнергии, функционирующего в Сибири, определяется несколькими факторами, среди которых можно выделить следующие.

- Большая доля электроэнергии (50-70%) производится на гидроэлектростанциях (Красноярская, Саяно-Шушенская, Братская, Усть-Илимская, Иркутская ГЭС) и, как следствие, энергобалансы Сибири подвержены нестабильным природным воздействиям, связанным с колебаниями стока рек.
- В Сибири (так сложилось исторически) функционируют в основном крупные генерирующие мощности. Кроме гидроэлектростанций можно выделить семь конденсационных станций мощностью более 1 000 МВт, которые разбросаны по большой территории и работают на местных углях. Масштаб производства, расстояние между потребителями определяют, в некоторой степени, их монопольное положение на прилегающих территориях. На базе этих станций достаточно сложно организовать здоровую конкуренцию.
- Потребление существенной части энергии крупными потребителями. Выпуск конкурентоспособной продукции некоторыми предприятиями возможен, в том числе благодаря тому, что электроэнергия в Сибири дешевле, чем в среднем по России.
- Имеют место повышенные требования к надёжности системы в связи с суровыми климатическими условиями. Важна тесная координация предприятий электроснабжения и коммунального хозяйства. Здесь появляются проблемы управления функционированием многопродуктовых производств, совмещающих тепло- и электроснабжение.
- Энергосистема Сибири работает практически изолированно, что связанно с плохой связью с другими энергозонами России.
- Большая протяжённость линий электропередач, определяемая низкой плотностью населения и очаговым характером развития экономики.

Учесть все эти особенности при моделировании рынка проблематично. С одной стороны, необходимы модели, согласовывающие интересы рыноч-

ных агентов (производителей электроэнергии и её потребителей) и лежащие в плоскости олигопольного взаимодействия, с другой стороны, учитывающие специфические сетевые ограничения, диктуемые естественно-монопольной средой. Мы ограничимся анализом стратегий экономических агентов и оценкой их рыночной силы без сетевых ограничений с помощью моделей, которые были описаны в части 1.

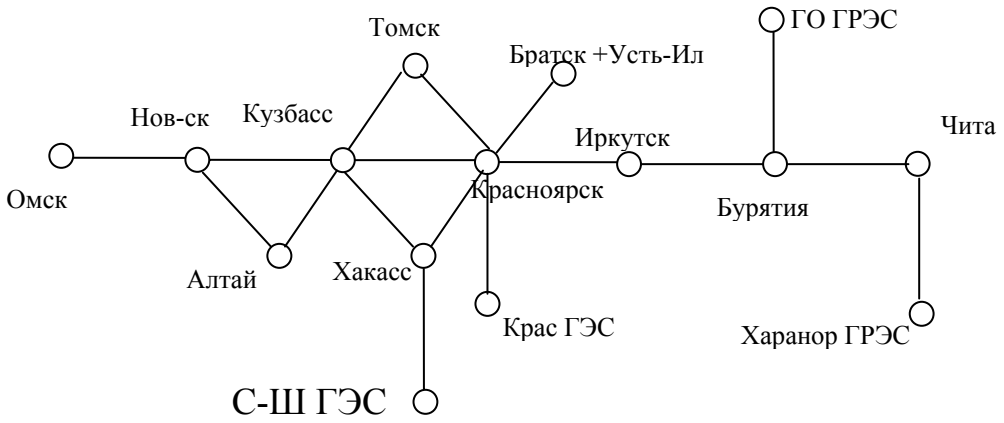


Рис. 1. Схема электроэнергетической системы «Сибирь»

На рис.1 представлена схема, состоящая из 14 узлов, для которой было смоделировано ценообразование с учётом стратегического взаимодействия генераторов на рынке. В табл. 2 и 3 представлены основные характеристики генерации и потребления (среднечасовое потребление и среднегодовые издержки генераторов). Для примера взят 2008 год.

Таблица 2

Спрос на электроэнергию (2008 г)

Узел	**Потребление, МВт**	**Узловые цены,** руб, МВт·ч
Иркутск	8268	247
Алтай	1629	275
Бурятия	795	281
Красноярск	6487	233
Новосибирск	2239	283
Омск	1586	339
Томск	1336	262
Чита	1087	319
Хакасия	2754	260
Кузбасс	5269	263

Определение функции спроса на электроэнергию в зоне Сибирь

Будем исходить из предположения, что спрос однороден для всех потребителей на всём географическом пространстве. Имея статистику потребления (цены и объёмы потребления) за 2005 и 2008 (табл. 2) годы в 14 узлах энергосистемы, мы допускаем, что имеем 28 точек наблюдаемого спроса. Агрегированный спрос определяется путём сложения спроса в отдельных точках. Например, если есть наблюдения, что объём 2000 МВт был потреблён по цене 156 руб. и объём 200 МВт по цене 300 рублей, то для оценки функции спроса имеем две точки: первая – объём потребления 2200 МВт по цене 156 руб. и объём 200 МВт по цене 300 руб. Далее полученную статистику о суммарном спросе оцениваем линейной функцией методом наименьших квадратов.

Существует несколько слабых мест такой оценки.

- Нет разделения потребителей на категории, такие как население, промышленные потребители, потребители сельскохозяйственной продукции. Эти группы существенным образом различаются по типу потребления. Для них не совпадают объёмы почасового потребления в течение суток, а также среднее потребление по месяцам в течение года. Мы предполагаем, что все эти потребители присутствуют на оптовом рынке, подавая заявки в виде своих функций спроса (объёмы и цены предполагаемого потребления промышленными предприятиями, сельскими потребителями и населением, которых представляют распределительные компании). Этот спрос суммируется и удовлетворяется в результате действия рыночного механизма.
- Предположение об однородности спроса в пространстве в случае Сибири далеко от реальности. Обеспеченность энергоресурсами по регионам Сибири различается существенным образом. Это диктует места размещения производств, а, следовательно, будет влиять и на формирование функции спроса. В данном случае можно предположить, что спрос населения достаточно однороден на рассматриваемом пространстве в связи со сходными климатическими условиями и развитием территорий.

Другим вариантом построения функции спроса может стать использование кривой загрузки мощностей в определённые моменты времени $N(t)$ и предположения о наклоне функции спроса γ (Стофт, 2006). Причём здесь возможно оценивать спрос для различных категорий и в различные временные интервалы. Следует отметить, что этот способ оценки тоже не является совершенным, так как использует экспертные оценки наклона прямой спроса, в то время как именно этот параметр является определяющим в формировании стратегии поведения генераторов на рынке.

Для данных табл. 2 параметр наклона функции спроса оценивается как $\gamma = 239{,}4$. А суммарная спрос выглядит как $D(P) = 77731 - 239{,}4p$.

Таблица 3

Характеристика часовой стоимости производства электроэнергии на ГРЭС, ТЭС и ГЭС (2008 г.)

Узел	**Часовая выработка, q, МВт**	a_i	c_i
Иркутск	4762	18	0.064
Гусиноозерская ГРЭС	2020	6.8	0.2
Харанорская ГРЭС	665	9.12	0.588
Красноярск	7906	21	0.042
Новосибирск	3214	12.4	0.09
Кузбасс	4297	20	0.08
С.-Ш.ГЭС	4781	-	-
Кр.ГЭС	3873	-	-
БГЭС-УИГЭС	9657	-	-

Производители (генераторы энергии) разделялись на стратегических (активно влияющих на цену) и ценополучателей. Во вторую группу входили гидроэлектростанции, имеющие по предположениям нулевые предельные затраты и участвующие на рынке только объёмами производимой энергии (Красноярская, Саяно-Шушенская, Братская, Усть-Илимская). Все станции имели ограничения на генерацию (столбец 2 в табл.3). Были рассчитаны цены по модели Курно и по модели равновесия линейных функций предложения (LSFE) с наличием конкурентного окружения. Полученные цены сравни-

вались с ценой монополии и ценой по Вальрасу (последняя формировалась путем приравнивания спроса и предложения, а свои функции предложения фирмы подавали в виде предельных издержек на единицу продукции).

Полученные характеристики цены, объемов генерации, прибыли отдельных генераторов приведены в табл. 4 и 5. В табл. 6 представлены сводные результаты расчётов по рассматриваемым моделям.

Таблица 4

Цена, установившаяся на рынке, и объемы производства из расчёта среднегодовых характеристик станций зоны Сибирь (на основе данных 2008 г.)

Цена (руб./МВт ч) и объёмы генерации (МВт)	Вид рынка (функции предложения)			
	Модель Вальраса	Модель LSFE	Модель Курно	Модель монополии
Цена (руб./МВт ч)	255	257	259	268
Иркутск	3707,81	3543,56	3534,91	3024,76
Гусиноозерская ГРЭС	1242,50	1228,08	1235,20	1023,92
Харанорская ГРЭС	418,50	419,64	421,96	344,32
Красноярск	5576,34	5430,79	5387,4	4537,74
Новосибирск	2696,19	2621,82	2618,47	2213,16
Кузбасс	2941,25	2825,41	2839,25	2394,81

Анализируя данные в таблицах, можно отметить, что большие объёмы производства будут при стратегиях, соответствующих модели LSFE. При этом, что закономерно, равновесие функций предложения даёт более низкие, чем при Курно, цены и прибыли отдельных генераторов. Суммарный излишек продавцов и потребителя будет возрастать при переходе от модели Курно к модели LSFE, но при этом доля прибыли в нём увеличивается и происходит перераспределение общего благосостояния в пользу производителей, относительно ухудшая положение потребителей. Этот пример показывает, что общепринятые в теории критерии не всегда хорошо работают на практике.

Таблица 5

Величина прибыли, которую получат генерирующие компании при реализации различных механизмов функционирования на электроэнергетическом рынке зоны Сибирь (на основе данных 2008 г.)

Величина прибыли генераторов (тыс. руб.)	Вид рынка (функции предложения)			
	Модель Вальраса	Модель LSFE	Модель Курно	Модель монополии
Иркутск	439,30	446,61	448,16	461,90
Гусиноозерская ГРЭС	154,38	156,97	157,58	162,09
Харанорская ГРЭС	51,17	52,40	52,61	54,09
Красноярск	653,26	664,39	666,52	685,96
Новосибирск	327,45	332,78	333,99	343,94
Кузбасс	346,03	351,51	353,00	363,31

Таблица 6

Основные характеристики равновесий при различных моделях функционирования рынка электроэнергии

	Цена (руб./ МВт ч)	Суммарный объём предложения (МВт)	Суммарная прибыль генераторов (тыс.руб.)	Выигрыш потребителя	Общественное благосос-тояние
Вальрас	255	16801,10	1991,35	577,8	2556,2
LSFE	257	16069,33	2009,69	540,1	2549,1
Курно	259	16037, 21	2020,51	520,7	2541,2
Монополия	268	13536,72	2071,13	383,7	2455,5

В табл. 7 приведены результаты расчётов по моделям равновесных цен для нескольких возможных сценариев. И та, и другая модель рассчитывались при наличии конкурентного окружения, куда входили гидроэлектростанции. Были рассмотрены модели с возможным недостатком и избытком генери-

рующих мощностей для фирм конкурентного окружения (предположения мало- и полноводного года), модель с эластичным и неэластичным спросом.

Таблица 7

Результаты расчётов равновесных цен по моделям Курно и конкуренции функций предложения в сравнении с ценой Вальраса при реализации различных сценариев

Условия	Цена Вальраса	Цена LSFE	Цена Курно
Базовые условия спроса	255	257	259
Неэластичный спрос (γ умень-шили на 10%)	255	297	370
Маловодный год (выработка ГЭС снизилась на 8%)	264	267	268

В результате моделирования стратегического взаимодействия на оптовом рынке электроэнергетической системы «Сибирь» моделями несовершенной конкуренции было определено:

1) меньшую равновесную цену дают модели, в которых все генераторы используют равновесные функции предложения с наличием конкурентного окружения (в нашем случае линейные функции предложения);

2) при функционировании с неэластичным спросом у фирм существует возможность значительно завышать цены относительно цен, ориентированных на предельные издержки (цены Вальраса);

3) при резком ограничении участия фирм конкурентного окружения (маловодный год) возрастание цен не сочетается с увеличением разброса цен, рассчитанных по разным моделям, в том числе модели Вальраса.

Последнее говорит о том, что рыночная власть стратегических фирм возрастает слабо, а значительное повышение цен определяется рыночной властью гидроэлектростанций. Соответственно, необходимо вводить их в модель как активных игроков.

В российском законодательстве функции предложения генерирующих компаний при подаче заявок должны представлять из себя линейные либо ступенчатые функции (не более трёх ступеней). Во всех обсуждаемых моделях стратегии производителя соответствуют именно таким функциям. Какой вид заявки будет выбран отдельным действующим агентом, зависит от многих факторов, в том числе информируемости и преследуемых целей. Нас заинтересовал вопрос, как изменятся характеристики рынка, если генерирующие компании будут учитывать своих конкурентов разными способами. Например, один из них будет руководствоваться стратегией, соответствующей линейным функциям предложения, а все остальные будут формировать заявки в соответствии с Курно (первый столбец в табл.8). Оказывается, генерирующая компания, следующая LSFE, получит прибыль больше, чем, если бы она со всеми следовала стратегии Курно, соответственно, прибыли других игроков снизятся. В то же время, если все действуют в соответствии с функцией предложения LSFE, то прибыль в целом уменьшается для каждого (второй столбец в табл.8). Однозначно определить, как будет действовать генератор, невозможно. Поэтому таких ситуаций может быть бесконечное множество и это существенно усложняет анализ и возможность эффективного моделирования рынка. С одной стороны, стратегия LSFE даёт возможность учесть как высокий, так и низкий спрос, подавая заявку в виде зависимости объёма от цены как возрастающую функцию. С другой стороны, следование такой стратегии всех игроков приведёт к меньшей цене и меньшим прибылям, чем если бы все оставались на заявках по типу Курно, где поставщик однозначно определяет цену и объём заявки в момент подачи.

Таблица 8

Характеристики рынка (цена и прибыль) в случае выбора компаниями разных типов заявок (1-й – LSFE (10), остальные – Курно (7)) в сравнении с условиями, когда все агенты будут придерживаться одинаковых способов формирования функций предложения

Прибыль (тыс. руб.)	Вид рынка (функции предложения)		
	Все – Курно, а 1-й (Иркутск) – LSFE	Модель LSFE	Модель Курно
цена	258,15	257	259
Иркутск	448,27	446,61	448,16
Гусиноозер.ГРЭС	157,49	156,97	157,58
Харанор. ГРЭС	52,58	52,40	52,61
Красноярск	665,96	664,39	666,52
Новосибирск	333,73	332,78	333,99
Кузбасс	352,70	351,51	353,00

Необходимо отметить, что один из возможных способов формирования заявок на рынке электроэнергии России на сутки вперёд организован так, что генерирующая компания подаёт заявку в виде ступенчатой функции на каждый час следующего дня. Функция может включать в себя не более трех ступенек, две из которых (верхняя и нижняя) по сути определяется техническими характеристиками станции. Игра (формирование стратегии) идёт только по средней ступеньке. Это означает, что стратегия участника рынка – однозначно определить в момент подачи цену и объем. В предположении ограниченного числа компаний, конкурирующих на рынке, для формирования заявки достаточно знать объёмы, предлагаемые другими агентами, и эластичность функции спроса. В этом случае всем компаниям выгодно придерживаться стратегии Курно с максимальными индексами влияния. Таким образом, правила нашего рынка заранее стимулируют генераторов придерживаться менее выгодной стратегии для общества – Курно. Формирование заявок с большим количеством ступеней, линейных, агрегированных по времени суток будет стимулировать компании формировать стратегии в соответствии с моделью равновесия функций предложения.

Заключение

Нерегулируемый рынок даёт значительную свободу производителю энергии при формировании своей стратегии. Отклонение от предельных издержек функций предложения, которые подает производитель оператору рынка, приводит к сокращению объёмов производства и увеличению цен. Это формирует неоптимальную для общества ситуацию, снижающую эффективность и надёжность снабжения потребителя электроэнергией. Необходимо ответить на вопрос, насколько выгодно отклоняться поставщикам от ситуации совершенной конкуренции. Для электроэнергетики рационально рассматривать в сравнении стратегии поведения двух типов (и наши исследования это подтвердили). Это стратегия, приводящая к модели Курно и учитывающая остаточный спрос (который остаётся неудовлетворённым после действий конкурентов), а также эластичность общей функции спроса. Второй тип – стратегии, приводящий к модели равновесия функций предложения, где фирма ориентируется на скорость изменения общего объёма производства рынка в зависимости от скорости изменения её собственного объёма (в Курно изменения в общем объёме выпуска, вносимые другими участниками рынка, игнорируются – индекс влияния максимальный).

Из теории известно, что чем мельче доля каждой фирмы, тем равновесная цена ближе к предельным издержкам и объёмы производства близки к оптимальным. Для электроэнергетики это не выполняется. Здесь играет роль эффект экономии от масштаба: чем крупнее генераторы, тем ниже предлагаемые цены, выше объёмы выпуска, надёжнее поставки. Неоправданному росту цен противодействует также значительная открытость информации на энергорынке. Знание издержек конкурентов даёт возможность формировать представления о влиянии на рынок действий по завышению и занижению цен своих и конкурентов, т.е. при правильном формировании механизма функционирования рынка имеются все предпосылки для реализации исходов, описывающихся моделью предполагаемых функций предложения. В то же время на сегодняшний день система подачи заявок сформирована таким об-

разом, что стимулирует производителей подавать заявки, приводящие к модели Курно.

В нашей работе выделены адекватные модели поиска равновесных цен, учитывающие различные стратегии поведения производителей, для электроэнергетических спотовых рынков при отсутствии ограничений на передачу энергии. Надо признать, что полученные результаты не позволяют в полной мере оценить и проанализировать все особенности электроэнергетического рынка Сибири. Это связано, в том числе, со следующими факторами: для таких энергосистем, как Сибирь, нельзя пренебрегать сетевыми ограничениями в силу расположения и протяжённости линий электропередач. Большая доля ГЭС в генерации диктует необходимость введения этих станций на рынок в качестве стратегических игроков, а это достаточно проблематично. Исторически сложилось, что предельные издержки гидроэлектростанций принято считать нулевыми, построение привычных моделей с такими функциями невозможно. Поэтому единственный путь – менять представления о затратах ГЭС, например, задаться ценностью воды. Можно строить модели со стохастической характеристикой объёмов, заявляемых гидростанциями, зависящих не только от приточности, но и от желаний владельцев ГЭС.

Рассмотренные в работе небольшие, легко анализируемые модели важны для проигрывания результатов большого количества вариантов функционирования и выбора наиболее эффективного механизма организации взаимодействия рыночных агентов. Моделирование может решить ряд вопросов: определение рыночной власти у агентов и её сила, перспективы развития системы и анализ возможных последствий воздействий на рыночную структуру, в том числе, выбор антимонопольного регулирования. Важно выбрать механизм организации, обладающий не только качествами, оптимизирующими рынок, но и стимулирующий его участников к устойчивому, предсказуемому поведению. Эти вопросы касаются области теории экономических механизмов (mechanism design) и требуют отдельного обсуждения.

Список литературы

[1] Аболмасов А., Колодин Д. Конкурентный рынок или создание монополий: структурные проблемы российского оптового рынка электроэнергии. ERRC final report, 2002.

[2] Беляев Л.С. Проблемы электроэнергетического рынка. Новосибирск: Наука, 2009.

[3] Давидсон М.Р., Догадушкина Ю.В., Крейнес Е.М., Новикова Н.М., Удальцов Ю.А., Ширяева Л.В. Математическая модель конкурентного оптового рынка электроэнергии в России. // Известия Академии Наук. Теория и системы управления. N3, 2004. – С. 72 – 83.

[4] Подковальников С.В., Хамисов О.В. Несовершенные электроэнергетические рынки. Моделирование и исследование развития генерирующих мощностей. // Известия Российской академии наук. Энергетика, 2011. – № 2. – С. 57-76.

[5] Стофт С. Экономика энергосистем. Введение в проектирование рынков электроэнергии: Пер. с англ. М.: Мир, 2006.

[6] Федеральный закон от 26 марта 2003 года N 35-ФЗ «Об электроэнергетике».

[7] Baldick R., Hogan W.W. Capacity Constrained Supply Function Equilibrium Models of Electricity Markets: Stability, Non-decreasing Constraints, and Function Space Iterations, POWER Working paper, Revised August, 2002.

[8] Baldick R., Grant R., Kahn E. Theory and application of linear supply function equilibrium in electricity markets // Journal of Regulatory Economics, 2004. – Vol. 25(2). – P. 143 – 167.

[9] Bompard E., Lu W., Napoli R., Jiang X. A supply function model for representing the strategic bidding of the producers in constrained electricity markets// Electrical Power and Energy Systems, 2002. – Vol. 32. – P. 678 – 687.

[10] Bulavsky V.A., Kalashnikov V.V. Equilibrium in generalized Cournot and Stackelberg models // Economics and Mathematical Methods (Ekonomika i Matematicheskie Metody), 1995. – Vol. 31. – P. 164 – 176.

[11] Bulavsky V.A., Kalashnikov V.V., Kalashnikova N.I., F. J.Castillo Pérez Mixed Oligopoly with Consistent Conjectures. // The European Journal of Operational Research, 2011. – Vol 210(3). – P. 729 – 735.

[12] Day C.J., Hobbs B.F., Pang J. Oligopolistic competition in power networks: a conjectured supply function approach. // IEEE transactions on power systems, 2002. – Vol. 17(3). – P. 597 – 607.

[13] Hobbs B., Metzler C., Pang J. Strategic Gaming Analysis for Electric Power Systems: An MPEC Approach // IEEE transactions on power systems, 2000. – Vol. 15(2). – P. 638 – 645.

[14] Hu X., Rulph D. Using EPECs to Model Bilevel Games in Restructured Electricity Markets with Locational Prices // Operations research, 2007. – Vol. 55(5). – P. 809 – 827.

[15] Klemperer P., Meyer M. Supply Function Equilibria in Oligopoly under uncertainty // Econometrica, 1989. – Vol. 57(6). – P. 1243 – 1277.

[16] Liu Youfei, Ni Y.X., Wu F.F., Cai Bin Existence and uniqueness of consistent conjectural variation equilibrium in electricity markets // International Journal of Electrical Power and Energy Systems, 2007. – Vol. 29.– P. 455 – 461.

[17] Mas-Colell A., Whinston M., Green J. Microeconomic Theory. New York, Oxford University Press, 1995.

[18] Vasin A.A., Vasina P.A. Models of supply functions competition with application to the network auctions. Moscow: EERC, 2005.

[19] Vasin A.A., Vasina P.A. Homogeneous Good Markets and Auctions // Working Paper 2005/047. Moscow, New Economic School, 2005.

УДК 519.856.3: 631.153

МОДЕЛИ ОПТИМИЗАЦИИ РАЗМЕЩЕНИЯ ПОСЕВОВ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР В УСЛОВИЯХ НЕПОЛНОЙ ИНФОРМАЦИИ

Астафьева М.Н.

Иркутская государственная сельскохозяйственная академия

E-mail: astafeva-mn@yandex.ru

Иваньо Я.М.

Иркутская государственная сельскохозяйственная академия

E-mail: iymex@rambler.ru

Аннотация. В статье рассматриваются модели оптимизации размещения сельскохозяйственных культур с вероятностными и интервальными параметрами, к которым относятся ряды урожайности различных сельскохозяйственных культур и производственных ресурсов. Показана возможность применения двух алгоритмов с использованием метода статистических испытаний для получения оптимальных планов структуры посевных площадей. Приведены результаты моделирования на основе данных по Иркутской области.

Ключевые слова. Задача математического программирования, интервальные оценки, вероятностный параметр, метод статистических испытаний, размещение сельскохозяйственных культур.

Введение

Применение оптимизационных моделей позволяет получить ощутимый результат, так как при планировании аграрного производства часто приходится сталкиваться с проблемой выбора оптимальных вариантов использования земли, трудовых и материально-денежных ресурсов, техники, удобрений и т.д. В настоящее время успешно решаются задачи размещения, специализации и концентрации сельскохозяйственного производства, определения оптимальных размеров предприятий по зонам, эффективности капиталовложений, планирования материально-технического снабжения, отраслевой структуры предприятий, оптимального распределения минеральных удобрений, оптимизации машинно-тракторного парка и его использования и др.

Планирование аграрного производства в условиях неполной информации является трудоемким процессом. Помимо этого, моделирование ослож-

няется влиянием на параметры сельскохозяйственного производства множества природных и антропогенных факторов. Поэтому в работах [1, 2] для оптимизации аграрного производства предлагается использовать имитационное моделирование, которое позволяет увеличить количество возможных вариантов решения задачи.

В работе объектом моделирования является растениеводческая отрасль. Задача математического программирования ориентирована на критерий минимизации затрат на производство. Применение этого критерия связано с неустойчивым характером производства сельскохозяйственной продукции, поскольку для многих предприятий получение продукции скотоводства и растениеводства является убыточным. При интеграционных процессах слабые хозяйства присоединяются к предприятиям, работающим устойчиво. Поэтому критерий в виде затрат в большей степени соответствует сегодняшнему состоянию сельского хозяйства.

В качестве переменных величин обычно применяют искомые размеры площадей сельскохозяйственных культур и объем получаемого валового сбора. С учетом этой специфики запишем математическую модель, которая учитывает особенности задачи и конкретную информацию о предприятии.

Математическая модель оптимизации размещения растениеводческой продукции записывается в следующем виде. Целевая функция направлена на обеспечение минимума затрат:

$$\sum_{i\in I}\sum_{s\in S} c_{is}x_{is} \to \min, \tag{1}$$

при условиях:

1) ограниченности производственных ресурсов

$$\sum_{s\in S} v_{lis}x_{is} \le V_{li},\ l\in L; \tag{2}$$

2) ограниченности размера растениеводческой отрасли

$$\underline{n}_{ri} \le \sum_{s\in S_r}(1+\alpha_s)x_{is} \le \overline{n}_{ri},\ r\in R; \tag{3}$$

3) производства конечной продукции не менее заданного объема

$$\sum_{s\in S} y_{isq} x_{is} \geq Y_{iq},\ q \in S; \tag{4}$$

4) по количеству вносимых удобрений и средств защиты растений

$$\sum_{s\in S} v_{mis} \leq V_{mi},\ m \in M; \tag{5}$$

5) неотрицательности переменных

$$x_{is} \geq 0; \tag{6}$$

6) по реализации произведенной продукции

$$\sum_{s\in S} y_{is} x_{is} \geq \sum_{k\in K} b_k, \tag{7}$$

$$\sum_{q\in Q} y_{iq} x_{iq} \geq \sum_{w\in W} a_w. \tag{8}$$

Здесь x_{is} – искомая переменная, вид сельскохозяйственных угодий *i*–го поля или площадь *s*–ой культуры (га), c_{is} – затраты на 1 га сельскохозяйственных угодий *i*–го поля или *s*–ой культуры (га), v_{lis} – расход l–го ресурса на единицу площади s–ой культуры или вида сельскохозяйственных угодий *i*–го поля (тыс. чел.-ч/га, тыс. руб./га), V_{li} – наличие ресурса l–го вида *i*–го поля, y_{isq} – соответственно выход товарной продукции q–го вида с единицы площади s–ой культуры *i*–го поля (ц/га), Y_{iq} – гарантированный (обязательный) объем производства продукции q–го вида с *i*–го поля (ц), $\bar{n}_{ri}$, $\underline{n}_{ri}$ – максимально и минимально возможная площадь культур r–ой группы *i*–го поля (га), v_{mis} – расход m–го удобрения (средства защиты растений) на единицу площади s-культуры или вида сельскохозяйственных угодий *i*–го поля (ц/га), V_{mi} – наличие удобрения m–го вида *i*–го поля (ц), b_k – объемы реализации продукции на *k*–ом месте сбыта (рынки, оптовые базы и др.), a_w – объемы реализации товарной продукции на *w*–ом рынке.

При решении задачи (1) – (8) возникают сложности, связанные с распределением площадей по предприятиям, расположенным на различных расстояниях друг от друга. Помимо этого, необходимо учитывать типы почв

(чернозем, краснозем, бурозем, дерново-подзолистые и др.), предшественник, периодичность внесения удобрений и др.

Следует отметить, что параметры ограничений (2), (4), (7), (8) являются, как правило, неопределенными величинами. В одних случаях урожайности сельскохозяйственных культур (y_{isq}) характеризуются значимыми трендами или высокими коэффициентами автокорреляции, а в других – представляют собой случайные или слабосвязные выборки. Аналогичные особенности имеют место для трудовых и некоторых других производственных ресурсов. При наличии трендов или высоких значимых коэффициентов автокорреляции условия (2), (4), (7) и (8) связаны с параметром *t*:

$$\sum_{s\in S} v_{lis}(t)x_{is} \le V_{li},\ l \in L, \tag{9}$$

$$\sum_{s\in S} y_{isq}(t)x_{is} \ge Y_{iq},\ q \in S, \tag{10}$$

$$\sum_{s\in S} y_{is}(t)x_{is} \ge \sum_{k\in K} b_k, \tag{11}$$

$$\sum_{q\in Q} y_{iq}(t)x_{iq} \ge \sum_{w\in W} a_w. \tag{12}$$

Если же коэффициенты ограничений (2), (4), (7) и (8) представляют собой случайные и слабосвязные выборки и зависят от вероятности *p*, они примут следующий вид:

$$\sum_{s\in S} v_{lis}^{p} x_{is} \le V_{li},\ l \in L, \tag{13}$$

$$\sum_{s\in S} y_{isq}^{p} x_{is} \ge Y_{iq},\ q \in S, \tag{14}$$

$$\sum_{s\in S} y_{is}^{p} x_{is} \ge \sum_{k\in K} b_k, \tag{15}$$

$$\sum_{q\in Q} y_{iq}^{p} x_{iq} \ge \sum_{w\in W} a_w. \tag{16}$$

Наличие в задаче (1) – (8) случайных параметров позволяет использовать метод статистических испытаний при ее решении. При этом оптимальный план задачи связывается с некоторой вероятностью, представляющей

собой сумму вероятностей урожайности сельскохозяйственных культур, трудовых и других ресурсов:

$$\xi^j = \sum_{l \in Li \in I} \sum_{s \in Si \in I} \sum p^j_{lisq}, \qquad (17)$$

где $j = \overline{1, N}$, N – число экспериментов.

На рис. 1 приведен алгоритм решения оптимизационной задачи размещения сельскохозяйственных культур при вероятностных значениях урожайности с использованием метода Монте-Карло.

Задача оптимизации размещения посевов сельскохозяйственных культур решается в следующей последовательности. Во-первых, вычисляются статистические параметры: средние значения ($\overline{y}_{isq}$, $\overline{v}_{lis}$) и коэффициенты вариации ($C_{v_{Y_{isq}}}$, $C_{v_{vlis}}$). По критерию согласия выбирается закон распределения вероятностей. Во-вторых, моделируются случайные числа, характеризующие ординаты функции распределения p_{lisq}. На основании формулы (17) вычисляются суммарные вероятности. В-третьих, определяются значения урожайности y_{isq} и ресурсов v_{lis}, соответствующие моделируемым вероятностям p_{lisq}. В-четвертых, по полученным значениям y_{isq} и v_{lis} вычисляется критерий оптимальности в виде затрат на получение сельскохозяйственной продукции f_{ξ^j}. По значениям целевой функции строится функция распределения вероятностей. Алгоритм приведен на рис. 1.

Очевидно, что предложенная задача является довольно сложной, если учесть, что в Иркутской области выращивается около 40 сельскохозяйственных культур различных сортов. Помимо этого, имеется большое количество полей, для обработки которых необходимы различные ресурсы.

Следует отметить, что в задаче (1)-(8) используемые ряды урожайности сельскохозяйственных культур и ресурсов являются неоднородными и непродолжительными. В этом случае можно применять модели с интервальными оценками:

$$\sum_{s\in S}\widetilde{y}_{isq}x_{is} \geq Y_{iq},\ q\in S, \tag{18}$$

$$\sum_{s\in S}\widetilde{y}_{is}x_{is} \geq \sum_{k\in K}b_k, \tag{19}$$

$$\sum_{q\in Q}\widetilde{y}_{iq}x_{iq} \geq \sum_{w\in W}a_w, \tag{20}$$

$$\underline{\widetilde{y}}_{isq} \leq \widetilde{y}_{isq} \leq \overline{\widetilde{y}}_{isq}, \tag{21}$$

$$\sum_{s\in S}\widetilde{v}_{lis}x_{is} \geq V_{li}, l\in L, \tag{22}$$

$$\underline{\widetilde{v}}_{lis} \leq \widetilde{v}_{lis} \leq \overline{\widetilde{v}}_{lis}. \tag{23}$$

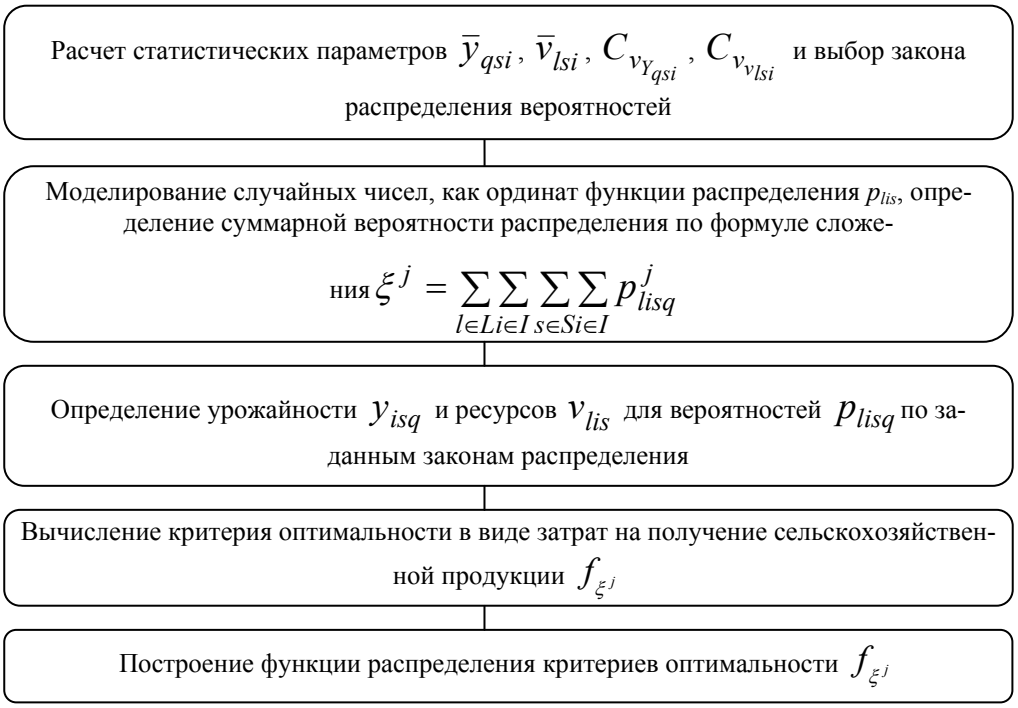


Рис. 1. Алгоритм оптимизации размещения сельскохозяйственных культур при вероятностных значениях урожайности

Знания о верхних и нижних оценках фактической урожайности сельскохозяйственных культур позволяет моделировать различные ситуации производства в условиях недостаточной и неоднородной информации с применением методов имитационного моделирования [3].

Рассмотрим алгоритм оптимизации размещения сельскохозяйственных культур на основе приведенных выше подходов оценки верхних и нижних значений урожайности сельскохозяйственных культур. При использовании

алгоритма сначала согласно пространственно-временному анализу с учетом природно-климатических особенностей рассматриваемых территорий определяются экстремальные значения урожайности сельскохозяйственных культур и ресурсов $y_{isq}^{\max}$, $y_{isq}^{\min}$, $v_{lis}^{\max}$, $v_{lis}^{\min}$. На втором этапе с использованием экстремумов моделируются ряды урожайности y_{isq}^{j} и ресурсов v_{lis}^{j} в виде случайных чисел, где $i \in I$, $s \in S$, $l \in L$, $j \in J$. Затем по полученным значениям строится некоторое число оптимизационных моделей размещения сельскохозяйственных культур с критерием f^{j}. Из них выбираются максимальное, минимальное значения и медиана целевой функции ($f_{\min}^{\max}$, $f_{\min}^{\min}$, $f_{\min}^{med}$), в качестве которой использованы затраты на производство сельскохозяйственной продукции. Описанный алгоритм показан на рис. 2.

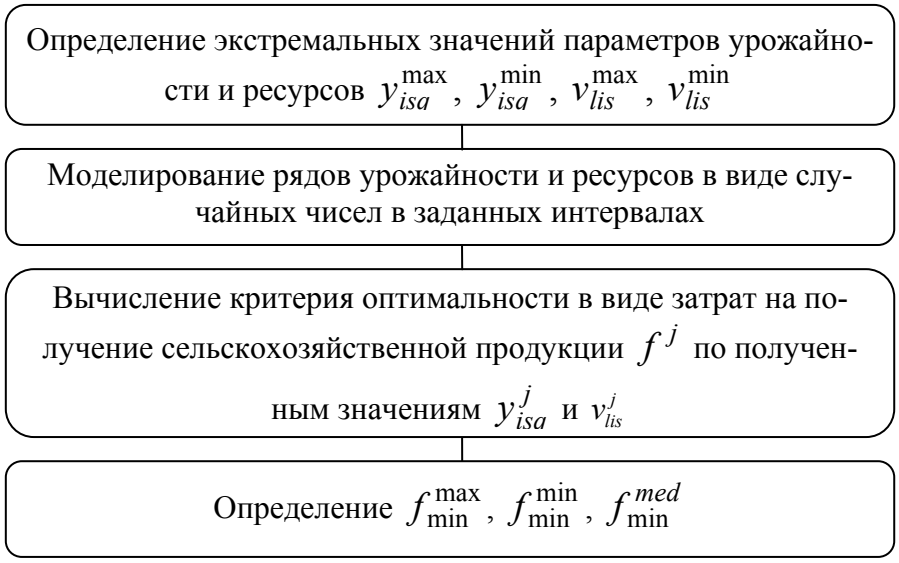


Рис. 2. Алгоритм оптимизации размещения сельскохозяйственных культур с использованием метода статистических испытаний для интервальных величин урожайности сельскохозяйственных культур

Приведенный алгоритм позволяет расширить возможности моделирования размещения сельскохозяйственных культур, поскольку ряды урожайности являются короткими и обладают неоднородностью, что приводит к значительным стандартным ошибкам параметров модели.

Первый алгоритм (рис. 1) реализован для оптимизации размещения посевов сельскохозяйственных культур Иркутской области [4]. Для описания параметров модели с помощью вероятности использована функция Гаусса. Ряды урожайности сельскохозяйственных и плодово-ягодных культур Иркутской области объединены в 4 группы: зерновые и зернобобовые, картофель и овощебахчевые, кормовые и плодово-ягодные. Число экспериментов составило 100 значений. Моделирование различных ситуаций согласно алгоритму показало степень изменчивости неизвестных величин в моделях.

На рис. 3 приведена связь значений целевых функции f^j и суммарных вероятностей урожайностей культур ξ^j, использованных при построении моделей.

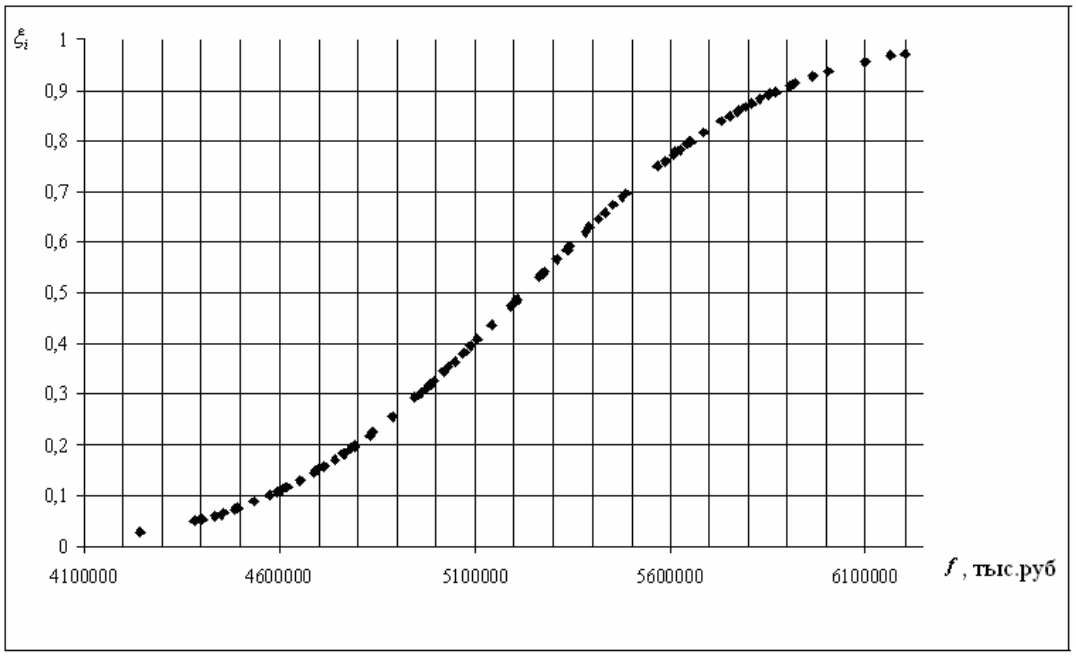


Рис. 3. Распределение вероятностей критерия оптимальности при количестве экспериментов 100

На основании функции распределения приведены результаты моделирования для вероятностей 0.9, 0.5 и 0.1 (табл. 1).

Таблица 1

Результаты решения задачи оптимизации размещения сельскохозяйственных и плодово-ягодных культур для суммарной вероятности 0.9, 0.5 и 0.1

Целевая функция, тыс. руб.		Значения посевных площадей *x,* га			
		Зерновые и зернобобовые культуры	Картофель и овощебахчевые культуры	Кормовые культуры	Плодово-ягодные культуры
$f_{0,1}$	4573338	383412	68377	142895	864
$f_{0,5}$	5209614	440296	68395	146145	1112
$f_{0,9}$	5869003	542731	68870	178789	1932

Применение моделей с высокими урожайностями показало, что для ситуации, соответствующей вероятности 0.1, затраты на производство необходимо объема продукции составили 1295665 тыс. руб. В противном случае, когда на урожайность влияют неблагоприятные факторы (вероятность превышения равна 0.9), критерий оптимальности увеличивается на 28%. Наибольшее расхождение между максимальными и минимальными площадями выявлено для плодово-ягодных культур (2.2 раза), а наименьшее – для картофеля и овощебахчевых (0.03%).

Второй алгоритм (рис. 2) реализован для лесостепной зоны Иркутской области [5]. В качестве исходных данных использованы урожайности зерновых культур, овощей и картофеля (n = 3). Число экспериментов изменялось, составив N = 10, 25 и 50. Моделирование различных ситуаций согласно алгоритму показало степень изменчивости неизвестных величин в моделях. Решения получены для следующего диапазона урожайности сельскохозяйственных культур, найденного путем усреднения данных по муниципальным районам лесостепной зоны: зерновые 13 – 17, овощи 74 – 259, картофель 128 – 152 ц/га.

Согласно табл. 2 коэффициенты вариации площадей как неизвестных величин модели для различных значений параметра N составляют 0,017-0,12.

Наименьшее рассеяние имеет место для картофеля, а наибольшее – для овощей. Между тем расхождение между минимальными и максимальными значениями относительно среднего значения варьирует для зерновых культур в пределах 9.5-13.3%, картофеля – 6.7-8.5%, а овощей – 35.2-91.3%, что в значительной степени сказывается на затратах при определении структуры площадей.

Таблица 2

Результаты решения задачи оптимизации размещения сельскохозяйственных культур с верхними и нижними оценками урожайности

Статистические параметры / Сельскохозяйственные культуры	Среднее значение $\bar{x}$, га	Стандартное отклонение σ, га	Минимальное значение $x_{\min}$, га	Максимальное значение $x_{\max}$, га
N=10				
Зерновые	172466	5359	153293	195420
Картофель	1425	24	1335	1544
Овощи	8061	931	5736	15328
N=25				
Зерновые	174211	3261	152381	195974
Картофель	1449	15	1356	1565
Овощи	9184	743	6194	17570
N=50				
Зерновые	173626	2140	151479	190145
Картофель	1445	10	1332	1565
Овощи	9051	460	5868	12686

В табл. 3 приведены результаты решения задачи размещения сельскохозяйственных культур для экстремальных значений целевой функции и ее медианы. Расхождение между минимальными и максимальными значениями целевой функции (N = 10) составило 624255 тыс. руб. или 36.8% относительно медианы. Во втором и третьем случаях (N = 25 и 50) разница достигла 37.0 и 26.4%. При этом отклонение между посевными площадями (x) относительно среднего значения в первом случае (N = 10) составило для зерновых культур 24.6, картофеля – 4.7, овощей – 106.2%. В свою очередь для зерновых этот показатель с увеличением N уменьшается, а для картофеля и овощей – увеличивается.

Таблица 3

Результаты решения задачи оптимизации размещения сельскохозяйственных культур для экстремальных значений целевой функции и ее медианы

Целевая функция, тыс. руб.		Значения посевных площадей *x,* га		
		Зерновые	Картофель	Овощи
$N = 10$				
$f_{\min}^{\min}$	1469433	153293	1341	5863
$f_{\min}^{med}$	1695951	165161	1380	8912
$f_{\min}^{\max}$	2093688	193939	1406	15327
N=25				
$f_{\min}^{\min}$	1474761	152381	1380	6193
$f_{\min}^{med}$	1777879	182857	1402	7485
$f_{\min}^{\max}$	2133293	192481	1461	17038
N=50				
$f_{\min}^{\min}$	1480394	152381	1400	5868
$f_{\min}^{med}$	1727231	185507	1538	6424
$f_{\min}^{\max}$	1935698	188235	1832	11741

Таким образом, модели оптимизации размещения сельскохозяйственных культур, построенные в работе, применимы для предприятий со сложной структурой производства.

Предложен алгоритм решения задачи оптимизации структуры посевов с вероятностными параметрами биопродуктивности и приведенных ресурсов, реализованный с использованием данных по муниципальным районам Иркутской области.

Поскольку ряды урожайности сельскохозяйственных культур являются короткими и неоднородными предложено решать задачи оптимизации распределения площадей посевов с интервальными оценками. При этом применим метод статистических испытаний, адекватно отражающий значения интервальных оценок. Алгоритм решения задачи математического программирования с интервальными оценками реализован на реальных объектах.

Предложенные модели усложняются при увеличении числа переменных и ограничений за счет включения в систему всего разнообразия сельскохозяйственных культур, возделываемых в различных предприятиях, сельскохозяйственных зонах, муниципальных образованиях и регионе. Приведенные результаты показывают, что затраты, а следовательно и площади посевов, чувствительно реагируют на значительную изменчивость урожайности сельскохозяйственных культур, что сказывается на управлении аграрным производством.

Список литературы

[1] Вашукевич Е.В., Елохин В.Р., Иваньо Я.М., Труфанова Е.С. Об использовании имитационного моделирования для решения задач аграрного производства // Сб. статей международ. науч.-практ. конф. «Природа и сельскохозяйственная деятельность человека». – Иркутск, 2011. – С. 185 – 188.

[2] Труфанова Е.С. Алгоритмы имитационного моделирования производственного потенциала земельных ресурсов региона // Вестник Брянского гос. тех. ун-та. – 2011. – № 1(29). – С. 79 – 84.

[3] Емельянов А.А., Власова Е.А., Дума Р.В. и др. Имитационное моделирование экономических процессов: Учебное пособие. – М.: Финансы и статистика, 2002. – 368 с.

[4] Астафьева М.Н., Иваньо Я.М. Модели оптимизации размещения сельскохозяйственных и плодово-ягодных культур с вероятностными параметрами в условиях неопределенности // Научно-практический журнал «Вестник ИрГСХА». – 2011. – Вып. 48. – С. 12 – 20.

[5] Астафьева М.Н., Иваньо Я.М. Оптимизация размещения посевов сельскохозяйственных культур с использованием имитационного моделирования // Научно-практический журнал «Актуальные проблемы аграрной науки». – 2011. – Вып. 1. – С. 59 – 67.

УДК 519.855: 63

ЗАДАЧИ ПАРАМЕТРИЧЕСКОГО ПРОГРАММИРОВАНИЯ С ВЕРОЯТНОСТНЫМИ И ИНТЕРВАЛЬНЫМИ ПАРАМЕТРАМИ ПРИМЕНИТЕЛЬНО К ЗАДАЧАМ ОПТИМИЗАЦИИ СЕЛЬСКОХОЗЯЙСТВЕННОГО ПРОИЗВОДСТВА

Барсукова М. Н.

Иркутская государственная сельскохозяйственная академия

E-mail: bmn1982@rambler.ru

Иваньо Я. М.

Иркутская государственная сельскохозяйственная академия

E-mail: iymex@rambler.ru

Аннотация. В работе приведены задачи параметрического программирования с одним и множеством параметров, с помощью которых можно оптимизировать производство растениеводческой, животноводческой продукции и их сочетания. Рассмотрены особенности данных, согласно которым модели характеризуются различной структурой. При наличии случайных параметров, подчиняющихся законам распределения вероятностей, для решения прикладных задач эффективным является метод статистических испытаний. Предложенные модели реализованы для агропромышленных предприятий региона.

Ключевые слова. Задача параметрического программирования, факторы, неопределенность, вероятностные оценки, интервальные оценки, сельскохозяйственное производство.

Введение

Экономико-математические модели, используемые для оптимизации производства сельскохозяйственной продукции, обычно не вкладываются в задачи линейного программирования. При разработке оптимального решения для сельскохозяйственного предприятия возникает проблема выбора адекватных математических методов и моделей, позволяющих отражать структуру производства, оперировать оценками экспертов, учитывать неясность, неточность данных. Другими словами, при решении практических задач можно использовать оптимизационные модели в условиях неопределенности.

Анализ результатов работы множества сельскохозяйственных предприятий показывает, что они могу быть выделены в группы, в одной из которых преобладают детерминированные производственно-экономические пока-

затели, а в другой - вероятностные и неопределенные параметры [4]. Детерминированные модели предпочтительнее для прогнозирования и планирования развития предприятия. В то же время модели в условиях неопределенности ближе к отображению реальной ситуации, но с их помощью получают многовариантные решения.

Соблюдение современных технологий получения сельскохозяйственной продукции, как правило, уменьшает пагубное влияние природных явлений на финансовое состояние предприятий. Однако в условиях резко континентального климата даже стабильно работающие хозяйства теряют значительную часть урожая в результате влияния на производственные процессы природных событий. Тем не менее для стабильных предприятий характерна неубывающая многолетняя тенденция производственно-экономических параметров.

Статистические исследования многолетних рядов характеристик деятельности устойчивых сельскохозяйственных предприятий позволили выявить следующие закономерности отраслей растениеводства и скотоводства [4]. Во-первых, в большинстве рассматриваемых случаев имеют место значимые линейные тренды. Во-вторых, многолетние ряды, описывающие скотоводство, характеризуются высокими первыми коэффициентами автокорреляции и значимыми авторегрессионными уравнениями при сдвиге на величину 1 – 2 года.

Из всего этого следует, что для оптимизации продукции растениеводства, скотоводства и их сочетания применимы модели задачи параметрического программирования.

В общем виде задача параметрического программирования с одним параметром записывается следующим образом:

$$f = \sum_{j \in J} c_j(t) x_j \to \max(\min), \tag{1}$$

$$\sum_{j \in J} a_{ij}(t) x_j = b_i(t),\ i \in I, \tag{2}$$

$$x_j \geq 0,\ j \in J, \tag{3}$$

где f – целевая функция; x_j – переменная; t – параметр; c_j, a_{ij}, b_i – коэффициенты, связанные с параметром t; j, i, w – индексы, принадлежащие соответствующим множествам J, I и W [1, 2].

Вид параметра задачи параметрического программирования связан со свойствами производственно-экономических показателей. При наличии значимых трендов в качестве параметра можно использовать время, а в случае высоких коэффициентов автокорреляции – предшествующее значение. Реальной является ситуация, когда правая часть ограничений является нелинейной функцией с верхним пределом производственных показателей.

На основе параметрической задачи построены модели для растениеводства, скотоводства и их сочетания. Модели реализованы для известного хозяйства СХОАО «Белореченское» Иркутской области. При решении реальных задач коэффициенты при неизвестных в целевой функции и коэффициенты левых частей ограничений приняты как постоянные значения. Только правые части условий описаны в виде значимых линейных уравнений.

Полученные модели проверены на адекватность и точность. Исследования остатков рядов и оценка значимости уравнений и их коэффициентов позволяют считать модели качественными. Кроме того, результаты ретроспективного прогноза с помощью оптимизационной модели с упреждением 1 год показали удовлетворительную сходимость модельных значений с реальными данными. Наименьшее расхождение между фактическими и прогностическими затратами на производство оказались для скотоводства, составив 3 – 6%. При этом предпочтительнее выглядит модель, включающая в себя авторегрессионные уравнения. Максимальное расхождение между значением целевой функции и фактическими затратами получено для растениеводства (25%), что вполне объяснимо, поскольку эта отрасль сельского хозяйства в наибольшей степени подвержена влиянию климатических условий и факторов подстилающей поверхности. По результатам моделирования сочетания отраслей отличие прогностических и реальных значений затрат составили около 20%.

Таким образом, моделирование производства продукции скотоводства на основе задачи линейного параметрического программирования позволяет прогнозировать параметры модели, связанные со временем и предшествующими значениями рядов. Приемлемые результаты получены для растениеводства и сочетания отраслей – для растениеводства и скотоводства. Однако ретроспективный прогноз показывает не столь высокое качество этих моделей по сравнению с моделью, описывающей скотоводство.

Добавим к сказанному, что наличие случайных составляющих в предложенных детерминированных моделях уменьшает точность полученных результатов. Поэтому помимо определения точечных значений неизвестных модели осуществлена их интервальная оценка при заданном уровне значимости. Для отрасли скотоводство разности между верхним и нижним значением прогнозов составили 5 – 6%. Этот показатель для растениеводства и скотоводства находится на уровне чисел, рассчитанных при сравнении точечных прогностических значений с реальными данными.

Полученные результаты показывают, что для предприятий с неубывающими показателями производства приемлемо моделирование оптимальной структуры отраслей и их сочетания на основе задачи параметрического программирования. Вместе с тем приведенные модели несут в себе ряд допущений и упрощений, не учитывая многие факторы, касающиеся как производственных, так и природных условий. В частности, даже устойчивые сельскохозяйственные предприятия не в состоянии предотвратить воздействие на производство продукции природных стихийных явлений. Так, засуха 2003 г. в Иркутской области отрицательно сказалась на урожайности зерновых культур. Последствия стихии для стабильно работающих организаций привели к уменьшению продукции на 25 – 30%.

Наряду со сложностью учета погодных факторов и разнообразия ландшафтов трудности при моделировании сельскохозяйственных процессов возникаю при оценке некоторых производственно-экономических параметров, входящих в оптимизационные модели. Во многих случаях приведенные за-

траты, стоимостные показатели и коэффициенты, характеризующие трудовые и земельные ресурсы, не являются постоянными величинами. В лучшем случае они колеблются в незначительных пределах и могут усредняться, а в худшем несут в себе неопределенность. Поэтому модели с детерминированными параметрами ограничены для практического использования.

Несмотря на то, что в большинстве устойчивых сельскохозяйственных предприятий наблюдаются значимые тенденции роста по основным производственно-экономическим показателям, многим из них свойственна неопределенность.

Зачастую необходимо знать верхние и нижние оценки $\psi(y)$ или лучший и худший варианты работы предприятия [5]. Для этого нужно решить две следующие задачи

$$\min \{\psi(y) : y \in R_y\}, \tag{5}$$

$$\max \{\psi(y) : y \in R_y\}. \tag{6}$$

Обозначим решение этих задач $\underline{\psi}$, $\overline{\psi}$. При этом, если $\overline{\psi} - \underline{\psi} \le \varepsilon$ (ε – малое положительное число), то задача (5) – (6) устойчива к возмущениям, в противном случае $(\underline{\psi} - \psi >> \varepsilon)$ эти возмущения могут привести к катастрофическим последствиям, и лицу, принимающему решение, следует использовать дополнительные меры для дальнейшего устойчивого функционирования системы.

Приведенная задача применена к определению оптимальной структуры производства отраслей сельского хозяйства и их сочетания в условиях неопределенности. Предложено минимизировать затраты на производство при неопределенных характеристиках в целевой функции и правой части ограничений. При этом использованы задачи как линейного, так и параметрического программирования.

Математическая модель сочетания отраслей с использованием неопределенных параметров записывается следующим образом. Критерий оптимальности имеет вид

$$\sum_{s\in S} c_s x_s + \sum_{h\in H} c_h x_h \to \min, \tag{7}$$

при условии, что затраты на получение продукции с 1 га (c_s) и 1 головы скота (c_h) находятся в некоторых пределах: $\underline{c_s} \le c_s \le \overline{c_s}$, $\underline{c_h} \le c_h \le \overline{c_h}$.

Модель связана с рядом условий:

1) ограниченностью производственных ресурсов

$$\sum_{s\in S} f_{ls} x_s + \sum_{h\in H} f_{lh} x_h \le F_l,\ l \in L; \tag{8}$$

2) использованием в животноводстве побочной продукции растениеводства

$$\sum_{s\in S} p_{js} x_s \ge x_j,\ \ j \in J; \tag{9}$$

3) ограниченностью размера отраслей, в том числе

а) по растениеводству

$$\underline{n}_r \le \sum_{s\in S_r} (1+\alpha_s) x_s \le \overline{n}_r,\ \ r \in R, \tag{10}$$

б) по скотоводству

$$x_k = \lambda_{hh'} x_{h'},\ h' \in H; \tag{11}$$

4) производством конечной продукции не менее заданного объема, в том числе

а) по растениеводству

$$\sum_{h\in H} v_{qh} x_s \ge V_q,\ q \in Q_1, \tag{12}$$

б) по скотоводству

$$\sum_{h\in H} v_{qh} x_s \ge V_q,\ q \in Q_2 \tag{13}$$

при условии, что $\underline{V_q} \le V_q \le \overline{V_q}$;

5) увязкой растениеводства и скотоводства, в том числе

а) по балансированию рационов животных по элементам питания

$$\sum_{s\in S} a_{is} p_s x_s + \sum_{j\in J} a_{ij} x_j \ge \sum_{h\in H} b_{ih} x_h,\ \ i \in I, \tag{14}$$

б) по структуре производства кормов

$$\sum_{h \in H} \underline{d_{hh}} x_h \le \sum_{s \in S_k} a_{1s} p_s x_s + \sum_{j \in J_k} a_{1j} x_j \,,\ k \in K \,, \qquad (15)$$

6) неотрицательностью переменных

$$x_s, x_h \ge 0 \,. \qquad (16)$$

Здесь x_s – искомая переменная площадь s–ой культуры или вида кормовых угодий, f_{ls} – расход l–го ресурса на единицу площади s–ой культуры или вида кормовых угодий, F_l– наличие ресурса l–го вида, V_q – гарантированный объем производства продукции q–го вида, p_{is} – выход с единицы площади s–го культуры j–го вида корма, x_j – количество кормов j–го вида, используемое скотоводством, $\overline{n_r}$, $\underline{n_r}$ – максимально и минимально возможная площадь культур r–ой группы, v_{qs} – соответственно выход товарной продукции q–го вида с единицы площади s–ой культуры, p_s – выход основной кормовой продукции с единицы площади s–ой культуры или вида кормовых угодий, $\underline{d}_{kh}$, $\overline{d}_{kh}$ – минимально и максимально допустимый нормативный размер потребности в кормах k–ой группы единицы поголовья h–го вида (группы) животных, выраженный в кормовых единицах, a_{is} – содержание i–го элемента питания в единице кормовой продукции, получаемое от s–ой культуры, a_{ij} – содержание i–го элемента питания в j–ом виде корма или компоненте кормосмеси, a_s – коэффициент, учитывающий площадь семенных посевов для s–ой культуры, a_{ij} – содержание i–го элемента питания в j–ом виде корма или компоненте кормосмеси, a_{ls} – содержание кормовых единиц в единице корма, получаемого от s–ой культуры, x_h– искомая переменная с h–ым видом скота, B_h – количество необходимого поголовья, f_{lh} – расход l–го ресурса на единицу поголовья h–го вида (группы) животных, v_{qh} – выход товарной продукции q–го вида с единицы поголовья h–го вида, b_{ih} – минимальная потребность в i–го элементе питания единицы поголовья h–го вида (группы).

При моделировании использованы данные крупного сельскохозяйственного предприятия СХОАО «Белореченское». Полученные результаты показывают возможности вариации параметрами моделей для управления реальными производственными процессами. Обращает на себя внимание тот факт, что разности верхних и нижних оценок в задаче параметрического программирования менее значительны.

Конечно, хозяйству в своей деятельности не следует ориентироваться на экстремальные реализации случайных и неопределенных факторов, однако, такие постановки не только возможны, но и необходимы для прогнозирования и оценки критических ситуаций.

Помимо рассмотренных задач однопараметрического программирования для описания сельскохозяйственного производства возможно использование многопараметрических задач, в которых коэффициенты при неизвестных в целевой функции, коэффициенты при неизвестных в системе уравнений и свободные члены системы уравнений линейно зависят от нескольких параметров.

Для сельского хозяйства задачи многопараметрического программирования имеют практическое и теоретическое значение, поскольку производственно-экономические показатели подвержены влиянию многих факторов. Задачу многопараметрического программирования можно записать в следующем виде [3]:

$$F = \sum_{j \in J} c_j(t_1, t_2, t_3, \ldots t_m) x_j \to \max(\min), \tag{17}$$

$$\sum_{j \in J} a_{ij}(t_1, t_2, t_3, \ldots t_m) x_j \le b_i(t_1, t_2, t_3, \ldots t_m),\ i \in I, \tag{18}$$

$$x_j \ge 0,\ j \in J. \tag{19}$$

где x_j – переменная, $t_1, t_2, t_3, \ldots, t_m$ – параметры, c_j, a_{ij}, b_i – коэффициенты, связанные с параметром t, j, i – индексы, принадлежащие соответствующим множествам J, I.

В частности, урожайность сельскохозяйственных культур зависит от таких случайных параметров как число дней бездождевого периода и сумма месячных осадков за вегетационный период [6].

Одним из методов построения математической модели изменчивости урожайности зерновых культур, является корреляционно-регрессионный анализ.

Многофакторные модели могут быть линейными и нелинейными. Часто для оценки зависимости y от факторов используются связи в виде полинома второй степени.

В общем виде задачи параметрического программирования с учетом линейного и нелинейного тренда в левых частях ограничений с двумя параметрами выглядят следующим образом:

$$\sum_{j\in J}(a_{i0}+a_{i1}t_1+a_{i2}t_2)x_j=b_i\,, \tag{20}$$

$$\sum_{j\in J}(a_{i0}+a_{i1}t_1+a_{i2}t_1^2+a_{i3}t_2+a_{i4}t_2^2+a_{i5}t_1t_2)x_j=b_i\,,\ i\in I\,. \tag{21}$$

Исследования показали, что урожайность зерновых культур можно моделировать с помощью двухфакторных линейных и нелинейных моделей. При этом выявлено преимущество вторых видов моделей над первыми. С учетом выражения (21) предложена задача параметрического программирования с параметрами в левых частях ограничений:

$$F=\sum_{j\in J}c_j(t_1,t_2)x_j\to\max(\min), \tag{22}$$

$$\sum_{j\in J}(a_{i0}+a_{i1}t_1+a_{i2}t_1^2+a_{i3}t_2+a_{i4}t_2^2+a_{i5}t_1t_2)x_j=b_i\,,\ i\in I\,, \tag{23}$$

$$x_j\ge 0\,,\ j\in J\,. \tag{24}$$

Эта задача имеет значение для оптимизации производства сельскохозяйственной продукции применительно к остепненной и лесостепной зонам Иркутской области. При этом уравнения нелинейной регрессии в левых частях ограничений имеют разный вид:

$$\sum_{j\in J}\left(a_{i0}+a_{i1}t_1+a_{i2}t_1^2+a_{i3}t_3+a_{i4}t_3^2\right)x_j=b_i\,, \tag{25}$$

$$\sum_{j\in J}\left(a_{i0}+a_{i1}t_1+a_{i2}t_1^2+a_{i3}t_2+a_{i4}t_2^2\right)x_j=b_i\,,\ i\in I\,, \tag{26}$$

$$\sum_{j\in J}\left(a_{i0}+a_{i1}t_1+a_{i2}t_1^2+a_{i3}t_2+a_{i4}t_1t_2\right)x_j=b_i\,,\ i\in I\,. \tag{27}$$

Первое и второе ограничения справедливы для остепненной зоны, а третье – для лесостепной. В выражениях (25) – (27) t_1 – число дней бездождевого периода, t_2 – сумма месячных осадков, t_3 – сумма средних месячных температур. При этом факторы характеризуют вегетационный период. Приведенные выражения отображают особенности рассматриваемых территорий. В первом случае на урожайность сельскохозяйственных культур влияние оказывает сочетание факторов тепла и увлажнения, а во втором – преобладают факторы увлажнения.

Для решения сформулированных многофакторных параметрических задач могут быть использованы методы имитационного моделирования, поскольку исследования параметров $t_1, t_2, t_3, \ldots, t_m$ показывает, что они являются случайными.

В некоторых ситуациях получение исходных статистических данных путем специально организованных экспериментов невозможно. В этом случае необходимый статистический материал может быть получен с помощью специально созданных математических моделей, в основу которых положен *метод статистических испытаний* (метод Монте-Карло).

Анализ параметров выражений (25) – (27) t_1, t_2 и t_3 показал, что они являются случайными величинами, поэтому возможно моделирование различных вариантов решения задачи математического программирования с использованием метода статистических испытаний. При этом задаются законы распределения вероятностей природных параметров.

С одной стороны такая постановка задачи усложняет процесс нахождения оптимального решения, а с другой – позволяет моделировать различные варианты развития событий в зависимости от климатических условий.

Приведем последовательность решения задачи (22) – (24) с использованием метода Монте-Карло.

1. Вычисляются статистические параметры t_1, t_2 и t_3, коэффициенты вариации C_v и средние значения. Согласно критерию согласия выбирается закон распределения вероятностей, соответствующий эмпирическим данным. Опыт показывает, что к таким законам относится нормальный и гамма-распределение.

2. Моделируются случайные числа, характеризующие ординаты функции распределения p_{mn} (m - номер параметра; n – номер эксперимента), по которым определяются значения параметров с помощью заданных законов распределения вероятностей. При этом $m \in M$, а $n \in N$.

3. На основании формулы (23) строятся ограничения.

4. Решается задача параметрического программирования с учетом вероятностных значений факторов урожайности сельскохозяйственных культур.

Многократное повторение эксперимента позволяет определять различные варианты решения задачи в зависимости от сочетания факторов, соответствующих смоделированным вероятностям. Алгоритм приведен на рис. 1.

Предложенный алгоритм может существенно дополнить результаты моделирования структуры производства сельскохозяйственной продукции на основе задач параметрического программирования. При этом моделируются возможные сочетания погодных условий, влияющих на урожайность различных сельскохозяйственных культур.

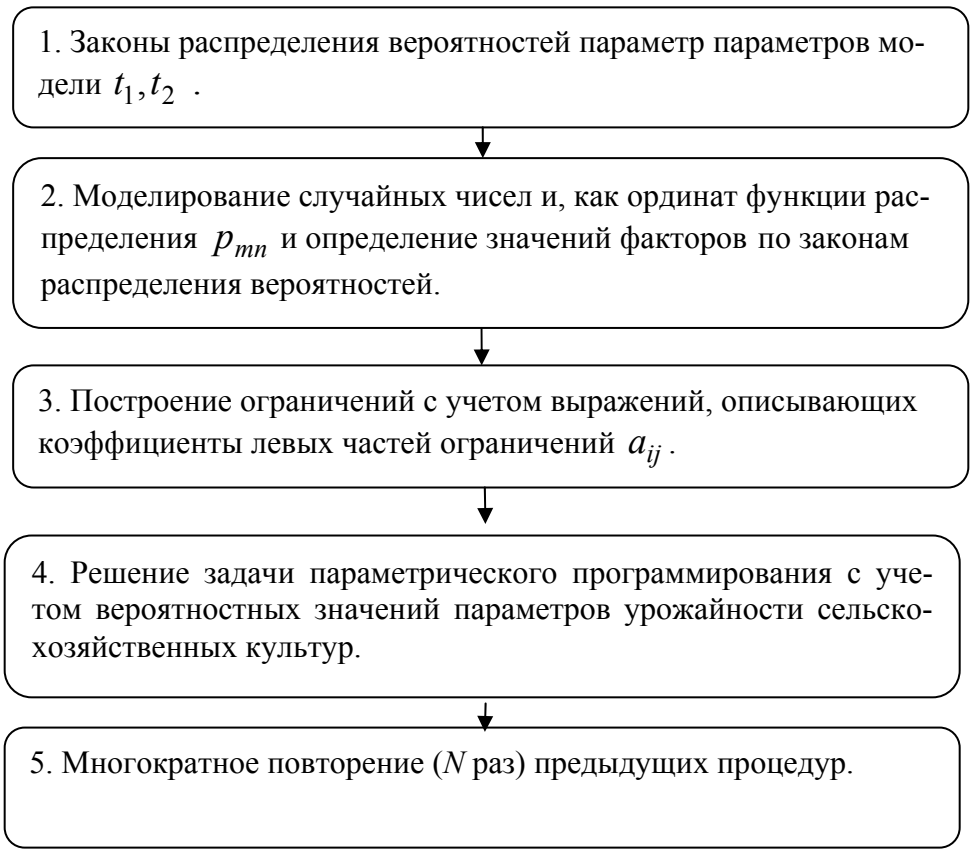


Рис. 1. Алгоритм решения задачи параметрического программирования с учетом вероятностных значений урожайности сельскохозяйственных культур

Кроме того, поскольку уравнения регрессии содержат случайные составляющие, метод статистических испытаний может быть использован для моделирования остатков рядов и обеспечения полной информацией об изменчивости параметров задачи математического программирования.

Список литературы

[1] Акулич И.Л. Математическое программирование в примерах и задачах : учеб. пособие. – 2-е изд., испр. и доп. – М. : Высш. шк., 1993. – 336 с.

[2] Вильямс Н. П. Параметрическое программирование в экономике. – М. : Статистика, 1976. – 398 с.

[3] Барсукова М.Н., Иваньо Я.М. Задача многопараметрического программирования для оптимизации сочетания отраслей сельскохозяйственного предприятия // Совместная деятельность сельскохозяйственных товаропроизводителей и научных организаций в развитии АПК Центральной

Азии: сб.материалов междунар. науч.-практ. конф., (Иркутск, 25 – 28 марта 2008 г.)/ ред. кол.: Я.М.Иваньо [и др.]; Иркут.гос. с.-х. акад. – Иркутск: Изд-во ИрГСХА, 2008. – Ч.4. – С.73 – 78.

[4] Барсукова М.Н., Иваньо Я.М. Модели с детерминированными и неопределенными параметрами применительно к оптимизации сельскохозяйственных процессов // Вестник Московского государственного университета леса – Лесной вестник, 2007. – № 6. – С. 170 – 177.

[5] Булатов В.П., Федурина Н.И. Об одном эффективном методе выпуклого программирования // Дискретный анализ и исследование операций. Сер. 2, 2004. – Т. 11. – № 1. – С. 1 – 5.

[6] Вашукевич Е.В. Статистическая оценка влияния факторов на агрономическую засуху // Сборник материалов международной научно-практической конференции «Совместная деятельность сельскохозяйственных товаропроизводителей и научных организаций в развитии АПК Центральной Азии». – Иркутск, 2008. – С. 89 – 94.

УДК 621.311.016.011.681.518

ОЦЕНКА ПЕРЕМЕННЫХ РЕЖИМА ЭЭС МЕТОДАМИ ВЕРОЯТНОСТНОГО ПОТОКОРАСПРЕДЕЛЕНИЯ

Болоев Е. В.

Ангарская государственная техническая академия

E-mail: boloev@mail.ru

Голуб И. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: golub@isem.sei.irk.ru

Аннотация. Представлены линейные и нелинейные аналитические методы вероятностного потокораспределения, позволяющие определить числовые характеристики переменных режима электроэнергетических систем (ЭЭС) и построить на их основе функции распределения и плотности вероятности с использованием разложения Грама-Шарлье. Проанализирована возможность сочетания линейного аналитического метода с методом сингулярного анализа для выделения в ЭЭС сенсорных переменных. Предложены модификации известных нелинейных аналитических методов, позволяющие получить вероятностные характеристики переменных, близкие к характеристикам, полученным методом Монте-Карло. Разработан учитывающий ограничения неравенства алгоритм вероятностного потокораспределения, который позволяет выбором управляющих воздействий обеспечить требуемую вероятность нахождения сенсорных переменных в заданных границах. Сравнение методов вероятностного потокораспределения приводится на примере тестовой схемы ЭЭС.

Ключевые слова. Сингулярный анализ, сенсорные переменные, методы вероятностного потокораспределения.

Введение

В процессе функционирования ЭЭС подвергается большим и малым внешним возмущениям и реагирует на них изменением переменных режима. Такая реакция зависят как от набора и величины возмущения, так и от таких инвариантных к режиму факторов, как топология и параметры элементов схемы сети.

Знание реакции переменных режима системы на внешние возмущения важно в том случае, если ее результатом является ухудшение таких критериев функционирования ЭЭС, как допустимость режима, статическая и динамическая устойчивость, оперативная надежность и экономичность.

Эксперименты показывают, что возмущения, локализуемые в разных местах ЭЭС, как правило, вызывают заметную реакцию модулей и фаз напряжения в одних и тех же узлах, перетоков мощности, потерь напряжения в одних и тех же связях. Такие элементы схемы сети, переменные режима которых в наибольшей степени изменяются при случайных внешних возмущениях, названы сенсорами [1].

Поскольку переменные режима сенсоров часто определяют критические состояния ЭЭС, их знание необходимо: для усиления сети при проектировании и управлении, определении наиболее ответственных точек их контроля и ускорения процедуры оценки их допустимости в реальном времени, синтезе законов управления.

Но прежде, чем определять значимость реакций с точки зрения упомянутых критериев управления, надо было найти простые способы выявления сенсорных элементов не столь громоздкие, как статистические испытания, а также способы выявления факторов, порождающих сенсоры, чтобы целенаправленно воздействовать на них как при эксплуатации, так и при развитии ЭЭС.

Для этой цели была разработана технология, в основе которой лежат спектральные и топологические свойства электрической сети и синтез законов управления, обеспечивающих наилучшие характеристики переменных режима. В состав этой технологии вошли методы [1] обнаружения сенсорных переменных и слабых мест, базирующиеся на сингулярном анализе матрицы линеаризованной системы уравнений установившегося режима, и проанализированы способы усиления слабых мест коррекцией их параметров, использованием регулирующих устройств и выработкой управляющих воздействий.

Разработанные на основе этих алгоритмов программы применялись для выделения сенсоров и слабых мест большого числа реальных энергосистем различной размерности. Эксперименты подтвердили их эффективность и работоспособность, тем не менее, они показали, что при наличии большого

числа близких по величине минимальных сингулярных значений процедура определения сенсоров и слабых мест, особенно для систем большой размерности, может оказаться весьма трудоемкой.

Другим недостатком технологии сингулярного анализа является невозможность одновременно с идентификацией сенсорных переменных оценить возможные диапазоны их изменения и вероятности выхода переменных за допустимые границы. Такая вероятность зависит от того, является ли переменная сенсорной и насколько близко ее текущее значение к предельному значению, каковы допустимые границы изменения переменной. В число контролируемых переменных при управлении в первую очередь нужно включить переменные, для которых вероятность выхода за допустимые границы максимальна, при этом возможна ситуация, когда сенсорная переменная не войдет в состав контролируемых параметров. Если вероятность нахождения переменной в допустимых границах ниже требуемого значения, возникает проблема выбора управляющих воздействий для увеличения такой вероятности.

Указанные сложности использования технологии сингулярного анализа заставили искать методы, менее трудоемкие, но не менее эффективные, чем сингулярный анализ. К таким методам относятся методы вероятностного потокораспределения, которые учитывают неопределенность задания исходной информации, внешние возмущения в них представляются случайным изменением нагрузок, а реакция ЭЭС на возмущения характеризуется числовыми характеристиками и функциями плотности вероятности, позволяющими оценить возможные границы изменения результатов потокораспределения и вероятности нахождения переменных в допустимых технологических границах.

Цель данной работы заключалась в выборе наиболее эффективного метода вероятностного потокораспределения для обнаружения сенсорных переменных, оценке их числовых характеристик, в определении вероятности

нахождения сенсорных переменных в допустимых границах и выборе управляющих воздействий, повышающих такую вероятность.

Подробный и глубокий обзор современных методов расчета вероятностного потокораспределения и их использования для решения различных задач электроэнергетики приводится в обзорах зарубежных работ [2] и российских работ [3, 10].

Основные российские работы, связанные с методами расчета вероятностного потокораспределения, включают методы: статистических испытаний, линейной и нелинейной аппроксимации, функционального преобразования.

Теоретические основы методов линейной и нелинейной аппроксимации представлены в [4, 5], а также в [6], где изложены основы метода статистической линеаризации. Эффективное практическое применение методов линейной и нелинейной аппроксимации для решения проблемы вероятностного потокораспределения в ЭЭС осуществлено в работах, перечисленных в [7].

Метод статистической линеаризации был впервые применен для расчета вероятностного потокораспределения в ЭЭС в работе [8] Для улучшения оценок моментов полученных методом статистической линеаризации в [9, 10] используется метод квадратичной аппроксимации и метод моментов.

В [11] метод квадратичной аппроксимации был усовершенствован за счет представления матрицы Гессе в прямоугольной форме, позволившей при вычислении ковариаций напряжений заменить операцию обращения матрицы решением систем уравнений.

Применение метода функциональных преобразований [3,4], являющегося точным аналитическим методом, базируется на общем принципе сравнения вероятностей. В [12] показана ограниченность применения этого метода, вызванная многомерностью совместной плотности распределения искомых переменных.

Линейный аналитический метод

Определение среднеквадратических отклонений (СКО) модулей и фаз узловых напряжений, в методе, который будем называть **линейным**, по заданным СКО узловых мощностей может быть получено с использованием выражения

$$\begin{pmatrix} \Delta\delta \\ \Delta U \end{pmatrix} = \begin{pmatrix} \dfrac{\partial P}{\partial \delta} & \dfrac{\partial P}{\partial U} \\ \dfrac{\partial Q}{\partial \delta} & \dfrac{\partial Q}{\partial U} \end{pmatrix}^{-1} \begin{pmatrix} \Delta P \\ \Delta Q \end{pmatrix} = J^{-1} \begin{pmatrix} \Delta P \\ \Delta Q \end{pmatrix}, \qquad (1)$$

связывающего в системе линеаризованных уравнений изменения фаз $\Delta\delta$ и модулей ΔU узловых напряжений с изменениями активных ΔP и реактивных ΔQ мощностей, J^{-1} – обратная матрица Якоби.

Математические ожидания $\mu_{\Delta\delta,\Delta U}$ и ковариации $\mu^2_{\Delta\delta,\Delta U}$ модулей и фаз напряжений определятся через математические ожидания $\mu_{\Delta P,\Delta Q}$ и дисперсии нагрузок $\mu^2_{\Delta P,\Delta Q}$ в точке решения нелинейной системы уравнений установившегося режима ЭЭС

$$\mu_{\Delta\delta,\Delta U} = J^{-1}\mu_{\Delta P,\Delta Q}, \qquad (2)$$

$$\mu_{2\Delta\delta,\Delta U} = J^{-1}\mu_{2\Delta P,\Delta Q}\left(J^{-1}\right)^T. \qquad (3)$$

Числовые характеристики нагрузок могут быть получены по статистической или прогнозной информации, при известной функции распределения нагрузок разложением ее характеристической функции в ряд Маклорена [5]. Предположение о нормальном законе распределения нагрузок позволяет определить их дисперсии с использованием функции Лапласа, называемой также функцией ошибок.

Для заданной вероятности P отклонения нормально распределенной случайной величины X от математического ожидания m на величину, не большую заданной точности $\Delta\varepsilon$,

$$P\left(\left|X-m\right|<\Delta\varepsilon\right)=\Phi\left(\Delta\varepsilon/\sigma\right). \tag{4}$$

может быть вычислено СКО σ. Значения $\Delta\varepsilon$ определяется погрешностью прогноза нагрузок или оценок нагрузок.

Более простое выражение для линейной модели (2), (3) может быть получено с использованием сингулярного разложения несимметричной матрицы Якоби

$$J=W\Sigma V^T=\sum_{\mathrm{j}=1}^{\mathrm{n}} w_j\,\sigma_j\,v_j^T, \tag{5}$$

где $W=\left(w_1,w_2,...,w_n\right)$ и $V=\left(v_1,v_2,...,v_n\right)$ – ортогональные матрицы, столбцы которых являются левым и правым сингулярными векторами, а Σ – диагональная матрица упорядоченных по возрастанию сингулярных значений $\sigma_1<\sigma_2<\sigma_3<...<\sigma_n$.

С учетом разложения (5) выражение (1) может быть представлено в виде

$$\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}=J^{-1}\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}=\sum_{i=1}^{n}v_i\left(w_i^T/\sigma_i\right)\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}=\sum_{i=1}^{n}v_i\,\Delta S^{(i)}\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}. \tag{6}$$

Если первое сингулярное значение $\sigma_1=\sigma_{min}$ существенно меньше остальных сингулярных значений, то, наибольший вклад в изменение фаз и модулей узловых напряжений вносит первое слагаемое суммы (6)

$$\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}^{(1)}=v_1\left(w_1^T/\sigma_1\right)\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}=v_1\,\Delta S^{(1)}, \tag{7}$$

где компоненты первого правого сингулярного вектора распределяют скалярную величину $\Delta S^{(1)}$ первого обобщенного возмущения между узлами сети

$$\Delta S^{(1)}=\left(w_1^T/\sigma_1\right)\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}. \tag{8}$$

Математические ожидания и ковариаций модулей и фаз напряжений, с учетом первого обобщенного возмущения $\Delta S^{(1)}$ могут быть выражены через

скалярные значения математического ожидания и дисперсии этого возмущения как

$$\mu_{\Delta\delta,\Delta U} = \frac{\nu_1 w_1^T}{\sigma_1}\mu_{\Delta P,\Delta Q} = \nu_1 \mu_{\Delta S_1}, \tag{9}$$

$$\mu_{2\Delta\delta,\Delta U} = \frac{\nu_1 w_1^T}{\sigma_1}\mu_{2\Delta P,\Delta Q}\left(\frac{\nu_1 w_1^T}{\sigma_1}\right)^T = \nu_1 \mu_{2\Delta S_1} \nu_1^T . \tag{10}$$

Количество вариантов возмущений бесконечно, они могут отличаться и по составу, и по величине возмущения. Такой подход не требует задания сценария изменения узловых мощностей, а позволяет по заданной величине обобщенного возмущения оценить множество сценариев возмущений по одному критерию. При этом необходимо только определить разумный диапазон изменения числовых характеристик обобщенного возмущения.

Выражения для числовых характеристик переменных, аналогичные (2), (3), (9), (10), могут быть записаны и для других переменных режима, таких как перетоки активной и реактивной мощности, разности модулей и фаз узловых напряжений. Найденные на основе этих выражений максимальные СКО позволяют не только выделить сенсорные переменные, но и оценить диапазоны их изменения.

Нелинейный аналитический метод

Сравнение СКО переменных режима, полученных на основе линейного аналитического подхода и метода Монте-Карло показало, что во многих случаях метод Монте-Карло дает большие значения СКО, чем линейный аналитический метод.

Это может сильно повлиять на заключение об адекватности полученных результатов, таких как вероятность нахождения контролируемых переменных в допустимой области. Снижение ошибки, связанной с линеаризацией уравнений потокораспределения может быть получено с использованием квадратичного вероятностного потокораспределения.

Квадратичная аппроксимация Тейлора уравнений установившегося режима в общем виде может быть представлена в виде

$$\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix}= J\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}+\frac{1}{2}\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}^T H\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}, \tag{11}$$

где кубическая матрица H, размера k^3, называемая матрицей Гессе, состоит из k слоев.

В [13] предложено применять другую форму матрицы H, заключающуюся в записи ее слоев в виде прямоугольной матрицы с k строками и k^2 столбцами

$$H=\begin{pmatrix}\frac{\partial^2 P}{\partial\delta\partial\delta} & \frac{\partial^2 P}{\partial\delta\partial U} & \frac{\partial^2 P}{\partial U\partial\delta} & \frac{\partial^2 P}{\partial U\partial U}\\ \frac{\partial^2 Q}{\partial\delta\partial\delta} & \frac{\partial^2 Q}{\partial\delta\partial U} & \frac{\partial^2 Q}{\partial U\partial\delta} & \frac{\partial^2 Q}{\partial U\partial U}\end{pmatrix}, \tag{12}$$

которая позволяет записать (11) в виде

$$\Delta Y = J\Delta X + \frac{1}{2}H\Delta X\otimes\Delta X, \tag{13}$$

$$\Delta Y=\begin{pmatrix}\Delta P\\ \Delta Q\end{pmatrix},\ \Delta X=\begin{pmatrix}\Delta\delta\\ \Delta U\end{pmatrix}.$$

Связь математических ожиданий и дисперсий узловых мощностей с математическими ожиданиями и ковариациями параметров состояния на основе (13) может быть представлена как

$$\mu_{\Delta Y} = J\mu_{\Delta X} + \frac{1}{2}Hb_{\mu_{2\Delta X}}, \tag{14}$$

$$\mu_{2\Delta Y} = J\mu_{2\Delta X}J^{\mathrm{T}} + \frac{1}{2}J\mu_{3\Delta X}H^T + \frac{1}{2}H\mu_{3\Delta X}^T J^T + \frac{1}{4}H\left(\mu_{4\Delta X} - b_{\mu_{2\Delta X}}b_{\mu_{2\Delta X}}^T\right)H^T, \tag{15}$$

где $\mu_{3\Delta X}$ и $\mu_{4\Delta X}$ – матрицы совместных центральных моментов третьего и четвертого порядков, размеров $\left(k^2\times k\right)$ и $\left(k^2\times k^2\right)$ соответственно, $b_{\mu_{2\Delta X}}$ – вектор, составленный из k столбцов, матрицы $\mu_{2\Delta X}$.

Система (14), (15) недоопределенная, поскольку в два ее уравнения входят четыре неизвестные матрицы моментов, то для получения единственного решения можно использовать различные формы записи уравнения (15).

В предложенном в [8] методе **статистической линеаризации** уравнение для ковариаций включает только первое слагаемое

$$\mu_{2\Delta Y} = J\mu_{2\Delta X}J^{\mathrm{T}}. \qquad (16)$$

В работе [9] разработан еще один подход к получению единственного решения системы (14), (15), в котором моменты третьего порядка полагают равными нулю, а моменты четвертого порядка учитываются введением коэффициента a, что позволяет представить уравнение (15) как

$$\mu_{2\Delta Y} = J\mu_{2\Delta X}J^{\mathrm{T}} + \frac{1}{4}aH\ b_{\mu_{2\Delta X}}b^{T}_{\mu_{2\Delta X}}H^{T}, \qquad (17)$$

где значение коэффициента a определяется видом плотности распределения.

С тех пор, когда впервые была предложена запись уравнений нелинейного вероятностного потокораспределения в виде (14), (15), существенно возросли возможности вычислительной техники и с точки зрения объемов оперативной памяти, и быстродействия, и способов программирования сложных выражений, позволившие использовать другие представления выражения (15).

Предположение о близости закона распределения переменных режима к нормальному закону позволяет выразить моменты третьего и четвертого порядков через равные нулю кумулянты.

Такой метод назван методом **двух моментов,** уравнение (15) в нем может быть записано как

$$\mu_{2\Delta Y} = J\mu_{2\Delta X}J^{\mathrm{T}} + \frac{1}{4}H\left(\mu_{4\Delta X} - b_{\mu_{2\Delta X}}b^{T}_{\mu_{2\Delta X}}\right)H^{T}. \qquad (18)$$

Итерационный процесс получения решения (14), (18) в общем виде может быть представлен следующим образом.

Задаются математические ожидания и дисперсии нагрузок, а также исходные приближения математических ожиданий параметров состояния.

Формируются матрицы Якоби и Гессе. Из системы (14) определяются $\mu_{\Delta\delta,\Delta U}$, а в соответствии с выражением (3) − исходное приближение матрицы $\mu_{2\Delta\delta,\Delta U}$ и вычисляемой на ее основе матрицы $\mu_{2\Delta\delta,\Delta U}$. Далее из (18) находится уточненная матрица $\mu_{2\Delta Y}$. Итерационный процесс повторяется до тех пор, пока небаланс в системе (14), (18) не станет меньше заданного значения.

Для дальнейшего повышения точности решения задачи вероятностного нелинейного потокораспределения можно использовать **метод трех моментов**, в котором систему (14), (15) предлагается дополнить уравнением для центральных моментов третьего порядка

$$\begin{aligned}\mu_{3\Delta Y} = {} & J\otimes J\mu_{3\Delta X}J^{\mathrm{T}} + \frac{1}{2}H\otimes J\left(\mu'_{4\Delta X} - b_{\mu_{2\Delta X}}\otimes b_{\mu_{2\Delta X}}\right)J^{\mathrm{T}} + \frac{1}{2}J\otimes H\left(\mu'_{4\Delta X} - \mu_{2\Delta X}\otimes b_{\mu_{2\Delta X}}\right)J^{\mathrm{T}} + \\ & + \frac{1}{2}J\otimes J\left(\mu_{4\Delta X} - b_{\mu_{2\Delta X}}b^{T}_{\mu_{2\Delta X}}\right)H^{\mathrm{T}} + \frac{1}{4}H\otimes H\left(\mu'_{5\Delta \mathrm{X}} - b_{\mu_{2\Delta X}}\otimes\mu_{3\Delta X} - \mu_{3\Delta X}\otimes b_{\mu_{2\Delta X}}\right)J^{\mathrm{T}} + \\ & \frac{1}{4}H\otimes J\left(\mu_{5\Delta X} - b_{\mu_{2\Delta X}}\otimes\mu^{\mathrm{T}}_{3\Delta X} - b_{\mu_{3\Delta X}}b^{T}_{\mu_{2\Delta X}}\right)H^{\mathrm{T}} + \frac{1}{4}J\otimes H\left(\mu_{5\Delta X} - \mu^{\mathrm{T}}_{3\Delta X}\otimes b_{\mu_{2\Delta X}} - b_{\mu_{3\Delta X}}b^{T}_{\mu_{2\Delta X}}\right)H\end{aligned} \qquad (19)$$

где $\mu'_{4\Delta X}$, $\mu_{5\Delta X}$, $\mu'_{5\Delta X}$, $\mu'_{6\Delta X}$ − матрицы совместных центральных моментов четвертого, пятого и шестого порядков, имеющие размеры $\left(k^3\times k\right)$, $\left(k^3\times k^2\right)$, $\left(k^4\times k\right)$, $\left(k^4\times k^2\right)$; $b_{\mu_{3\Delta X}}$, $b_{\mu_{4\Delta X}}$ – векторы, составленные из элементов столбцов матриц центральных моментов $\mu_{3\Delta X}$ и $\mu_{4\Delta X}$ Элементы матриц $\mu'_{4\Delta X}$, $\mu'_{5\Delta X}$, $\mu'_{6\Delta X}$ и $\mu_{4\Delta X}$, $\mu_{5\Delta X}$, $\mu_{6\Delta X}$ одинаковые, но расположены по-разному.

Система трех уравнений (14), (15), (19) из-за наличия в ней неизвестных матриц моментов четвертого, пятого и шестого порядков является недоопределенной. Для получения единственного решения центральные моменты четвертого, пятого и шестого порядков представляются через кумулянты, которые полагаются равными нулю. Алгоритм итерационного решения задачи вероятностного потокораспределения методом трех моментов аналогичен алгоритму для метода двух моментов.

В [14] предложен наименее трудоемкий, названный здесь **безитерационным,** метод расчета вероятностного нелинейного потокораспределения, не предполагающий проведения итерационного уточнения решения, а позво-

ляющий только уточнить математические ожидания и моменты второго порядка, полученные на основе линейной аппроксимации, с учетом матрицы Гессе. Слои матрицы Гессе в этом методе записываются один под другим, что позволяет представить (11) в виде

$$\Delta Y = J\Delta X + \frac{1}{2}\begin{pmatrix} \Delta X^T H_1 \Delta X \\ \Delta X^T H_2 \Delta X \\ \dots \\ \Delta X H_k \Delta x \end{pmatrix}, \tag{20}$$

А решение относительно вектора состояния как

$$\Delta X = J^{-1}\Delta Y - \frac{1}{2}J^{-1}\begin{pmatrix} \Delta Y^T \left(J^{-1}\right)^T H_1 J^{-1}\Delta Y \\ \Delta Y^T \left(J^{-1}\right)^T H_2 J^{-1}\Delta Y \\ \dots \\ \Delta Y \left(J^{-1}\right)^T H_k J^{-1}\Delta Y \end{pmatrix} = \Delta X' - \Delta X'', \tag{21}$$

где второй член решения $\Delta X''$ корректирует переменные $\Delta X'$, полученные на основе линейной аппроксимации. На основе (21) могут быть получены очевидные выражения для математических ожиданий $\mu_{\Delta X}$ и ковариаций $\mu_{2\Delta X}$.

Повышение точности вероятностных оценок метода [15] может быть достигнуто при представлении матрицы Гессе в форме (12), позволяющей переписать (21) в виде

$$\Delta X = J^{-1}\Delta Y - \frac{1}{2}J^{-1}H\left(J^{-1}\otimes J^{-1}\right)\left(\Delta Y \otimes \Delta Y\right), \tag{22}$$

и включением в алгоритм процедуры итерационного уточнения решения.

Математическая формулировка такого, названного **модернизированным,** метода вероятностного потокораспределения, записанная для трех моментов, будет иметь вид

$$\mu_{\Delta X} = A\mu_{\Delta Y} + B\mu_{2\Delta Y}, \tag{23}$$

$$\mu_{2\Delta X} = A\mu_{2\Delta Y}A^T + A\mu_{3\Delta Y}B^T + B\left(\mu_{3\Delta Y}\right)^T A^T + B\mu_{4\Delta Y}B^T, \tag{24}$$

$$\begin{aligned}\mu_{3\Delta X} = {} & A\otimes A\mu_{3\Delta Y}A^T + A\otimes A\mu_{4\Delta Y}A^T + A\otimes B\mu'_{4\Delta Y}A^T + B\otimes A\mu'_{4\Delta Y}A^T + A\otimes B\mu_{5\Delta Y}B^T + \\ & + B\otimes A\mu_{5\Delta Y}B^T + B\otimes B\mu'_{5\Delta Y}A^T + B\otimes B\mu'_{6\Delta Y}B^T),\end{aligned} \tag{25}$$

где $A = J^{-1}$, $B = -\frac{1}{2} J^{-1} H\left(J^{-1} \otimes J^{-1}\right)$.

Точность вероятностных оценок на основе линейного и нелинейных методов, может быть оценена при их сравнении с оценками, полученными численным Монте-Карло. Этот метод, признан наиболее точным и является тестовым при оценке точности упрощенных методов вероятностного потокораспределения. Основным недостатком этого метода является его трудоемкость.

Сравнение методов вероятностного потокораспределения

В качестве тестовой схемы при сравнении методов вероятностного потокораспределения использовалась приведенная на рис. 1. тестовая схема ЭЭС, содержащая 14 узлов и 15 связей.

Исходная информация о математических ожиданиях, дисперсиях и моментах более высоких порядков для нагрузок, которые задавались во всех узлах расчетной схемы, была получена на основе датчика случайных чисел для нормального распределения.

Для того, чтобы отстроиться от нормального закона распределения нагрузок было проведено ограниченное число испытаний. В качестве исходной информации при формировании случайных чисел выступали математические ожидания узловых мощностей и СКО этих мощностей, которые полагались равными 12 % от их математических ожиданий, что соответствует 20 % погрешности прогноза нагрузок для вероятности отклонения случайной величины от математического ожидания, равной 0.9. О достижении поставленной цели свидетельствовали как моменты высших порядков, так и кривые плотностей вероятности для нагрузок, рис. 2, построенные на их основе, с использованием разложения Грама-Шарлье.

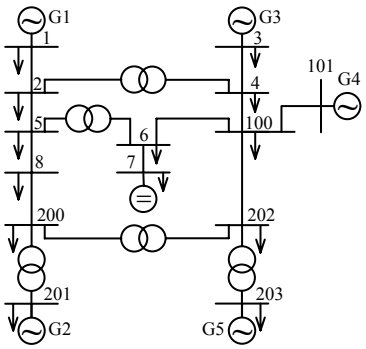


Рис. 1. Схема 14 узловой тестовой сети

Аппроксимируемая плотность распределения $f(x)$ переменной x рядом Грама-Шарлье [15] может быть представлена в виде

$$f(x)=\sum_{j=0}^{\infty} c_j\, H_j(x)\varphi(x), \tag{26}$$

где функция $\varphi(x)=\frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$ является функцией нормального распределения, $H_j(x)$ – ортогональные полиномы Эрмита, а c_j – коэффициенты, построенные на основе моментов.

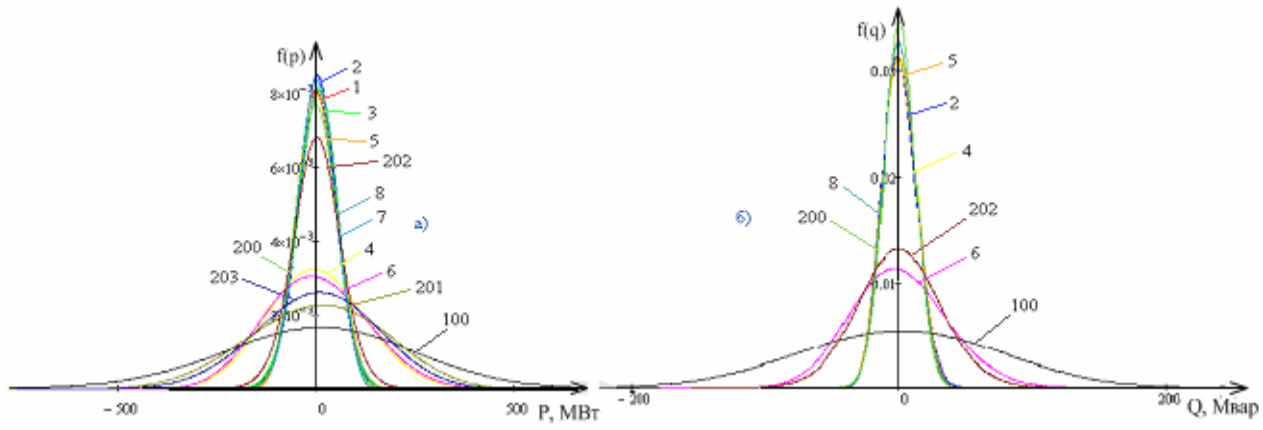


Рис. 2. Кривые плотностей вероятности для изменений активных а) и реактивных б) узловых мощностей в узлах тестовой схемы, полученные по четырем моментам с использованием разложения Грама-Шарлье

Сравнение методов вероятностного потокораспределения проиллюстрировано на примере модулей узловых напряжений.

На рис. 3 а приведены полученные на основе соответствующих ковариационных матриц СКО изменений модулей напряжений в узлах тестовой

схемы для: двух линейных методов, пяти нелинейных методов и метода Монте-Карло. Все методы выделили узел 8, как наиболее сенсорный.

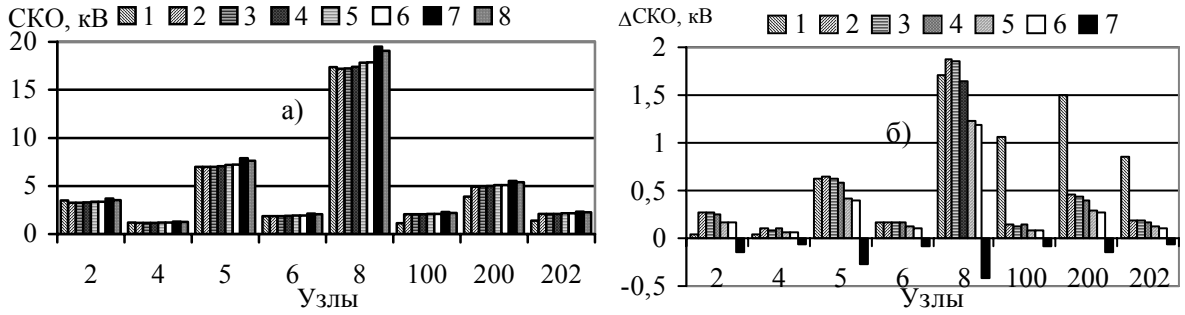


Рис. 3 СКО изменений модулей узловых напряжений а) полученные на основе методов: 1 – обобщенного возмущения, 2 – линейного, 3 – безитерационного, 4 – двух моментов, 5 – статистической линеаризации, 6- модернизированного, 7 – трех моментов, 8 – Монте-Карло и разность СКО б) для метода 8 и методов 1 – 7

На рис. 3 б показаны разности СКО, полученных для метода Монте-Карло, и СКО для линейных и нелинейных методов, позволившие сделать следующие выводы:

- для всех методов максимальная разность СКО по отношению к методу Монте-Карло отмечается в 8 узле с сенсорным модулем напряжения, этот же узел определяется как сенсорный и на основе сингулярного анализа [1];
- СКО, полученные линейным, безитерационным методами и методом обобщенного возмущения близки между собой;
- СКО на основе безитерационного метода, незначительно отличаются от отклонений для линейного метода, что связано небольшими значениями добавок $\Delta X''$ к $\Delta X'$, σ;
- СКО для метода трех моментов максимально приближаются к СКО по методу Монте-Карло, но в отличие от других методов, превышают их;
- следующими после метода трех моментов, имеющими СКО, близкие к СКО по методу Монте-Карло, являются метод статистической линеаризации, модифицированный метод и метод двух моментов.

Преимущество метода трех моментов с точки зрения точности получаемого решения следует и из сравнения кривых плотностей вероятности для модуля напряжения в 8 сенсорном узле, построенных для шести аналитических методов и метода Монте-Карло.

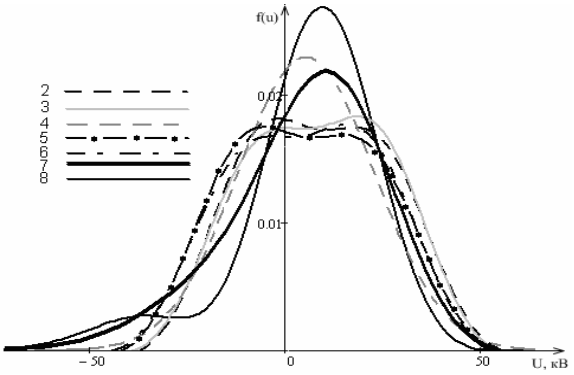


Рис. 4. Кривые плотностей вероятности модуля напряжения в сенсорном узле 8, построенные на основе разложения Грама-Шарлье (номера методов те же, что на рис.3)

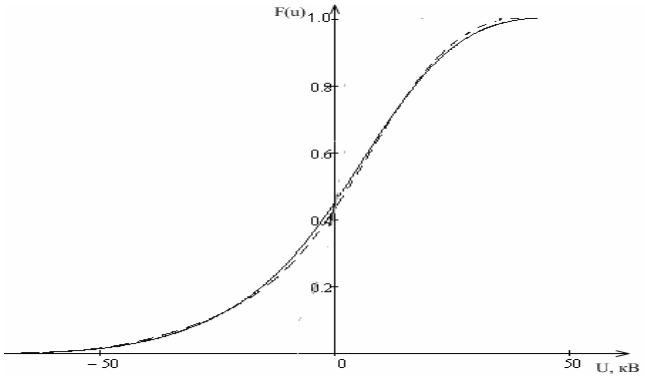


Рис. 5. Графики функций распределения модуля напряжения в 8 узле, построенные на основе разложения Грама-Шарлье для метода трех моментов (сплошная линия) и метода Монте-Карло (пунктирная линия)

Из рис. 4 видно, что кривая плотности вероятности для метода трех моментов наиболее близка к тестовой кривой плотности, полученной для метода Монте-Карло.

Сравнение графиков функций распределения для модуля напряжения в 8 узле для всех аналитических методов с графиком для метода Монте-Карло показало, что наиболее близки между собой функции распределения для метода трех моментов и метода Монте-Карло, что показано на рис. 5.

Вероятностное потокораспределение с учетом ограничений

Если в результате расчета вероятностного потокораспределения окажется, что вероятность нахождения контролируемых переменных, к которым в первую очередь относятся сенсорные переменные, в допустимых границах, ниже требуемой, то для увеличения такой вероятности осуществляется выбор управляющих воздействий.

Для решения указанной проблемы предлагается использовать метод, аналогичный методу детерминированного эквивалента [16], заключающемуся в последовательном итерационном решении детерминированной и вероятностной задач. Однако процедура ввода контролируемых параметров в допустимую область отличается от процедуры, используемой в методе детерминированного эквивалента, и заключается не в сужении допустимого интервала для случайного контролируемого параметра, а в перемещении математического ожидания в точку, являющуюся медианой его плотности распределения, усеченной ограничениями [17, 18].

В детерминированной задаче расчета потокораспределения с учетом ограничений для обеспечения требуемой вероятности нахождения контролируемых переменных в допустимых пределах должны быть выработаны соответствующие управляющие воздействия. Среди них для каждой контролируемой переменной с низкой вероятностью нахождения в допустимой области, надо найти управления, к которым она является наиболее чувствительной.

Другими словами, минимальное изменение управляющего параметра должно привести к максимальному изменению контролируемой переменной, позволяющему увеличить вероятность ее нахождения в допустимой области. Такая задача успешно решается при сочетании методов приведенного градиента и задачи квадратичного программирования, решаемой на каждом шаге последовательной линеаризации [19].

Проанализируем проблему определения желаемых вероятностных характеристик контролируемых переменных, обеспечивающих необходимую вероятность их нахождения в допустимой области. Будем считать, что в результате расчета вероятностного потокораспределения для контролируемой сенсорной переменной g найдены ее математическое ожидание μ_g, СКО σ_g, моменты до четвертого порядка $\mu_{2\,g}$, $\mu_{3\,g}$, $\mu_{4\,g}$ и известен допустимый интервал ее изменения $(g_{\min,}\ g_{\max})$.

Вероятность попадания переменной в допустимый интервал может быть определена либо по информации о математическом ожидании и СКО, либо на основе разложения Грама-Шарлье.

Если вероятность нахождения переменной в заданном интервале ниже, чем требуемое значение вероятности, то существует две возможности ее увеличения. Первая заключается в поиске подходов к снижению СКО, что может быть достигнуто, например усилением слабых мест или выбором управлений, приводящих к уменьшению потери напряжения в связи. И вторая может быть достигнута совмещением медианы m_c усеченной ограничениями кривой плотности распределения со значением математического ожидания.

Вторая возможность наиболее очевидна для нормального распределения, поскольку ее кривая плотности вероятности симметричная, то максимальная вероятность попадания контролируемой переменной в интервал будет при совмещении ее математического ожидания с медианой, расположенной в центре допустимого интервала

$$m_c = \left(g_{min} + g_{max}\right)/2. \qquad (27)$$

Если кривая плотности распределения получена для нескольких моментов с использованием разложения Грама-Шарлье, то задача поиска ее рационального смещения усложняется.

В результате реализации управляющего воздействия, связанного со смещением математического ожидания контролируемой переменной, происходит изменение и других переменных, что может привести к увеличению СКО контролируемой переменной, а, следовательно, не к увеличению, а к снижению вероятности ее попадания в допустимую область. Кроме того, могут увеличиться вероятности выхода и других контролируемых переменных за допустимые границы.

Основные этапы алгоритма заключаются в следующем.

1. Расчет установившегося режима ЭЭС. Индекс итерации $k = 0$.

2. Расчет вероятностного потокораспределения, включающего определенных числовых характеристик контролируемых параметров, выделение сенсорных контролируемых переменных, в частности, $\mu_{g_i}^k$, $i \in I_c$, где I_c – множество индексов сенсорных переменных.

3. Определение вероятности нахождения сенсорных переменных в допустимых границах. Завершение работы алгоритма, если требуемая вероятность для всех сенсоров обеспечена. Если нет, то определение для каждого сенсора $j \in I_{cv}$, где $I_{cv} \subset I_c$ – множество индексов сенсорных переменных, для которого заданная вероятность соблюдения ограничения $(g_{j\min}, g_{j\max})$ не выполняется, и смещения его математического ожидания $\Delta_j^k = m_{cj} - \mu_{g_j}^k$.

4. Определение из решения детерминированной задачи вектора управляющего воздействия $\Delta Y = Y - Y^k$, при котором сенсорные переменные $j \in I_{cv}$ принимают значения $\mu_{g_j}^k + \Delta_j^k$. Если такое решение не может быть найдено, то ищется решение, при котором вероятность нахождения в допустимых границах будет наибольшей.

Целевая функция детерминированной оптимизационной задачи может быть записана в виде выражения

$$\min_Y \sum_{j=1}^{I_{cv}} \left[g_j(Y) - \left(\mu_{g_j}^k + \Delta_j^k \right) \right]^2, \qquad (28)$$

которое должно быть дополнено ограничениями на контролируемые переменные и параметры управления.

Для тестовой сети в табл. 1 приведены Математические ожидания и СКО модулей напряжений в узлах тестовой сети, полученные линейным методом, приведены в табл. 1. Там же записаны принятые допустимые диапазоны изменения напряжений и вероятности попадания модулей напряжения в допустимые интервалы, вычисленные на основе μ_g и σ_g.

Таблица 1

Вероятностные характеристики модулей напряжений в узлах тестовой схемы для исходного режима

Узлы	m_U (кВ)	σ_U	$\underline{\Delta U}$	$\Delta\overline{U}$	P
2	522.340	3.482	–30	30	0.9861
4	231.497	1.201	–25	25	1.0000
5	512.056	6.972	–30	30	0.9950
6	225.172	1.874	–25	20	1.0000
8	508.447	17.203	–30	30	0.8822
100	229.241	2.054	–25	25	1.0000
200	528.159	4.935	–30	30	0.6454
202	233.627	2.082	–25	25	1.0000

Несмотря на то, что модуль напряжения 8-го узла является более сенсорным, чем модуль напряжения 200-го узла, что следует из сравнения их СКО, разность математического ожидания и номинального напряжения 500кВ для 200-го узла существенно выше, чем для 8-го узла.

Таблица 2

Вероятности нахождения модулей напряжений в допустимых границах, полученные различными методами для исходного режима

Узлы	Методы вероятностного потокораспределения*						
	2	3	4	5	6	7	8
2	0.991	0.991	0.995	0.994	0.994	0.99	0.991
4,6,100,202	1	1	1	1	1	1	1
5	0.995	0.995	0.997	0.996	0.996	0.994	0.994
8	0.882	0.882	0.901	0.893	0.893	0.868	0.87
200	0.645	0.645	0.708	0.705	0.704	0.697	0.691

* номера методов те же, что на рис. 3

Последнее свойство является определяющим в том, что вероятность нахождения модуля напряжения в допустимых границах для 200 узла, равная 0.6454, ниже вероятности для 8 сенсорного узла.

Таблица 3

Вероятности нахождения модулей напряжений в допустимых границах, полученные различными методами в конечной точке

Узлы	Методы вероятностного потокораспределения						
	2	3	4	5	6	7	8
2, 4, 6, 200, 202	1	1	1	1	1	1	1
5	0.999	0.999	0.999	0.999	0.999	0.998	0.998
8	0.928	0.928	0.92	0.913	0.913	0.889	0.89
100	0.987	0.987	0.993	0.991	0.991	0.987	0.988

Такие же низкие вероятности для 200 узла получены и при использовании дополнительно к линейному методу пяти нелинейных методов и метода Монте-Карло, табл. 2.

В табл. 3 приведены значения вероятностей, полученные в результате последовательности управляющих воздействий. Достигнутое значение вероятности нахождения напряжения в сенсорном узле 8 в допустимых границах для всех методов составило в среднем 0.9.

Влияние увеличения проводимости слабой связи 100–202 [1] на уменьшение СКО в узлах тестовой схемы проиллюстрировано для линейного метода на рис. 6.

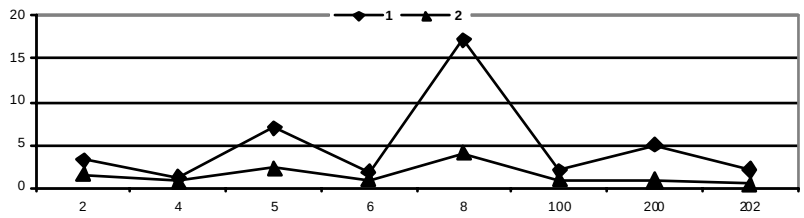


Рис. 6. Среднеквадратические отклонения изменений модулей узловых напряжений для линейного метода для исходной схемы (1) и при увеличении проводимости слабой связи 100-202 (2)

Выводы

1. Методы вероятностного потокораспределения позволяют обнаружить те же сенсорные переменные в ЭЭС, которые могут быть выделены на основе сингулярного анализа.
2. Предложено использовать сочетание аналитического вероятностного метода и скалярной величины первого обобщенного возмущения для получения вероятностных показателей переменных в неоднородной сети.
3. Предложены модификации методов вероятностного нелинейного потокораспределения, включающие методы двух и трех моментов с использованием кумулянтов и модификация безитерационного метода, заключающаяся в коррекции матриц Якоби и Гессе в процессе итераций.
4. Проведено экспериментальное сравнение аналитических методов вероятностного потокораспределения, показавшее несомненное преимущество метода трех моментов по точности получаемого решения по сравнению с другими методами.
5. Предложен подход для решения проблемы выбора управляющих воздействий, обеспечивающих требуемую вероятность нахождения контролируемых параметров в допустимых границах.

Список литературы

[1] Войтов О.Н., Воропай Н.И., Гамм А.З. и др. Анализ неоднородностей электроэнергетических систем − Новосибирск: Наука. Сибирская издательская фирма РАН, 1999. − 256 с.

[2] Chen P., Chen Z., Bak-Jensen B. Probabilistic load flow: A review // Electric Utility Deregulation and Restructuring and Power Technologies. Third International Conference. – 2008. – Pp. 1586 – 1591.

[3] Крумм Л.А. Методы оптимизации при управлении электроэнергетическими системами − Новосибирск: Наука. Сибирское отделение, 1981. – 317 с.

[4] Вентцель Е.С. Теория вероятностей. – М.: Наука, 1969. – 576 с.

[5] Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.

[6] Казаков И.Е., Доступов Б.Г. Статистическая динамика нелинейных автоматических систем – М.: Физматгиз, 1962. – 332 с.

[7] Контарева Е.Б., Пусеп П.А. Библиографический указатель публикаций профессора Манусова В.З. [Электронный ресурс] // Новосибирский государственный технический университет. – 2001. – URL:http://rudocs.exdat.com/docs/index-163024.html (дата обращения 24.08.11).

[8] Манусов В.З., Лыкин А.В. Вероятностный анализ установившихся режимов электрических систем // Электричество. – 1981. – № 4. – С. 7–13.

[9] Манусов В.З., Шепилов О.Н. Использование вероятностных свойств ретроспективной диспетчерской информации для планирования нормальных режимов ЭЭС // Алгоритмы обработки данных в электроэнергетике. – Иркутск: СЭИ СО АН СССР, 1982. – С. 163 – 170.

[10] Манусов В.З., Могирев В.В., Шепилов О.Н. Исследование режимов ЕЭС СССР с учётом случайного характера исходной информации // Электричество. – 1983. – № 10. – С. 3 – 6.

[11] Кучеров Ю.Н. Усовершенствование аналитических методов вероятностного анализа установившихся режимов электрических систем в предельных условиях // Изв. СО АН СССР. Сер. техн. Наук. – 1986. – Вып. 2, № 10. – С. 111 – 117.

[12] Болоев Е.В. Проверка математических моделей электроэнергетических систем, предназначенных для выполнения расчетов установившихся режимов в вероятностно-определенных условиях // Современные технологии. Системный анализ. Моделирование. – 2010. – № 3 (27). – С. 202 – 208.

[13] Iwamoto S., Tamure Y.A. A load flow calculation method for IEE-condition power system // IEEE Trans. Power Apparatus and systems. – 1981. – Vol. PAS-101. No.4. – Pp. 3261 – 3268.

[14] Li X., Chen X., Yin X., Xiang T., Liu H. The Algorithm of Probabilistic Load Flow Retaining Nonlinearity // Proceedings of 2002 Power Con, Int. Conf. on Power System Technology, Kunming. – 2002. – Vol. 4. – Pp. 2111 – 2115.

[15] Федорченко В.А. Теория многомерных распределений. – М.: Русь, 2003. – 576 с.

[16] Валдма М.Х., Крумм Л.А., Охорзин Ю.А. Методы решения стохастических задач комплексной оптимизации режимов сложных ЭЭС // Фактор неопределенности при принятии решений в больших системах энергетки. – Иркутск: СЭИ СО АН СССР, 1974. – С. 96 – 111.

[17] Болоев Е.В. Вероятностный расчет допустимого режима ЭЭС // Системные исследования в электроэнергетике: труды молодых ученых ИСЭМ СО РАН. – Иркутск: ИСЭМ СО РАН, 2003. – вып. 33. – С. 15 – 21.

[18] Гамм А.З. Вероятностные модели режимов электроэнергетических систем – Новосибирск: Наука. Сибирская издательская фирма, 1993. – 133 с.

[19] Мурашко Н.А., Охорзин Ю.А., Крумм Л.А. и др. Анализ и управление установившимися состояниями электроэнергетических систем. – Новосибирск: Наука, 1987. – 417 с.

УДК 519.856.3: 63

МОДЕЛИ ОПТИМИЗАЦИИ ВЗАИМОДЕЙСТВИЯ УЧАСТНИКОВ В АГРОПРОМЫШЛЕННЫХ КЛАСТЕРАХ С ВЕРОЯТНОСТНЫМИ ПАРАМЕТРАМИ

Бузина Т. С.

Иркутская государственная сельскохозяйственная академия

E-mail: buzinats@mail.ru

Иваньо Я. М.

Иркутская государственная сельскохозяйственная академия

E-mail: iymex@rambler.ru

Аннотация. В статье рассматривается агропромышленный кластер, как сложная система, характеризуемая множеством параметров, часть из которых является детерминированной, а другая – неопределенной. Для описания взаимодействия участников агропромышленных кластеров в регионе предлагаются две группы моделей: с учетом и без учета влияния природных событий, к которым относятся засухи, дождевые паводки, весенние половодья, интенсивные ливни, ураганы и др. Поставлены и решены задачи оптимизации взаимодействия участников зерновых кластеров с вероятностными параметрами, в качестве которых использованы продуктивность сельскохозяйственных угодий и цены на продукцию. Получены вероятностные распределения целевых функций и различные значения неизвестных моделей в зависимости от степени влияния на производство природных событий.

Ключевые слова. Задачи математического программирования, вероятностные параметры, агропромышленные кластеры, природные события.

Введение

Одним из эффективных направлений развития сельскохозяйственного производства является формирование агропромышленных кластеров с определением оптимального взаимодействия его участников для стратегического планирования на основе моделей, увязывающих условия функционирования различных предприятий, действующих в едином цикле производства, переработки и реализации продукции. В зависимости от различных условий производства и реализации продукции на территории Иркутской области возможно выделение трех видов агропромышленных кластеров: молочные, мясные и зерновые.

В общем виде модель агропромышленного кластера описывает взаимодействие множества различных участников: трех категорий сельскохозяйственных товаропроизводителей, перерабатывающего предприятия, сбытовых, научных, банковских, страховых и других организаций.

Поскольку возникает необходимость описания всех участников объединения как единой системы, математическая модель регионального агропромышленного кластера может иметь блочный вид. Каждый блок представляет собой группу ограничений модели, описывающую сельскохозяйственное производство, переработку и сбыт продукции, включающих множество параметров, значения которых могут колебаться, что необходимо учитывать при решении практических задач.

При моделировании агропромышленных кластеров особенно распространенными являются ситуации, когда выбор решения осуществляется в условиях рисков: существует неопределенность в виде множества частных исходов результата принятия решения, причем вероятности появления этих исходов либо определяемы тем или иным способом, либо неизвестны или не имеют смысла [5].

При этом необходимо учитывать, что на устойчивость сельскохозяйственного производства в значительной мере оказывают влияние природно-климатические факторы. Ежегодные потери в Иркутской области от гидрометеорологических явлений, к которым относятся: засухи, паводки, половодья, заторы, зажоры, ливневые осадки, заморозки, раннее выпадение снежного покрова, исчисляются сотнями миллионов рублей. В ранжированном ряду природных стихий по нанесению ущербов сельскому хозяйству региона особое место занимают засухи, дождевые паводки и весенние половодья.

Перечисленные природные события влияют на ресурсы, цены, продуктивность сельскохозяйственных угодий и животных.

Подобная ситуация описывается задачей стохастического программирования, которая имеет вид:

$$f(x) = M\left(\sum_{j \in J} c_j x_j \right) \to \min(\max), \quad (1)$$

$$p\left\{ \sum_{j=1}^{n} a_{ij} x_{ij} \le b_i \right\} \ge \alpha_i \ (i \in I), \quad (2)$$

где $f(x)$ – целевая функция, x_j – искомая переменная, c_j – коэффициенты целевой функции, a_{ij}, b_i – параметры ограничений, a_i – заданная вероятность выполнения системы, p – вероятность выполнения каждого ограничения [3, 4, 7].

Так как представленная задача является сложной, возможным методом ее решения является переход к детерминированному эквиваленту. В основе этого перехода лежит использование закона распределения случайной величины. В практике при описании параметров моделей, отражающих производство, наиболее часто используется семейство нормальных законов распределения и гамма–распределение [6]. Для планирования производства продукции в агропромышленном кластере может быть использован вариант задачи (1) – (2), когда коэффициенты ограничений, целевой функции и правых частей условий c_j, a_{ij}, и b_i представляют собой случайные величины, связанные с вероятностью превышения *p*. При этом в распределениях параметров модели отражено влияние природно-экономических факторов. Между тем возможна непосредственная оценка влияния природных событий на производство.

Поэтому модели оптимизации взаимодействия участников агропромышленных кластеров можно классифицировать на две группы: с учетом и без учета влияния природных событий (рис. 1).

Рис. 1. Модели оптимизации с учетом и
без учета влияния природных событий

В моделях оптимизации взаимодействия участников в агропромышленных кластерах первой группы коэффициентами при неизвестных в целевой функции являются различные стоимостные характеристики, которые могут описываться законами распределения вероятностей. В этом случае критерий оптимальности представляет собой случайную величину, связанную с вероятностью превышения p, и может быть описан следующим образом:

$$f = \sum_{j=1}^{J} \sum_{k=1}^{K} c_{jk}^{p} x_{jk} \to \max, \qquad (3)$$

где j - виды продукции $(j = 1,2,...,J)$, k – индекс категории предприятий $(k = 1,2,...,K)$; x_{jk} – объем производства продукции; c_{jk}^{P} – прибыль от единицы j-вида продукции в k-категории предприятий, представляющая собой случайную величину, связанную с вероятностью превышения p.

Если же случайными величинами являются сельскохозяйственные угодья, поголовье животных и другие производственные ресурсы в кластере, то ограничения в общем виде примут следующий вид

$$\sum_{j=1}^{J} a_{ijk} x_{jk} \le (\ge)\, b_{ik}^{p} \quad (i = \overline{1,I},\ k = \overline{1,K}), \tag{4}$$

где i – виды производственных ресурсов $(i = 1,2,...,I)$; a_{ijk} – норма затрат ресурсов i–го вида на единицу j–го вида продукции в k–ой категории предприятий; b_{ik}^{p} – объем ресурсов i–го вида в k–ой категории предприятий, представляющий собой случайную величину, связанную с вероятностью превышения p.

Кроме того, различными законами распределения могут характеризоваться такие параметры модели, как продуктивность животных и сельскохозяйственных культур. В этом случае ограничения модели записываются следующим образом:

$$\sum_{j=1}^{J} a_{ijk}^{p} x_{jk} \le (\ge)\, b_{ik} \quad (i = \overline{1,I},\ k = \overline{1,K}), \tag{5}$$

где a_{ijk}^{p} – параметр модели, характеризующий урожайность сельскохозяйственных культур как случайную величину.

Для второй группы моделей в условиях появления природного события (непосредственное влияние на коэффициенты целевой функции) цены на продукцию повышаются за счет дополнительной составляющей γ_{jk}^{p}, компенсирующей потери в результате влияния экстремальных природных явлений. Если изменение цен можно описать с помощью закона распределения вероятностей, то целевая функция примет вид

$$f = \sum_{J=1}^{J} \sum_{k=1}^{K} (c_{jk} + \gamma_{jk}^{p}) x_{jk} \to \max. \tag{6}$$

В дополнение к этому в результате проявления природных событий значения описанных выше параметров ограничений модели уменьшаются за

счет слагаемых φ_{ijk}^{p} и ψ_{ik}^{p}, характеризующих потери продуктивности и ресурсов в кластере в результате проявления природных событий. Поэтому ограничения модели примут вид

$$\sum_{j=1}^{J}(a_{ijk} - \varphi_{ijk}^{p})x_{jk} \leq (\geq)\, b_{ik} - \psi_{ik}^{p}. \qquad (7)$$

Собственные исследования и работы многих авторов [2, 3] показывают, что в агропромышленных кластерах с помощью вероятностных параметров может быть описана изменчивость урожайности зерновых культур, закупочных цен на сельскохозяйственную продукцию и прибыли от реализации продукции.

Эти особенности информации использованы при создании модели оптимизации взаимодействия участников в зерновом кластере в виде блоков, представляющих собой сельскохозяйственные организации, крестьянские (фермерские) хозяйства и перерабатывающее предприятие.

В этой задаче критерием оптимальности является максимум прибыли от реализации продукции всеми категориями предприятий

$$f = \sum_{j\in J}\sum_{k\in K} c_{jk}x_{jk}^{III} + \sum_{v\in V} c_v x_v^{I} - \sum_{i\in I}\sum_{k\in K} c_{ik}^{'} x_{ik}^{IV} - \sum_{q\in Q} c_q^{'} x_q^{II}, \qquad (8)$$

где x_v^I - искомая переменная, характеризующая объемы реализованной продукции вида v, c_v – цена реализации продукции вида v, c_q – закупочная цена на сельскохозяйственную продукцию, подлежащую переработке, x_q^{II} – искомая переменная, соответствующая объему q–го вида продукции, подлежащей переработке, V и Q – множества видов конечной продукции кластера и продукции, подлежащей переработке, $c_q^{'}$ – приведенные затраты перерабатывающего предприятия, c_j – выручка от реализации продукции товаропроизводителями, $c_i^{'}$ – затраты на одну голову животных или на один гектар площади культур, x_j^{III} – объем производства сельскохозяйственной продукции вида j, x_i^{IV} – поголовье животных или площадь культур i–го

вида, I и J – множества видов производственных ресурсов и сельскохозяйственной продукции.

Условия модели описываются следующими выражениями.

Развитие отраслей в каждой категории предприятий ограничивается имеющимися и выделяемыми производственными ресурсами (посевные площади, численность поголовья животных и т.п.)

$$\sum_{j \in J} a_{ijk} x_{jk}^{III} \le A_{ik}, \ (i \in I, k \in K). \qquad (9)$$

Ограничения, связывающие между собой отрасль животноводства и растениеводства в каждой категории предприятий, имеют вид

$$\sum_{j \in J} a_{ijk} x_{jk}^{III} \ge \sum_{j \in J} b_{ijk} x_{jk}^{III}, \quad (i \in I, \ k \in K). \qquad (10)$$

Условия по учету ограниченности производственных мощностей перерабатывающего предприятия записываются следующим образом

$$\sum_{q \in Q} x_q^{II} \le W, \qquad (11)$$

где W – производственная мощность перерабатывающего предприятия.

Условия по развитию обслуживающих отраслей инфраструктуры имеют вид

$$\sum_{i \in I} \sum_{k \in K} g_{\omega ik} x_{ik}^{IV} + \sum_{q \in Q} g_{\omega q} x_q^{II} = x_\omega^V, \ (\omega \in W'), \qquad (12)$$

где x_ω^V – затраты труда в обслуживающих отраслях агропромышленного кластера; $g_{\omega ik}$ – нормативы потребности в услугах обслуживающей отрасли в расчете на единицу площади или вида животных в *k*–ой категории предприятий; $g_{\omega q}$ – трудовые затраты обслуживающих отраслей на единицу продукции *q*–го вида; ω – вид обслуживающих отраслей; W' – множество обслуживающих отраслей.

В модели необходимо учесть интересы всех участников кластера, к которым относятся три категории товаропроизводителей, перерабатывающие предприятия и сбытовые организации:

$$\sum_{i\in I} c_{ik} x_{ik}^{IV} - \sum_{i\in I} c'_{ik} x_{ik}^{IV} - \sum_{d\in D} c'_{dk} x_{dk}^{VI} \geq R_k \ \ (k \in K), \tag{13}$$

где c'_{dk} – себестоимость единицы d–го вида корма в k–ой категории предприятий, x_{dk}^{VI} – количество кормов из состава покупных кормов и побочной продукции в k-категории предприятий, R_k – прибыль от реализации продукции k –ой категории предприятий; D – множество видов кормов.

Для учета распределения инвестиций между отраслями введено ограничение по распределению основных фондов

$$\sum_{i\in I}\sum_{k\in K} \phi_{ik} x_{ik}^{IV} + \sum_{q\in Q} \phi_q x_q^{II} + \sum_{\omega\in W'} \phi_\omega x_\omega^{V} \leq F', \tag{14}$$

где $\phi_{ik}, \phi_{\mathrm{q}}, \phi_\omega$ – количество основных фондов, приходящихся на единицу площади или вида животных, единицу продукции q–го вида, единицу трудовых затрат обслуживающей отрасли, F' – общий объем основных фондов в агропромышленном кластере.

Все переменные модели должны быть неотрицательны

$$x_v^I, x_q^{II}, x_{jk}^{III}, x_{ik}^{IV}, x_\omega^V, x_{dk}^{VI} \geq 0. \tag{15}$$

К ограничениям, учитывающим вероятности p, относятся условия:

- по производству зерновой продукции не менее задаваемой потребности

$$\sum_{i\in I}\sum_{k\in K} \mu_{jik}^p x_{ik}^{IV} \geq M_j, \ (j \in J), \tag{16}$$

где μ_{jik}^p – урожайность зерновой культуры, как случайная величина, M_j – требуемый объем производства продукции j–го вида;

- по реализации продукции из всех категорий предприятий в количестве, не превышающем объемов ее производства

$$\sum_{i\in I}\sum_{k\in K} \mu_{jik}^p x_{ik}^{IV} \geq \sum_{k\in K} x_{vk}^I, \ (v \in V, j \in J), \tag{17}$$

- учету пропорциональности между производством и переработкой продукции

$$\sum_{i\in I}\sum_{k\in K} \mu_{jik}^p x_{ik}^{IV} - \sum_{j\in J}\sum_{k\in K} \eta_{jk} x_{jk}^{III} = x_q^{II}, \ (q \in Q). \tag{18}$$

Модель оптимизации взаимодействия участников зернового кластера, учитывающая вероятность в левых частях ограничений, описанная выражениями (8)-(18), реализована на примере Балаганско-Заларинского кластера. Для моделирования распределения урожайности зерновых использована функция нормального распределения вероятностей.

Решение задачи показывает, что для вероятности превышения, изменяющейся от 0,1 до 0,9, прибыль находится в пределах 97 – 152 млн. руб. (рис. 2).

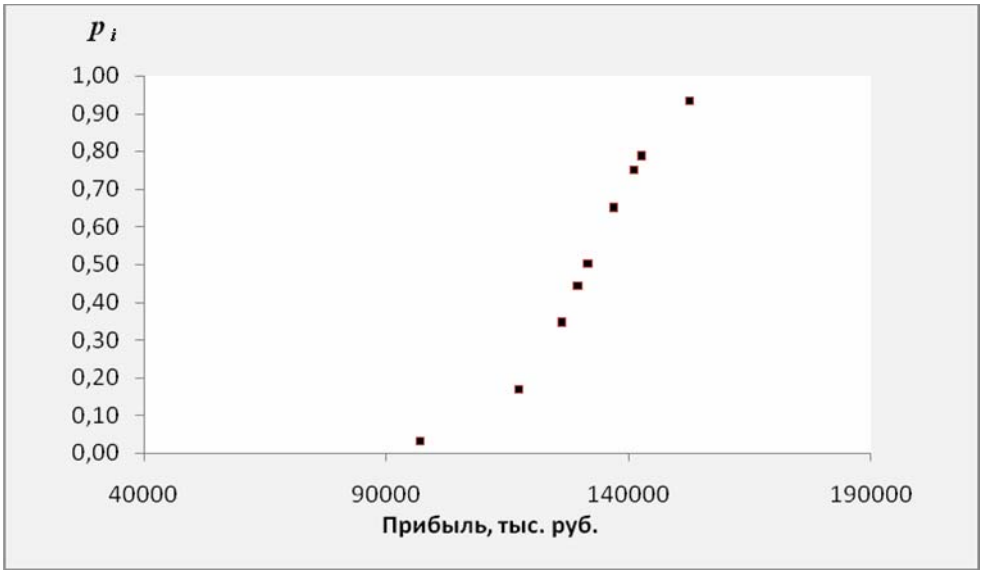


Рис. 2. Изменение прибыли зернового кластера в зависимости от вероятности превышения

Помимо решения задачи оптимизации взаимодействия участников агропромышленного кластера, относящейся к первой группе моделей (без учета влияния природных событий), решена задача с учетом влияния на параметры природных событий. В модели в качестве параметров, подверженных влиянию экстремальных природных явлений, использованы коэффициенты при неизвестных в целевой функции и левых частях ограничений. Для их описания использован закон распределения вероятностей Гаусса.

В этом случае критерий оптимальности и условия, отражающие влияние природных событий, примут вид:

$$f = \sum_{j\in J}\sum_{k\in K}(c_{jk} + \gamma^{p}_{jk})\, x^{III}_{jk} + \sum_{v\in V} c_v x^{I}_{v} - \sum_{i\in I}\sum_{k\in K} c'_{ik} x^{IV}_{ik} - \sum_{q\in Q} c'_q x^{II}_{q}, \quad (19)$$

$$\sum_{i\in I}\sum_{k\in K}(\mu^{p}_{jik} - \varphi^{p}_{ijk}) x^{IV}_{ik} \geq M_j,\ (j \in J), \quad (20)$$

$$\sum_{i\in I}\sum_{k\in K}(\mu^{p}_{jik} - \varphi^{p}_{ijk}) x^{IV}_{ik} \geq \sum_{k\in K} x^{I}_{vk},\ (v \in V, j \in J), \quad (21)$$

$$\sum_{i\in I}\sum_{k\in K}(\mu^{p}_{jik} - \varphi^{p}_{ijk}) x^{IV}_{ik} - \sum_{j\in J}\sum_{k\in K}\eta_{jk} x^{III}_{jk} = x^{II}_{q},\ (q \in Q). \quad (22)$$

Задача решена применительно к Балаганско-Заларинскому зерновому кластеру. При решении принята гипотеза о том, что изменение дополнительной цены на пшеницу (коэффициенты при неизвестных критерия оптимальности) характеризуется нормальным законом распределения с коэффициентом вариации, равным 0,15. Кроме того, в модели учтено влияние природных событий на урожайность пшеницы (левые части ограничений), для описания изменения которой также принята гипотеза о нормальном распределении с коэффициентом вариации – 0,30.

На рис. 3 показано, что прибыль от реализации продукции в кластере в условиях проявления природных событий, может изменяться от 30,5 до 156 млн. руб. с вероятностью превышения 0,05-0,95.

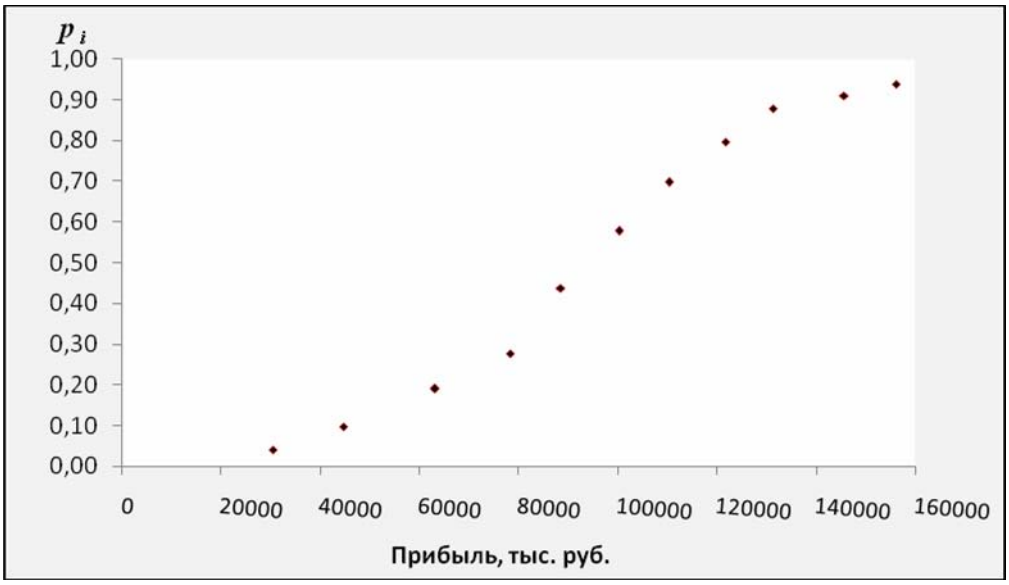


Рис. 3. Изменение прибыли зернового кластера в зависимости от вероятности превышения с учетом влияния природных событий

Следует отметить, что в моделях сложных систем, к которым относятся агропромышленные кластеры, для решения задач с вероятностными параметрами можно использовать метод статистических испытаний, позволяющий случайным образом моделировать интервальные величины и параметры, подчиненные законам распределения вероятностей. Возможность использования метода обусловлена адекватным отображением имитационных значений реальным данным. При этом на предварительном этапе необходимо оценить верхние и нижние оценки параметров, определить законы распределения, которым они подчиняются. С помощью методов имитационного моделирования можно оценить устойчивость результатов в зависимости от различной степени возмущений, влияющих на рассматриваемую систему [1, 2].

Для решения задач с неопределенными параметрами разработаны алгоритмы многократного получения решений с применением метода Монте-Карло для выбора необходимых для управления оптимальных результатов в зависимости от вероятностей случайных параметров (рис. 4).

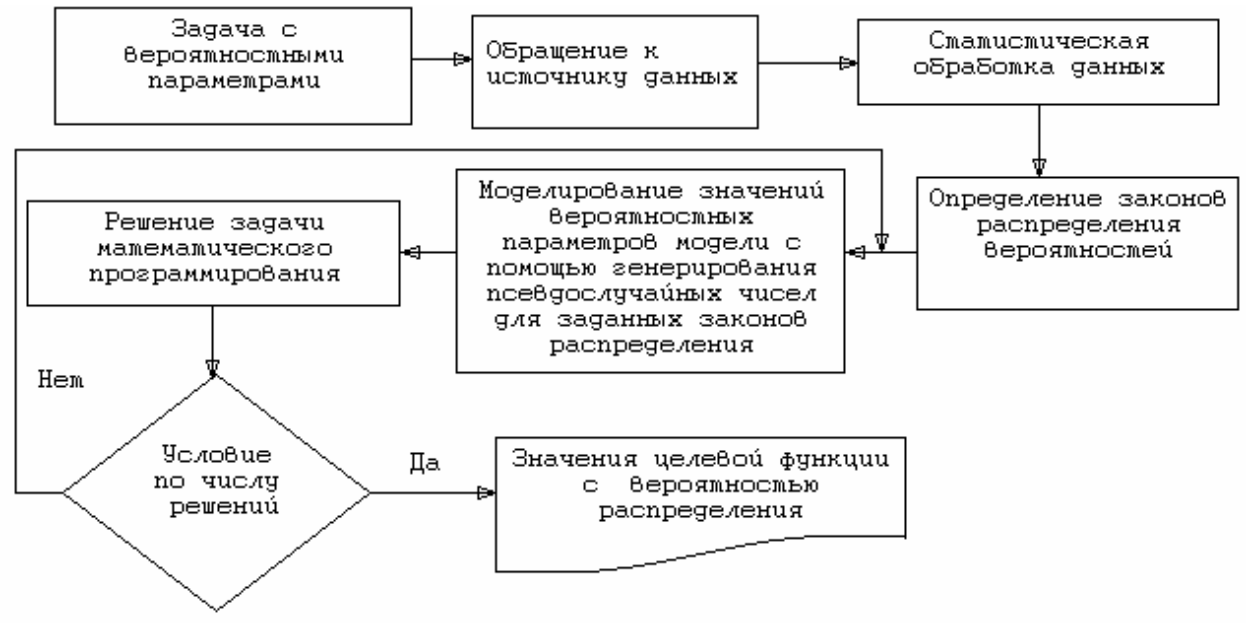


Рис. 4. Алгоритм получения оптимальных решений
в задаче со случайными параметрами

После обращения к данным и статистической обработки полученных сведений по муниципальным районам, выявляются случайные параметры,

входящие в целевую функцию или ограничения задачи. Для них определяются законы распределения вероятностей. Согласно заданным функциям распределения моделируются значения случайных параметров с помощью генерирования псевдослучайных чисел. Решается задача математического программирования с условием по заданному числу значений случайных величин, и находятся оптимальные планы с вероятностью распределения.

В заключение отметим, что рассмотрены две группы задач оптимизации взаимодействия участников агропромышленных кластеров: без учета и с учетом влияния природных событий. Приведены различные варианты моделей с вероятностными параметрами, реализованные для муниципальных районов Иркутской области. Выделены задачи оптимизации взаимодействия участников агропромышленных кластеров с учетом воздействия природных событий, позволяющие оценить их влияние на оптимальные планы с различной вероятностью, что позволяет лицу, принимающему решение, варьировать планированием в условиях природных рисков.

Список литературы

[1] Бузина Т.С., Иваньо Я.М. Программный комплекс оптимизации взаимодействия участников агропромышленного кластера // Вестник ИрГСХА. – 2011. – № 45. – С. 120 – 128.

[2] Бузина Т.С., Иваньо Я.М. Информационное обеспечение моделей агропромышленных кластеров // Вестн. Воронежского гос. техн. ун-та. – 2010. – Т. 6. – № 3. – С. 53 – 57.

[3] Иваньо Я.М. Моделирование сельскохозяйственного производства с учетом экстремальных природных событий // Фундаментальные проблемы изучения и использования воды и водных ресурсов: Материалы науч. конф. – Иркутск: Изд-во Ин-та географии СО РАН, 2005. – С. 230 – 232.

[4] Кардаш В.А. Модели управления производственно-экономическими процессами в сельском хозяйстве. – М.: Экономика, 1981. – 184 с.

[5] Красс М.С., Чупрынов Б.П. Математика для экономистов – СПб.: Питер, 2007. – 464 с.

[6] Орлов А.И. Математика случая: Вероятность и статистика – основные факты: учеб. пособие. – М.: МЗ-Пресс, 2004. – 110 с.

[7] Фидаров В.В., Герасимов Б.И., Романов А.П. Формирование товарно-ассортиментной политики организации в условиях неопределенности: моногр. – Тамбов: Изд-во Тамб. гос. техн. ун-та, 2004. – 152 с.

УДК 519.8

МАКРОМОДЕЛЬ ЭНЕРГЕТИКИ И ЭКОНОМІЧЕСКОГО РОСТА

Горбачук В. М.

Институт кибернетики имени В. М. Глушкова НАН Украины

E-mail: gorbachukvasyl@netscape.net

Пепеляев В. А.

Институт кибернетики имени В. М. Глушкова НАН Украины

E-mail: pepelaev@yahoo.com

Аннатоция. В работе рассмотрена макромодель, куда входят межотраслевая модель затраты-выпуск с оцениваемыми коэффициентами, зависящими от использования производительных факторов капитала, труда, энергии, материалов, и эконометрическая модель поведения национальной экономики, зависящего от бюджетно-налоговых параметров. Цель работы – прогнозирование сценариев развития пяти видов энергетических отраслей (добычи угля, сырой нефти и природного газа, нефтепереработки, услуг электроснабжения, услуг газового хозяйства) в зависимости от динамики цен на энергию.

Ключевые слова. Коэффициенты затраты-выпуск, оценивание параметров, энергетические отрасли, эконометрическая модель.

Введение

Данная работа представляет новый подход к количественному анализу энергетической политики государства, основанный на интеграции эконометрического моделирования и анализа затраты-выпуск. Этот подход объединяет новую методологию для оценки влияния экономической политики на спрос и предложение энергии в условиях полной эконометрической модели экономики государства [33]. Модель состоит из моделей производства для 9–ти промышленных секторов, модели потребительского спроса и макроэконометрической модели роста для экономики государства. Вначале эта модель используется для проектирования экономической деятельности и использования энергии с временным горизонтом 25 лет в предположении неизменности энергетической политики, а потом эта модель применяется для проектирования налоговой программы для стимулирования энергосбережения и уменьшения зависимости от импортированных источников энергии. Общий

вывод состоит в том, что существенного уменьшения энергопотребления можно достичь без значительных экономических затрат.

Повышение мировых цен нефти на порядок, связанное с нефтяным эмбарго ОПЕК в октябре 1973 г., вызвало потребность в новом подходе к количественному анализу экономической политики. Эконометрические модели, которые разработали Нобелевские лауреаты 1969 г. Тинберген (Tinbergen) и 1980 г. Клейн (Klein), оказались весьма полезными в изучении влияния экономической политики на агрегированный спрос. Критический вклад в макроэконометрическое моделирование экономики США совершила модель Klein–Goldberger [35], ставшая основой других макроэконометрических моделей США [29]. Вместе с тем эти модели не обеспечивают оценки влияния экономической политики на предложение (производство). Анализ затраты-выпуск в форме, который предложил Нобелевский лауреат 1973 г. Леонтьев (Leontief) [38] для подробного анализа предложения, базировался на фиксированной во времени технологии [20, 37]. Анализ затраты-выпуск не обеспечивает средств оценки влияния изменений в технологии, вызванных ценовой изменчивостью, связанной с решениями по экономической политике.

Цель работы

Цель данной работы – представить новый подход к количественному анализу энергетической политики государства. Этот подход подробнее представлен в заключительном отчете проекта энергетической политики США [34]. Подход, основанный на интеграции эконометрического моделирования и анализа затраты-выпуск, включает инновационную методологию для оценки влияния экономической политики на предложение. Комбинируя детерминанты спроса и предложения энергии в рамках единой схемы, установим связь между сценариями экономического роста государства и спросом-предложением. Этот подход можно использовать для проектирования экономического роста и использования энергоресурсов государства для любой предлагаемой энергетической политики государства. Он может применяться для исследования влияния мер отдельной политики на спрос и предложение

энергии, цену и стоимость энергии, экспорт и импорт энергии, а также на экономический рост государства.

Первая составляющая данной схемы для анализа энергетической политики – это эконометрическая модель межотраслевых трансакций для 9–ти отраслей отечественной экономики. В деловом секторе экономики государства выделим 9 отраслевых групп: 1) сельское хозяйство, строительство, добыча полезных ископаемых, кроме топлив; 2) переработка (обрабатывающая промышленность), кроме нефтепереработки; 3) транспорт; 4) коммуникации, торговля и услуги; 5) добыча угля; 6) добыча сырой нефти и природного газа; 7) нефтепереработка; 8) коммунальные услуги электроснабжения; 9) коммунальные услуги газового хозяйства.

Межотраслевая модель

Межотраслевая модель представляет межотраслевые трансакции в диаграммном виде следующим образом. К квадратной матрице 9×9 затраты-выпуск для вышеупомянутых 9–ти секторов, называемой **матрицей межотраслевых трансакций**, добавим секторы-строки (rows) первичных входов: 10R) импорт; 11R) услуги капитала; 12R) услуги труда. Прямоугольную матрицу (размерности) 12×9 затрат из секторов 1)–9) и выпуска из секторов-строк 10R)–12R) называют **матрицей первичных входов**. Добавим секторы-столбцы (columns) конечного спроса: 10C) затраты личного потребления; 11C) валовые отечественные (внутренние) частные инвестиции; 12C) правительственные закупки товаров и услуг; 13C) экспорт. Прямоугольную матрицу 12×4 затрат из секторов-столбцов 10C) – 13C) и выпуска из секторов 1) – 9), 10R) – 12R) называют **матрицей конечного спроса**. Введем также сектор-строку 13R) 1×13 – **общие затраты** длиной первых 13 секторов-столбцов и сектор-столбец 14C) 12×1 – **общий выпуск** высотой первых 12 секторов-строк.

Межотраслевая модель состоит из балансовых уравнений между спросом и предложением для продуктов любого из 9–ти базовых секторов. Модель также включает бухгалтерские тождества между стоимостью внутрен-

него ассортимента этих продуктов и суммой стоимостей промежуточных затрат в каждой отрасли, добавленной стоимости в отрасли, импорта конкурирующих продуктов. Спрос на продукты включает спрос на их использование как входов (затрат) любым из 9–ти базовых секторов. Остаток внутреннего ассортимента распределяется среди 4–х категорий конечного спроса, задаваемых секторами-столбцами 10С)–13С).

Для проектирования спроса и предложения энергии берем уровни конечного спроса для всех отраслей из макроэконометрической модели роста. Кроме того, для 5–ти энергетических секторов модели считаем цену и объем импорта экзогенными. Для 4–х неэнергетических секторов считаем цены импорта экзогенными и определяем объемы импорта вместе с объемами услуг капитала и труда в каждой отрасли. Импорт энергии касается только нефтепродуктов и природного газа. В период 1958–1972 гг. импорт нефти в США подлежал системе квот [19]. Импорт газа в США подлежит регулированию Федеральной энергетической комиссией (Federal Power Commission). Цены услуг капитала и труда определяются в рамках макроэконометрической модели. Полагаем объемы экспорта и правительственных закупок в выпуске каждой отрасли экзогенными. Распределение инвестиций среди базовых отраслей считаем заданным.

Затраты личного потребления также включают неконкурентный импорт, жилищные услуги и потребительские товары длительного пользования (ПТДП). Стоимость затрат личного потребления каждой из 11–ти товарных групп модели сектора домашних хозяйств определяется проектированным уровнем затрат личного потребления из макроэконометрической модели, из цен имеющегося внутреннего выпуска каждого сектора межотраслевой модели, из цен неконкурентного импорта, услуг ПТДП и жилищных услуг. Цены неконкурентного импорта считаются экзогенными. Цены капитальных услуг (услуг капитала) для услуг ПТДП и жилищных услуг определяются ценами капитальных услуг из макроэконометрической модели.

Уравнения, представляющие баланс спроса и предложения для каждого из 9–ти секторов межотраслевой модели, устанавливают имеющееся внутреннее предложение равным сумме промежуточного спроса и конечного спроса. Промежуточный спрос определяется одновременно с уровнями выпуска каждой отрасли при данных коэффициентах затраты-выпуск, вычисляемых в модели поведения производителя. Коэффициенты затраты-выпуск определяются одновременно с ценами имеющегося внутреннего выпуска каждой отрасли. Наконец, стоимости услуг капитала и труда для всех секторов и конкурентного импорта для 4–х неэнергетических секторов определяются имеющимися внутренними уровнями и соответствующими коэффициентами затраты-выпуск. Эти значения можно сравнивать с проектированными уровнями макроэконометрической модели.

Модель межотраслевых трансакций

Модель межотраслевых трансакций включает описание и приложения для проектирования спроса и предложения энергии. Введем обозначения: YI_{ij} – промежуточный (intermediate) спрос на выпуск отрасли i от отрасли j; YF_i – конечный (final) спрос на выпуск отрасли i; XD_i – имеющийся внутренний (domestic) выпуск отрасли i; P_i – цена (price) выпуска отрасли i (i–го товара).

Для упрощения обозначений считаем цену выпуска каждой отрасли одинаковой. Дефляторы для каждой категории промежуточного и конечного спроса могут отличаться. Проектируя спрос и предложение энергии, считаем экзогенным отношение дефлятора для индивидуальной категории спроса к дефлятору для имеющегося внутреннего выпуска отрасли.

Межотраслевая модель состоит из равенства между спросом и предложением для каждого из 9–ти секторов модели. Выпишем балансовые уравнениями для 9–ти секторов:

$$XD_i = \sum_{j=1}^{9} YI_{ij} + YF_i\,,\ i = 1, \ldots, 9.$$

Кроме того, модель включает бухгалтерские тождества между стоимостью внутреннего наличия и суммой стоимостей промежуточных затрат в отрасли, добавленной стоимости в отрасли и (если отрасль является одним из 4–х неэнергетических секторов) импорта конкурирующих продуктов:

$$P_i \times XD_i = \sum_{j=1}^{9} P_j \times YI_{ij} + P_K \times K_i + P_L \times L_i + PCIM_i \times CIM_i \ , \ i = 1, \ldots, 9,$$

где P_K – цена услуг капитала (capital); K_i – объем услуг капитала в отрасли i; P_L – цена услуг труда (labor); L_i – объем услуг труда в отрасли i; $PCIM_i$ – цена конкурентого (competitive) импорта (imports) в отрасли i; CIM_i – конкурентный импорт в отрасли i.

Межотраслевая модель включает модели поведения производителя в каждом из 9–ти промышленных секторов. Эти модели можно вывести из границы допустимых цен (price possibility frontier) для 9–ти секторов $i = 1, \ldots, 9$:

$$A_i \times P_i = G_i(P_1, \ldots, P_9, P_K, P_L, PCIM_i),$$

где A_i – индекс уровня нейтральной по Хиксу (Hicks) технологии в отрасли i. Границу допустимых цен для каждого сектора можно вывести из границ допустимых цен для каждой из 3–х подмоделей, используемых в нашем анализе производственной структуры (подробнее интерпретация границы допустимых цен описана в [27]). Эти 3 подмодели таковы:

(1) модель, дающая цену выпуска как функцию цен 4–х агрегированных входов в каждый сектор – капитал K, труд L, энергетику E и материалы M;

(2) модель, дающая цену агрегированного энергетического входа в каждый сектор как функцию цен 5–ти типов энергии модели (уголь, сырая нефть и природный газ, переработанные нефтепродукты, электричество и газ как продукт коммунальных услуг);

(3) модель, дающая цену агрегированного неэнергетического входа в каждый сектор как функцию цен 5–ти типов неэнергетического входа в каждый

сектор (сельское хозяйство, переработка, транспорт, коммуникации и (для 4–х неэнергетических секторов) конкурентный импорт).

Зная цены услуг капитала, труда и конкурентного импорта в каждом из 4–х неэнергетических секторов, можно определить цены P_i внутреннего (отечественного) наличия выпуска для всех 9–ти секторов $i = 1, \ldots, 9$. Для вычисления этих цен решаем систему из 27-ми уравнений для цен внутреннего наличия, цен агрегированного энергетического входа и цен агрегированного неэнергетического входа в каждый из 9–ти секторов. Эта система из 27-ми уравнений состоит из 3–х уравнений для каждого сектора. Эти 3 уравнения соответствуют производственным возможностям каждой из 3–х подмоделей сектора. Применяем теорему незамещения Нобелевского лауреата 1970 г. Самуэльсона [41], по которой при заданных ценах факторов производства и конкурентного импорта цены имеющегося внутреннего выпуска каждого сектора не зависят от состава конечного спроса.

Следующий шаг нашего анализа межотраслевых трансакций – вывод коэффициентов затраты-выпуск для каждого из 9–ти промышленных секторов межотраслевой модели. Коэффициенты затраты-выпуск можно выразить как функции цен. Прежде всего, относительную долю j-го промежуточного входа можно определить из тождества

$$\frac{\partial \ln P_i}{\partial \ln P_j} = \frac{P_j \times YI_{ji}}{P_i \times XD_i} = A_{ji} \frac{P_j}{P_i},$$

где A_{ji} – коэффициент затраты-выпуск, соответствующий YI_{ji}, $i, j = 1, \ldots, 9$; этот коэффициент представляет затраты отрасли j на единицу выпуска отрасли i. Подобные тождества определяют относительные доли услуг капитала, труда и конкурентного импорта (для моделирования поведении производителя).

Далее можно разделить эту относительную долю на отношение цены P_j имеющегося внутреннего выпуска j-й отрасли к цене P_i i-й отрасли, получая коэффициенты затраты-выпуск для $i, j = 1, \ldots, 9$:

$$\frac{YI_{ji}}{XD_i} = A_{ji}(P_1,\ldots,P_9,P_K,P_L,PCIM_i), \quad (1)$$

$$\frac{K_i}{XD_i} = A_{K_i}(P_1,\ldots,P_9,P_K,P_L,PCIM_i),$$

$$\frac{L_i}{XD_i} = A_{L_i}(P_1,\ldots,P_9,P_K,P_L,PCIM_i),$$

$$\frac{CIM_i}{XD_i} = A_{CIM_i}(P_1,\ldots,P_9,P_K,P_L,PCIM_i).$$

Для каждой отрасли выводим коэффициенты затраты-выпуск в два шага. Во-первых, определяем коэффициенты затраты-выпуск для агрегированных входов – капитала (K), труда (L), энергии (E) и материалов (M). Во-вторых, определяем коэффициенты затраты-выпуск для входа каждого типа энергетических затрат на единицу общего энергетического входа и для входа каждого типа неэнергетических затрат на единицу общего неэнергетического входа. Чтобы получить коэффициенты затраты-выпуск, нужные в нашей межотраслевой модели, умножаем коэффициенты затраты-выпуск для каждого типа энергии на коэффициент затраты-выпуск для общего энергетического входа. Аналогично умножаем коэффициенты затраты-выпуск для каждого типа неэнергетического входа на коэффициент затраты-выпуск для общего неэнергетического входа. Получаем коэффициенты затраты-выпуск для услуг капитала, труда, 5–ти типов энергетических входов в каждый сектор и 5–ти типов неэнергетических входов в каждый сектор.

Конечный спрос на имеющийся внутренний выпуск каждого из 9–ти секторов $i, j = 1, \ldots, 9$ модели распределяется среди потребления, инвестиций, правительственных закупок и экспорта:

$$YF_i = C_i + I_i + G_i + Z_i, \quad (2)$$

где C_i – затраты на личное потребление выпуска отрасли i; I_i – валовые частные внутренние инвестиции в выпуске отрасли i (сумма валовых частных фиксированных инвестиций и изменения чистых запасов); G_i – правительст-

венные закупки в отрасли i; Z_i – экспорт в выпуске отрасли i (экспорт минус импорт для 5–ти энергетических секторов).

Проектирование спроса и предложения энергии определяется уровнями конечного спроса для всех отраслей из макроэкономических проекций. Межотраслевая и макроэконометрическая модели связаны тождествами:

$$P_C \times C = \sum_{i=1}^{9} P_i \times C_i ,$$

$$P_I \times I = \sum_{i=1}^{9} P_i \times C_i ,$$

$$P_G \times G = \sum_{i=1}^{9} P_i \times G_i ,$$

$$P_Z \times Z = \sum_{i=1}^{9} P_i \times Z_i ,$$

где $P_C \times C$ – стоимость общих затрат на личное потребление; $P_I \times I$ – стоимость частных внутренних инвестиций; $P_G \times G$ – стоимость правительственных закупок; $P_Z \times Z$ – стоимость экспорта. Эти величины из макроэконометрической модели считаются равными суммам соответствующих категорий затрат по всем 9–ти отраслям межотраслевой модели.

В макроэконометрической модели правительственные затраты на товары и услуги разделяются на две части:

$$P_G \times G = PI_G \times GI + PC_G \times GC ,$$

где GI – объем правительственных (government) затрат на инвестиционные (investment) товары; GC – объем правительственных затрат на потребительские (consumption) товары; PI_G – цена правительственных затрат на инвестиционные товары; PC_G – цена правительственных затрат на потребительские товары.

В нашей макроэконометрической модели чистый экспорт товаров и услуг считается экзогенным. Для целей проектирования спроса и предложения энергии разделяем чистый экспорт на экспорт и импорт, размещаем экспорт

среди 9–ти промышленных групп модели и проектируем цены конкурентного импорта для каждого из 9–ти секторов модели. Поскольку чистый экспорт в текущих и фиксированных ценах является экзогенным для макроэконометрической модели, то такая трактовка чистого экспорта не влияет на структуру полной модели. Для проектировании спроса и предложения энергии используем цены импорта для каждого промышленного сектора, уровни услуг капитала и труда для каждого сектора.

Завершающий шаг в определении конечного спроса на имеющийся внутренний выпуск каждого сектора межотраслевой модели – это размещение затрат на личное потребление среди продуктов 9–ти секторов модели, затрат на неконкурентный импорт и услуги ПТДП, не охваченные продуктами 9–ти секторов. С этой целью применяем эконометрическую модель поведения потребителя. Эта модель основана на косвенной функции полезности [28], которая может быть представлена в форме:

$$\ln V = \ln V\left(\frac{P_1}{PC \times C}, \frac{P_2}{PC \times C}, \ldots, \frac{P_{11}}{PC \times C}\right),$$

где V – уровень полезности. В модель поведения потребителя включены 11 товарных групп: одна группа для каждого из 9–ти промышленных секторов, за исключением сектора 6), и одна группа для каждого из секторов жилищных услуг, услуг ПТДП, неконкурентного импорта.

Учитывая соотношения (1) и (2), уравнения, представляющие баланс спроса и предложения для каждого из 9–ти промышленных секторов $i = 1, \ldots,$ 9 межотраслевой модели, можно представить в виде

$$XD_j = \sum_{j=1}^{9} YI_{i\,j} + YF_i = \sum_{j=1}^{9} A_{i\,j} XD_i + C_i + I_i + G_i + Z_i \,.$$

Коэффициенты затраты-выпуск $A_{i\,j}$ определяются вместе с ценами P_i имеющегося внутреннего выпуска каждой отрасли. Зная цены и уровни агрегированных затрат на личное потребление, определяем уровни затрат C_i на личное потребление. Остальные составляющие конечного спроса I_i, G_i, Z_i про-

ектируется по соответствующей отрасли. Эти проекции совместимы с уровнями валовых частных внутренних инвестиций, правительственных закупок, экспорта из макроэконометрической модели.

В матричной форме уравнения баланса спроса и предложения для модели можно представить как:

$$x = A\,x + y,$$

где x и y – векторы выпуска и конечного спроса:

$$x^T = \begin{pmatrix} x_1 & x_2 & \cdots & x_9 \end{pmatrix},$$

$$y^T = \begin{pmatrix} y_1 & y_2 & \cdots & y_9 \end{pmatrix}.$$

Здесь x^T – вектор-строка (транспонированный вектор-столбец x), A – матрица коэффициентов затраты-выпуск:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{19} \\ A_{21} & A_{22} & \cdots & A_{29} \\ \vdots & \vdots & \ddots & \vdots \\ A_{91} & A_{92} & \cdots & A_{99} \end{pmatrix}.$$

Уровни имеющегося внутреннего выпуска в каждом векторе получаем путем решения системы уравнений:

$$x = (E - A)^{-1} y,$$

где E – единичная квадратная матрица 9×9, A^{-1} – обратная к A матрица.

Стоимости услуг капитала и труда, конкурентного импорта в отраслях $i = 1, \ldots, 9$, определяются уровнями внутреннего наличия и соответствующими коэффициентами затраты-выпуск:

$$K_i = A_{K_i} \times XD_i,$$

$$L_i = A_{L_i} \times XD_i,$$

$$CIM_i = A_{CIM_i} \times XD_i.$$

Коэффициенты затраты-выпуск для услуг капитала и труда, конкурентного импорта являются функциями цен выпуска 9–ти секторов межотраслевой

модели, цен услуг капитала и труда, цен конкурентного импорта для 4–х неэнергетических секторов модели.

Полная эконометрическая модель

Полная эконометрическая модель для межотраслевых трансакций представляется в схематической форме информационных потоков между определенными блоками. Среди них выделим экзогенные (входные) блоки:

PE – уровень затрат к производственной эффективности (production efficiency) выпуска для каждого сектора;

PP – первичные цены (primary prices) импорта, капитала, труда из макромодели;

PM – граничные допустимые цены для каждого из 9–ти производственных секторов моделей производства (production models);

TE – общая (total) стоимость затрат (expenditures) на личное потребление, инвестиции, правительственные закупки по макромодели;

E – стоимость экспорта (exports) от каждого сектора;

ED – энергетические данные (energy data): цены топлива в базовом году, исторические отношения физических единиц (ФЕ) к постоянной денежной единице (ПДЕ), отношения Btu к ПДЕ.

Промежуточные блоки – это функциональные межблочные зависимости:

PD = PD (PE, PP, PM) – одновременное нахождение ценовых границ и определение цен (price determination) выпуска 9–ти секторов;

IOC = IOC (PP, PM, PD) – логарифмические частные производные ценовых границ, приближаемые при равновесных ценах, дают доли затрат в каждом секторе. Такие доли вместе с ценами дают массив 12×9 коэффициентов затраты-выпуск (input-output coefficients);

CM = CM (TE, PD) – модель (model) реального потребительского (consumption) спроса на выпуск каждого сектора;

IG = IG (TE, PD) – пропорционально расщепляем инвестиционные (investment) затраты на спрос для каждого сектора и правительственные (government) закупки на спрос для каждого сектора;

FD = FD (E, IG, CM) – общий реальный конечный спрос (final demand) на выпуск каждого сектора;

IOM = IOM (FD, CM, IOC) – в модели затраты-выпуск (input-output model) находим: общий выпуск каждого сектора при заданных коэффициентах затраты-выпуск и заданном реальном конечном спросе; реальные межотраслевые и первичные трансакции для выпусков секторов и коэффициентов затраты-выпуск.

Конечными исходными блоками являются следующие:

TP (IOM) – формируем матрицу трансакций (transactions) в ПДЕ, текущих ДЕ и ценовых (price) индексах;

EF (ED, CM, IOM, IOC) – из реальных трансакций определяем потоки энергии (energy flows) в Btu и ФЕ, а из ценовых индексов – цены топлива.

Поведение производителя основано на модели взаимосвязи относительного спроса на энергию, услуги капитала и труда, неэнергетических входов. Определим группы затрат, агрегирующие 12 входов модели межотраслевой производственной структуры. Этими товарными группами являются: капитал K; труд L; энергия E (эта группа состоит из входов секторов 5)–9)); материалы M (эта группа состоит из входов секторов 1)–4), 10R)–12R)). Агрегированную модель называют $KLEM$.

Эконометрические спецификации

Эконометрические спецификации агрегированной модели $KLEM$ видны на примере границы допустимых трансцендентальных логарифмических цен, или транслог-цен ($i = 1, \ldots, 9$) [26]:

$$\ln A_i + \ln P_i = \alpha_0^i + \sum_{u=K,L,E,M} \alpha_u^i \ln P_u + 0.5 \sum_{u,v=K,L,E,M} \beta_{uv}^i (\ln P_v)(\ln P_u),$$

где P_E – цена энергии; P_M – цена материалов. Тогда уравнения для относительных долей 4–х входных агрегатов приобретают вид:

$$(P_K K_i)(P_i XD_i)^{-1} = \alpha_K^i + \sum_{u=K,L,E,M} \beta_{Ku}^i \ln P_u ,$$

$$(P_L L_i)(P_i XD_i)^{-1} = \alpha_L^i + \sum_{u=K,L,E,M} \beta_{Lu}^i \ln P_u ,$$

$$(P_E E_i)(P_i XD_i)^{-1} = \alpha_E^i + \sum_{u=K,L,E,M} \beta_{Eu}^i \ln P_u ,$$

$$(P_M M_i)(P_i XD_i)^{-1} = \alpha_M^i + \sum_{u=K,L,E,M} \beta_{Mu}^i \ln P_u , \qquad (3)$$

где E_i – объем входа энергии; M_i – объем входа материалов. Модель *KLEM* для общей переработки (сектора 2)) США, основанная на границе допустимой транслог-цены, была разработана в работе [18]. В ряде работ [15–17] разрабатывались модели замещения между капиталом и трудом для переработки США, основанные на производственной транслог-функции, двойственной к границе допустимой транслог-цены.

Зависимая переменная в каждой из 4–х функций, порожденных границей допустимой транслог-цены (3), – это относительная доля соответствующего входа. Чтобы вывести коэффициент затраты-выпуск для этого входа, разделим эту относительную долю на отношение цены входа к цене выпуска сектора. Например, коэффициентами затраты-выпуск для услуг капитала являются

$$A_{K_i} = \frac{K_i}{XD_i} = \frac{\alpha_K^i + \sum\limits_{u=K,L,E,M} \beta_{K\,u}^i}{\frac{P_K}{P_i}}, \; i = 1, \ldots, 9 .$$

Подобные выражения можно получить из коэффициентов затраты-выпуск для услуг труда (L), энергии (E), материалов (M).

Стоимость имеющегося внутреннего выпуска каждого сектора $i = 1, \ldots, 9$, равняется сумме стоимостей услуг капитала и труда в этом секторе плюс стоимости энергетических и неэнергетических входов в этот сектор:

$$P_i XD_i = \sum_{u=K,L,E,M} P_u \times u_i \; .$$

Исходя из этого бухгалтерского тождества, сумма относительных долей 4–х агрегированных входов в каждый сектор составляет 1. Параметры 4–х сумм относительного спроса для услуг капитала и труда, энергетических и неэнергетических входов должны удовлетворять следующим ограничениям:

$$\sum_{u=K,L,E,M} \alpha_u^i = 1,$$

$$\sum_{u=K,L,E,M} \beta_{uv}^i = 0 \ , \ v = K,L,E,M \ . \qquad (4)$$

Зная оценки параметров любых 3–х уравнений для относительных долей (4), можно получить оценки параметров 4–го уравнения.

Логарифм границы допустимой цены для каждого сектора является дважды дифференцированным по логарифмам цен входов, откуда следует, что гессиан (Hessian) такой функции симметричен. Это порождает множество ограничений относительно параметров смешанных частных производных. Для агрегированной модели *KLEM* 3 таких ограничения – явные:

$$\beta_{uv}^i = \beta_{vu}^i \ , \ u \ , \ v = K,L,E \ .$$

Кроме того, исходя из уравнений (4), оцениваем параметры β_{MK}^i, β_{ML}^i, β_{ME}^i:

$$\beta_{Mu}^i = - \sum_{v=K,L,E} \beta_{uv}^i \ , \ u \ , \ v = K,L,E \ .$$

Отсюда 3 дополнительных ограничения – неявные:

$$\beta_{uM}^i = \beta_{Mu}^i , \ u = K,L,E \ .$$

Таким образом, для каждого из 9–ти промышленных секторов агрегированная модель *KLEM* включает 6 ограничений симметрии.

Подробно представим агрегированную модель *KLEM* в модели поведения производителя, Для подмодели энергии можем выписать границу допустимой транслог-цены в виде

$$\ln P_E = \alpha_0^{E_i} + \sum_{j=1}^{5} \alpha_j^{E_i} P_{E_j} + 0.5 \sum_{j,k=1}^{5} \beta_{jk}^{E_i} \ln(P_{E_j}) \ln(P_{E_k}) ,$$

где P_{E_1} – цена угля, P_{E_2} – цена сырой нефти и природного газа, P_{E_3} – цена продуктов нефтепереработки, P_{E_4} – цена электричества, P_{E_5} – цена газа как продукта коммунальных услуг. Вид подмоделей энергии E и материалов M аналогичен виду агрегированной модели. Подобным образом можно записать границу допустимой транслог-цены для подмодели материалов в виде

$$\ln P_M = \alpha_0^{M_i} + \sum_{j=1}^{5} \alpha_j^{M_i} P_{M_j} + 0.5 \sum_{j,k=1}^{5} \beta_{jk}^{M_i} \ln(P_{M_j}) \ln(P_{M_k}),$$

где P_{M_1} – цена сельского хозяйства, нетопливной добычи и строительства (сектора 1)), P_{M_2} – цена переработки, кроме нефтепереработки (сектора 2)), P_{M_3} – цена транспорта (сектора 3)), P_{M_4} – цена коммуникаций, торговли и услуг (сектора 4)), P_{M_5} – цена конкурентного импорта (сектора 10R)).

Стоимость каждого агрегата равна сумме стоимостей товарных групп, составляющих этот агрегат. Например, стоимость энергии равняется сумме стоимостей 5–ти типов энергии:

$$P_E E_i = \sum_{j=1}^{5} P_{E_j} E_{ji}, \qquad (5)$$

где E_{1i} – объем угля, E_{2i} – объем сырой нефти и природного газа, E_{3i} – объем продуктов нефтепереработки, E_{4i} – объем электричества, E_{5i} – объем газа как продукта коммунальных услуг.

Оценивание параметров агрегированной модели *KLEM* производится для каждого из 9–ти промышленных секторов межотраслевой модели на 4–х уравнениях системы (3). Для 4–х промышленных секторов (секторов 1), 2), 4), 7)) подмодель энергии (E) состоит из 5–ти уравнений для относительных долей секторов 5) – 9). Для этих промышленных секторов можно оценивать лишь 4 из 5–ти уравнений. Если не оценивать, например, уравнение, соответствующее сектору 9) (газу как коммунальной услуги), то относительную долю сектора 9) можно определить из оцениваемых 4–х уравнений и тождества (5).

Для 4–х промышленных секторов межотраслевой модели подмодель энергии (E) состоит из 4–х уравнений для относительных долей 4–х типов энергии: для транспорта, добычи угля, коммунальных услуг электроснабжения (секторов 3), 5), 8)) относительная доля сектора 6) (сырой нефти и природного газа) равна 0; для сектора 9) (газа как коммунальной услуги) относительная доля сектора 8) равна 0. Для этих 4–х секторов 3), 5), 8), 9) вид подмодели энергии (E) аналогичен виду агрегированной модели ($KLEM$). Можно оценивать лишь 3 из 4–х уравнений для относительных долей 3–х типов энергии, за исключением уравнения для относительной доли сектора 9) (газа как продукта коммунальных услуг). Тогда относительную долю сектора 9) можно определить из тождества (5). Учитывая ограничения симметрии в этих 3–х уравнениях, уменьшаем число неизвестных параметров до 9-ти.

Для сектора 6) (сырой нефти и природного газа) межотраслевой модели подмодель энергии (E) состоит из 3–х уравнений для относительных долей 3–х типов энергии, поскольку для этого сектора относительные доли секторов 5) и 9) равны 0. Можно оценивать лишь 2 из этих 3–х уравнений, соответствующих секторам 6) – 8): если не оценивать уравнение, соответствующее сектору 7) (услугам электроснабжения), тогда относительную долю сектора 7) можно определить из тождества (5). Учитывая ограничения симметрии в этих 2–х уравнениях, уменьшаем число неизвестных параметров до 5-ти.

Для 4–х неэнергетических секторов межотраслевой модели подмодель материалов (M) состоит из 5–ти уравнений для относительных долей секторов 1) – 4), 10R) (сельского хозяйства, переработки, транспорта, коммуникаций и услуг, конкурентного импорта). Для 5–ти энергетических секторов 5)–9) подмодель материалов состоит из 4–х уравнений для относительных долей секторов 1) – 4). Вид подмодели материалов (M) для 4–х неэнергетических секторов аналогичен виду подмодели энергии (E) из 5–ти уравнений для относительных долей 5–ти типов энергии. Можно оценивать 4 из этих 5–ти уравнений, соответствующих секторам 1) – 4), 10R): если не оценивать урав-

нение, соответствующее сектору 10R) (конкурентному импорту), тогда относительную долю сектора 10R) можно определить из тождества для 5–ти типов неэнергетических входов, аналогичного тождеству (5). Каждая из этих подмоделей включает 14 неизвестных параметров.

Для 5–ти неэнергетических секторов 1) – 4), 10R) оцениваем 3 уравнения для относительных долей неэнергетических входов, за исключением входов сектора 4) (коммуникаций, торговли и услуг). Подмодели материалов (M) для этих секторов аналогичны агрегированной подмодели ($KLEM$) и включают 9 неизвестных параметров.

Для каждого из 9–ти промышленных секторов межотраслевой модели производства, все 3 подмодели – агрегированная подмодель ($KLEM$), подмодели энергии (E) и материалов (M) – оценивались на годовых данных США 1947 – 1971 гг. межотраслевых трансакций, услуг капитала и труда, конкурентного импорта [30]. Метод оценивания – поиск минимального расстояния для нелинейных одновременных уравнений, где цены конкурентного импорта считаются экзогенными переменными. Такой метод обсуждается в работе [39]. В каждой из этих подмоделей соответствующая система уравнений нелинейна по переменным, но линейна по параметрам.

Заслуживают внимания знаки оценок параметров границы допустимой транслог-цены для каждой из 3–х подмоделей эконометрической модели поведения производителя во всех 9–ти промышленных секторах.

Знаки оценок параметров границы допустимой транслог-цены для агрегированной подмодели ($KLEM$) в 9–ти промышленных секторах экономики США на данных 1947 – 1971 гг. таковы: α_K^i, α_L^i, α_E^i, α_M^i, $\beta_{KK}^i > 0$ при всех $i = 1, \ldots, 9$; $\beta_{KL}^i < 0$ при $i = 1, 3, 5, 6, 7, 9$; $\beta_{KE}^i < 0$ при $i = 1, 2, 3, 6, 7, 8, 9$; $\beta_{KM}^i < 0$ при $i = 1, \ldots, 6$; $\beta_{LL}^i < 0$ при $i = 5, \ldots, 9$; $\beta_{LE}^i < 0$ при $i = 7$; $\beta_{LM}^i < 0$ при $i = 2, 4, 5$; $\beta_{EM}^i < 0$ при $i = 1, 2, 4, 5, 6, 7, 9$; $\beta_{MM}^i < 0$ при $i = 3, 8, 9$; знаки остальных оценок параметров положительны.

Агрегированная граница допустимой цены определяется из цен капитала (K), труда (L), энергии (E), материалов (M). Для каждого из 9–ти промышленных секторов параметры агрегированной подмодели оцениваются из системы 3–х уравнений. Зависимыми переменными в этих уравнениях являются относительные доли капитала, труда, энергии в стоимости общего выпуска. Параметры в уравнении для относительной доли материалов оцениваются из ограничений, обусловленных учетным тождеством между стоимостью выпуска и стоимостью затрат. Используем ограничения, обусловленные условиями симметрии на границе допустимой цены для каждого сектора, уменьшая количество оцениваемых параметров до 9-ти. Кроме того, используем ограничения выпуклости, где они имеют место, чтобы дальше уменьшать количество оцениваемых параметров.

Знаки оценок параметров границы допустимой транслог-цены для подмодели энергии (E) в 9–ти промышленных секторах экономики США на данных 1947–1971 гг. таковы: $\alpha_j^{E_i} > 0$ при всех $i = 1,\ldots,9$, $j = 1,\ldots,5$; $\beta_{11}^{E_i} < 0$ при $i = 5, 6, 7, 9$; $\beta_{12}^{E_i} < 0$ при $i = 1, 2, 3, 4$; $\beta_{13}^{E_i} < 0$ при $i = 1, 2, 3, 4, 6, 9$; $\beta_{14}^{E_i} < 0$ при $i = 2, \ldots, 9$; $\beta_{15}^{E_i} < 0$, $\beta_{45}^{E_i} < 0$ при $i = 2, 3, 4, 5, 8$; $\beta_{22}^{E_i} < 0$ при $i = 2,\ 9$; $\beta_{23}^{E_i} < 0$ при $i = 2, 4, 7$; $\beta_{24}^{E_i} < 0$ при $i = 1, 4, 6, 7, 9$; $\beta_{25}^{E_i} < 0$ при $i = 1, 2, 4$; $\beta_{33}^{E_i} < 0$ при $i = 1, 5, 6, 9$; $\beta_{34}^{E_i} < 0$ при $i = 3, 4, 7, 8$; $\beta_{35}^{E_i} < 0$ при $i = 2, 3, 4$; $\beta_{44}^{E_i} < 0$, $\beta_{55}^{E_i} < 0$ при $i = 1$; знаки остальных оценок параметров неотрицательны.

Граница допустимой цены энергии дает цену энергии для каждого сектора как функцию от 5–ти типов энергетических входов: (E_1) уголь; (E_2) сырая нефть и природный газ; (E_3) продукты нефтепереработки; (E_4) электричество; (E_5) газ как продукт коммунальных услуг.

Наконец, знаки оценок параметров границы допустимой транслог-цены для подмодели материалов (M) в 9–ти промышленных секторах экономики

США на данных 1947 – 1971 гг. таковы: $\alpha_j^{M_i}$, $\beta_{33}^{M_i}$, $\beta_{44}^{M_i} > 0$ при всех $i = 1, \ldots, 9$, $j = 1, \ldots, 5$; $\beta_{11}^{M_i} < 0$ при $i = 4$; $\beta_{12}^{M_i} < 0$, $\beta_{35}^{M_i} < 0$ при $i = 1, 2, 3, 5, 6, 7$; $\beta_{13}^{M_i} < 0$ при $i = \overline{3, 9}$; $\beta_{14}^{M_i}$, $\beta_{24}^{M_i}$, $\beta_{34}^{M_i} < 0$ при $i = 1, \ldots, 9$; $\beta_{15}^{M_i} < 0$ при $i = 3, 5, 6, 7$; $\beta_{22}^{M_i}$, $\beta_{55}^{M_i} < 0$ при $i = 9$; $\beta_{23}^{M_i} < 0$ при $i = 1, \ldots, 7, 9$; $\beta_{25}^{M_i} < 0$ при $i = 1, \ldots, 7$; $\beta_{45}^{M_i} < 0$ при $i = 3, \ldots, 7$; знаки остальных оценок параметров неотрицательны.

Для 4–х неэнергетических секторов граница допустимой цены материалов определяется на ценах 5–ти типов неэнергетических входов: (M_1) сельское хозяйство, нетопливная добыча, строительство; (M_2) переработка, за исключением нефтепереработки; (M_3) транспорт; (M_4) коммуникации, торговля и услуги; (M_5) конкурентный импорт.

Макромодель долгосрочного экономического роста

Макромодель долгосрочного экономического роста основана теории экономического роста и теории поведения частного сектора экономики США. Поведение правительственного сектора 12С) и поведение зарубежного сектора считаются экзогенными. Демографические тренды – динамика населения, рабочей силы, безработицы – также полагаются экзогенными. Основная детерминанта роста производительности – это формирование капитала. Рост производительности свыше роста, обусловленного формированием капитала, считается экзогенным в модели. Проектируем демографические тренды роста производительности на базе послевоенного опыта США. Подробные проекции представлены в работе [32].

Переменные макромодели

Переменные макромодели входят в систему 4–х групп уравнений макроэконометрической модели роста.

Группа из 5–ти поведенческих (behavioral) уравнений описывает поведение секторов домохозяйств и бизнеса:

$$\frac{P_{IS} \times IS}{P_{KD} \times KD} = b_1 - b_2(\ln CS - \ln IS),$$

где IS – предложение (supply) инвестиционных (investment) товаров частными предприятиями, P_{IS} – неявный дефлятор (price) для IS, KD – услуги капитала K, P_{KD} – неявный дефлятор (price) для KD, CS – предложение (supply) потребительских (consumption) товаров частными предприятиями; спрос на труд задается уравнением

$$\frac{P_{LD} \times LD}{P_{KD} \times KD} = b_3,$$

где LD – частные закупки услуг труда (labor), P_{LD} – неявный дефлятор (price) для LD.

Граница производственных возможностей задается уравнением

$$\ln KD = -b_3 \ln LD + b_4 \ln CS - b_5 \ln A + b_1 \ln IS + 0.5 b_2 (\ln CS - \ln IS)^2,$$

где A – общая производительность факторов (затраты относительно выпуска).

Потребительский спрос задается уравнением

$$P_C \times C = b_6 W(-1) + b_7 (P_L \times LH + EL),$$

где C – затраты на личное потребление, включая услуги ПТДП, P_C – неявный дефлятор (price) для C, $W(-1)$ – частное национальное богатство (wealth) в предшествующий период времени, L – предложение услуг труда (labor), P_L – неявный дефлятор (price) для L, LH – имеющееся время, EL – правительственные трансфертные платежи лицам, отличные от фондов социального страхования.

Спрос на отдых задается уравнением

$$P_L \times LJ = b_8 W(-1) + b_9 (P_L \times LH + EL),$$

где LJ – длительность отдыха (leisure), $b_1, \ldots, b_9$ – некоторые положительные параметры, оцениваемые эконометрически.

Модель сектора домохозяйств предложена в работе [21], а модель сектора бизнеса – в работе [24]. Спрос на труд определяется общим уровнем

производства, количеством имеющихся услуг капитала и относительными ценами услуг труда и капитала. Выпуск инвестиционных товаров определяется ценой инвестиционных товаров, ценами услуг капитала и имеющегося предложения услуг капитала, объемом производственной мощности для выпуска потребительских товаров и услуг. Производство, которое имеет место в частном секторе США для потребительских или инвестиционных товаров, ограничивается общей производственной мощностью, которая, в свою очередь, зависит от имеющегося предложения услуг капитала и труда, а также от уровня технологии.

Уровень затрат домохозяйств на потребительские товары и услуги определяется богатством и ресурсами, которые содержатся в секторе домохозяйств, включая фонд времени сектора. Желаемый объем затрат труда, обеспечиваемого сектором домохозяйств, определяется общим количеством имеющегося времени, уровнем зарплаты и размером других ресурсов, которые имеют домохозяйства в форме богатства и трансфертных платежей.

Поведенческие уравнения макроэконометрической модели роста были оценены на исторических данных США за период 1929 – 1971 гг. Данные представлены в ряде статей [22 – 25].

В дополнение к поведенческим уравнениям, модель включает учетные тождества для основного капитала, инвестиций и услуг капитала, для стоимости затрат и выпуска, для сбережений и богатства, для стоимости потребительских товаров, инвестиционных товаров, услуг капитала и труда. Эти учетные тождества объединяют бюджетные ограничения для секторов домохозяйств и бизнеса с потоком каждого продукта и фактора производства в текущих ценах:

$$K = AI \times I + (1 - m)\,K(-1),$$

где K – основной капитал, AI – инвестиции (investment) в основной капитал, I – валовые частные внутренние инвестиции, включая закупки ПТДП, m – темп замещения для частных внутренних материальных активов;

$$KD = AK \times K(-1),$$

где AK – основной капитал после услуг капитала.

Стоимости затрат и выпуска задаются уравнением

$$P_{IS} \times IS + P_{CS} \times CS = P_{KD} \times KD + P_{LD} \times LD,$$

где P_{CS} – неявный дефлятор (price) для C.

Стоимость потребительских товаров задается уравнением

$$(1+TC)\,P_{CS} \times CS + P_{CE} \times CE = \sum_{u=C,CG,CI,CR} P_u \times u,$$

где TC – эффективная налоговая (tax) ставка для C, CE – предложение потребительских (consumption) товаров правительственными предприятиями (enterprises), CG – правительственные (government) закупки потребительских (consumption) товаров, CI – изменение в деловых запасах (inventories) потребительских (consumption) товаров, CR – чистый экспорт потребительских (consumption) товаров минус доход, порождаемый остальным (rest) миром, P_{CE} – неявный дефлятор (price) для CE, P_{CG} – неявный дефлятор (price) для CG, P_{CI} – неявный дефлятор (price) для CI, P_{CR} – неявный дефлятор (price) для CR.

Стоимость потребительских товаров задается уравнением

$$(1+TI)P_{IS} \times IS + P_{CI} \times CI = \sum_{v=I,IG,IR} P_v \times v,$$

где TI – эффективная налоговая (tax) ставка для I, IG – правительственные (government) закупки инвестиционных (investment) товаров, IR – чистый экспорт инвестиционных (investment) товаров, P_I – неявный дефлятор (price) для I, P_{IG} – неявный дефлятор (price) для IG, P_{IR} – неявный дефлятор (price) для IR.

Стоимость услуг капитала задается уравнением

$$(1-TK)[P_{KD} \times KD - TP \times P_I(-1) \times AW(-1) \times K(-1)] =$$

$$= n \times PI(-1) \times AW(-1) \times K(-1) + d \times PI \times AL \times K(-1) + PI(-1) \times AW(-1) \times K(-1) -$$

$$- PI \times AL \times K(-1),$$

где TK – эффективная налоговая (tax) ставка для услуг капитала K, TP – эффективная налоговая (tax) ставка для основного капитала, AW – инвести-

ции в богатство (wealth), n – номинальная ставка отдачи для частных материальных активов, d – темп амортизации (depreciation) для частных внутренних материальных активов, AL – инвестиции в основной капитал с запаздыванием (lag).

Стоимость услуг труда задается уравнением

$$(1-TL)\sum_{u=LD,LGE,LGG,LR} P_u \times u = P_L \times L,$$

где TL – эффективная налоговая (tax) ставка для услуг труда L, LGE – закупки правительственных предприятий (government enterprises) услуг труда (labor), P_{LGE} – неявный дефлятор (price) для LGE, LGG – общие правительственные (general government) закупки услуг труда (labor), P_{LGG} – неявный дефлятор (price) для LGG, LR – чистый экспорт услуг труда (labor), P_{LR} – неявный дефлятор (price) для LR.

Валовые частные национальные сбережения (saving) задаются уравнением

$$S = P_I \times I + P_G[G - G(-1)] + P_R[R - R(-1)],$$

где G – чистые требования к оплате правительством (government), P_G – неявный дефлятор (price) для G, R – чистые требования к оплате остальным (rest) миром, P_R – неявный дефлятор (price) для R.

Уравнение богатства

$$W = P_I \times AW \times K + P_G \times G + P_R \times R.$$

Модель завершают балансовые уравнения между спросом и предложением продуктов и факторов производства в постоянных ценах, а также уравнения агрегации, определяющие накопление запаса потребительских товаров. Хотя в модели определяются валовые частные внутренние инвестиции, их распределение между фиксированными инвестициями и накоплением запасов не определяется. Распределение между накоплением запасов в форме потребительских товаров и других составляющих валовых частных внутренних инвестиций требуется для баланса между спросом и предложением потреби-

тельских и инвестиционных товаров. В модели роста балансовые уравнения описывают потребление

$$CS + CE = C + CG + CI + CR,$$

инвестиции

$$IS + CI = I + IG + IR,$$

время

$$LH = L + LJ,$$

труд

$$L = LD + LGE + LGG + LR + LU.$$

Уравнения агрегации – это уравнение

$$P_{CI} = P_C \times APC,$$

где APC – неявный дефлятор (price) потребительских (consumption) товаров относительно неявного дефлятора изменения деловых запасов потребительских товаров, и уравнение

$$P_{CI} \times CI = P_I \times I \times ACI,$$

где ACI – инвестиции (investment) относительно изменения деловых запасов потребительских (consumption) товаров. Переменная с именем AX является экзогенной переменной агрегации, соответствующей базовой переменной X.

Если макроэконометрический подход к секторам экономики в целом [2] и к энергетическим секторам в частности [8] показывает свою практичность, то подход затраты-выпуск еще не показал всех своих возможностей [36] в силу неполного учета поведения производителей [3, 13]. Например, в секторе электроэнергетики поведение производителей существенно зависит от регулирования рынков [4, 5, 12]. В условиях технологических рисков, свойственных атомной энергетике, рациональное поведение производителей должно брать во внимание редкие события, не считая их выбросами [1, 11, 14]. В услових рыночных рисков оценивание эластичности замещения видов энергии [40] должно учитывать лаги временных рядов [10]. Например, замещения угля и природного газа [7], бензина и дизтоплива [9] на реальных данных выглядит не таким очевидным, как это может представляться теоретиче-

ски. Наконец, доставка энергии к потребителям является отдельным видом деятельности, отличающимся от производства энергии [6, 31].

Список литературы

[1]. Голодников А.Н., Ермольев Ю.М., Кнопов П.С. Оценивание параметров надежности в условиях недостаточной информации // Кибернетика и системный анализ, 2010. – № 3. – С. 109 – 125.

[2]. Горбачук В.М. Макроекономічні методи. – К.: Альтерпрес, 1999. – 263с.

[3]. Горбачук В.М. Методи індустріальної організації. Кейси та вправи. Економіка та організація виробництва. Економічна кібернетика. Економіка підприємства. – К.: А.С.К., 2010. – 224 с.

[4]. Горбачук В.М. Прибутки конкуруючих енергогенеруючих фірм / Якість економічного розвитку: глобальні та локальні аспекти. – Дніпропетровськ: ДНУ ім. О. Гончара, 2011. – Т.1. – С. 63 – 65.

[5]. Горбачук В.М. Організація неповних поєднаних енергоринків // Вісник Дніпропетровського університету. Економіка, 2011. – № 3. – С. 223 – 227.

[6]. Горбачук В.М., Гаркуша Н.І., Стецюк П.І., Єгоров Ю. До розрахунку транзитних обсягів і ставок // PDMU-2011 (Алупка). – К.: КНУ ім. Т. Шевченка, 2011. – С. 61 – 62.

[7]. Горбачук В.М., Гирич А.П. Чи є природний газ і вугілля диференційованими продуктами на щомісячних даних України 2010 року? / Теоретико-методологічні і науково-практичні засади інвестиційного, фінансового та облікового забезпечення розвитку економіки. Ч. 2. – Кам’янець-Подільський: Подільський державний аграрно-технічний університет, 2011. – С. 318 – 322.

[8]. Горбачук В.М., Дроб’язко А.О. Енергетичний сектор в економічній моделі України / Електроенергетика України: стратегія ефективності. І.Р. Юхновський (гол.ред.), В.Г. Бар’яхтар, В.М. Горбачук, В.А. Копилов, М.М. Кулик, В.Т. Меркушов, Г.Г. Півняк, С.Б. Тулуб (ред.) – К.:

Міжвідомча аналітично-консультативна рада з питань розвитку продуктивних сил і виробничих відносин, 2001. – С. 63 – 67.

[9]. Горбачук В.М., Сирніков П.В. Чи є бензин і дизпальне диференційованими продуктами на щомісячних даних України 2010 року? / Наукові дослідження: шлях від теоретичного пошуку до практичної реалізації. – Тернопіль: ТНЕУ, 2011. – С. 32 – 34.

[10]. Горбачук В.М., Толубко І.Є. До перевірки моделей загальної рівноваги / Сучасна наука XXI століття. Ч. 2. – К.: ТК Меганом, 2012. – С. 27–35.

[11]. Кнопов П.С., Голодніков О.М., Пепеляєв В.А. Робастне оцінювання ризику. – К.: Ін-т кібернетики імені В.М. Глушкова, 2009. – 96 с.

[12]. Зоркальцев В.И. Некоторые аспекты реформирования электроэнергетики России / Теория и методы согласованных решений. В.И. Зоркальцев, А.Ю. Филатов (ред.) – Новосибирск: Наука, 2009. – С. 170 – 192.

[13]. Зоркальцев В.И., Хамисов О.В. Равновесные модели в экономике и энергетике. – Новосибирск: Наука, 2006. – 221 с.

[14]. Михалевич В.С., Кнопов П.С., Голодников А.Н. Математические модели и методы оценки риска на экологически опасных производствах // Кибернетика и системный анализ. – 1994. – № 2. – С. 121–138.

[15]. Berndt E.R., Christensen L.R. The translog function and substitution of equipment, structures, and labor in U.S. manufacturing, 1929–1968 // Journal of econometrics, 1973, March. – Pp. 81 – 114.

[16]. Berndt E.R., Christensen L.R. The specification of technology in U. S. manufacturing // University of British Columbia discussion paper 73-17, 1973, November.

[17]. Berndt E.R., Christensen L.R. Testing for the existence of a consistent aggregate index of labor inputs // American economic review, 1974, November. – Pp. 391–404.

[18]. Berndt E.R., Wood D.W. Technology, prices, and the derived demand for energy // University of British Columbia discussion paper 74-09. – 1974, May.

[19]. Burrows J.C., Domencich T. An analysis of the United States oil import quota. – Lexington, MA: Heath-Lexington, 1970.

[20]. Carter A.P., Brody A. Contributions to input-output analysis. – Amsterdam: North-Holland, 1970.

[21]. Christensen L.R., Jorgenson D.W. Intertemporal optimization and the explanation of consumer behavior / Winter meetings of the Econometric Society, 1968.

[22]. Christensen L.R., Jorgenson D.W. The measurement of U. S. real capital input, 1929 – 1967 // Review of income and wealth, 1969, December. – Pp. 293 – 320.

[23]. Christensen L.R., Jorgenson D.W. U. S. real product and real factor input, 1929–1969 // Review of income and wealth, 1970, March. – Pp. 19 – 50.

[24]. Christensen L.R., Jorgenson D.W. Measuring economic performance in the private sector / The measurement of social and economic performance. M. Moss (ed.) // Studies in income and wealth. – New York: Columbia University Press, 1973. – Pp. 233 – 351.

[25]. Christensen L.R., Jorgenson D.W. U. S. income, saving, and wealth, 1929 – 1969 // Review of income and wealth, 1973, December. – Pp. 329 – 362.

[26]. Christensen L.R., Jorgenson D.W., Lau L.J. Conjugate duality and the transcendental logarithmic utility function // Econometrica, 1971, July. – Pp. 255 – 256.

[27]. Christensen L.R., Jorgenson D.W., Lau L.J. Transcendental logarithmic production frontiers // Review of economics and statistics, 1973, February. – Pp. 28 – 45.

[28]. Christensen L.R., Jorgenson D.W., Lau L.J. Transcendental logarithmic utility functions // American economic review, 1974.

[29]. Econometric models of cyclical behavior. Hickman B. (ed.) / Studies in income and wealth. – New York: Columbia University Press, 1972.

[30]. Faucett J. and Associates. Data development for the I-O energy model. Final report to the Energy Policy Project. – Washington, DC, 1973.

[31]. Gorbachuk V., Chumakov B. Who should lead: a producer or a supplier? / Modelare matematica, optimizare si tehnologii informationale. – Chisinau: Academia de transporturi, informatica si comunicatii, 2008. – Pp. 122 – 130.

[32]. Hudson E.A., Jorgenson D.W. U. S. economic growth, 1973 – 2000 / Long term projections of the U. S. economy. – Lexington, MA: Data Resources, Inc., 1974. – P. 161 – 236.

[33]. Hudson E.A., Jorgenson D.W. U. S. energy policy and economic growth, 1975 – 2000 // Bell journal of economics and management science. – 1974, Autumn. – Pp. 461 – 514.

[34]. Jorgenson D.W., Berndt E.R., Christensen L.R., Hudson E.A. U. S. energy resources and economic growth. Final report to the Energy Policy Project. – Washington, DC, 1973.

[35]. Klein L.R., Goldberger A.S. An econometric model of the United States, 1929 – 1952. – Amsterdam: North-Holland, 1955.

[36]. Kurz H.D. Who is going to kiss sleeping beauty? On the «classical» analytical origins and perspectives of input-output analysis // Review of political economy, 2011, June. – Pp. 25 – 47.

[37]. Leontief W.W. Introduction to a theory of the internal structure of functional relationships // Econometrica. – 1947, October. – Pp. 361 – 373.

[38]. Leontief W.W. Studies in the structure of the American economy. – New York: Oxford University Press, 1953.

[39]. Malinvaud E. Statistical methods of econometrics. – Amsterdam: North-Holland, 1970.

[40]. Messner S, Golodnikov A, Gritsevskii A. A stochastic version of the dynamic linear programming model MESSAGE III // Energy, 1996. – Vol. 21. – № 9. – Pp. 775 – 784.

[41]. Samuelson P. A. Nonsubstitution theorems / The collected scientific papers of Paul A. Samuelson. Vol. 1. J. Stiglitz (ed.) – Cambridge, MA: M.I.T. Press. – Pp. 513 – 536.

УДК 630.832

О СОЗДАНИИ ЭНЕРГЕТИЧЕСКИХ ПЛАНТАЦИЙ В РОССИИ И МИРЕ

Губий Е. В.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: egubiy@gmail.com

Аннотация. В статье рассматривается опыт зарубежных стран и оценка эффективности создания в России искусственных плантаций быстрорастущих деревьев для теплоснабжения отдаленных населенных пунктов. Изучена возможность создания энергетических плантаций древесных культур вокруг населенных пунктов, для отопления жилищ в которых используется получаемая биомасса в виде дров или топливных брикетов. Построена математическая модель для анализа эффективности энергетических плантаций. Произведен оценочный расчет эффективности плантации применительно к природно-климатическим условиям Иркутской области. Произведена оценка экономической эффективности мероприятия.

Ключевые слова. Энергетические плантации, биотеплоэнергетика, топливные брикеты.

Введение

Получение энергии из биомассы сегодня является одним из наиболее динамично развивающихся направлений во многих странах мира. Этому способствуют ее большой энергетический потенциал, возобновляемый характер и экологическая безопасность (потребление углекислого газа из атмосферы в процессе роста биомассы соответствует его эмиссии в атмосферу при сжигании). Кроме того, переработка древесины обеспечивает местное население дополнительными рабочими местами, что очень важно в сельской местности, а деньги, выплаченные энергогенерирующими предприятиями за местное сырье, остаются в регионе и способствуют его экономическому развитию. Однако использование биомассы экономически рентабельно только в местах ее сосредоточения. Наличие древесины и ее потенциал в качестве биотоплива, для замены нефти в будущем, неравномерно распределено в мире. Производство дровяной древесины является значительным всего лишь в нескольких промышленных странах, таких как Индия, Китай, Бразилия, Канада, США, Финляндия, Швеция и Австрия. Однако поскольку дровяная древесина часто используется в частных домохозяйствах и зачастую ее сбор и торговля

проходят неофициально, статистические данные являются не достаточно точными. По некоторым данным в мире существует около 100 миллионов гектаров земли, используемой для плантаций древесных культур [6].

Предполагается, что древесину можно собирать на топливо в существующих лесах (невостребованная биомасса в виде древесных остатков) и с искусственно выращенных энергетических плантаций. Эффективность использования древесной биомассы обратно пропорциональна расстоянию ее транспортировки от места образования к месту использования. Такая закономерность объясняется тем, что стоимость древесной биомассы при транспортировке резко возрастает. Поэтому в первом случае древесина используются в ограниченном количестве из-за высоких затрат на их сбор и транспортировку к местам потребления. Для второго способа получения биомассы необходимы крупные инвестиции. Однако учитывая низкую стоимость исходного сырья и постоянно растущие цены на традиционные источники энергии, использование древесной биомассы, выращенной на специальных плантациях получает все большее распространение. Проблема транспортировки древесной биомассы решается путем создания специальных энергетических плантаций из быстрорастущих пород древесины во многих странах мира, таких как Швеция, Дания, Финляндия, США, Канада, КНР, Бразилия и др. [8]

Использование энергии биомассы в России и мире

Из биомассы в мире ежегодно получают порядка 2 млрд. т.у.т. энергии, что составляет около 14 % общего потребления первичных энергоносителей. В развивающихся странах этот показатель превышает 30%, (иногда доходит до 50-80%), однако в промышленно развитых государствах он составляет в среднем 2-3%. Вклад биомассы в общий энергобаланс Европейского Союза в 2004 г. составлял 3,6%. В Финляндии биотопливом покрывается 19,6% общего потребления первичной энергии, в Швеции –16,1%, в Австрии – 11%, в Дании – 10,3%, в Польше – 4,5%, в Германии – 2,1% и т.д. Значительных успехов достигли Латвия (28%), Эстония (10,5%), Литва (7,6%). В России

официальная доля биоэнергетики в общем балансе составляет всего лишь 0,56 % [7].

Первые энергетические плантации были созданы в Германии ещё в 30 – 40-х годах прошлого века (во времена Третьего рейха). Однако широкое распространение такой способ получения энергии получил в конце 70-х годов в результате очередного энергетического кризиса, когда в ряде стран занялись поиском эффективных технологий энергетического использования древесины. Наиболее популярна идея создания энергетических лесов оказалась в Швеции. В 1980-е годы ивовая энергетика стала основой стратегического плана развития этой страны. На сегодняшний день энергетические посадки ивовых деревьев организованы на 16000 га болотных земель Швеции. Уборка ежегодного прироста древесины осуществляется в зимнее время комбайнами, когда болота замерзают.

Крупнейшая на сегодня в Европе электростанция, работающая на специально выращенной древесной биомассе, находится в г. Зиммеринг в Австрии. Мощность этой электростанции составляет 66 МВт, она ежегодно потребляет 190 тысяч тонн биомассы, собираемой в радиусе 100 км от станции. Посадка ивы на площади 324 га осуществлена в Великобритании с целью получения фитомассы, служащей топливом на ТЭС мощностью 10 МВт (количественные данные по состоянию на 2006 г.). В Германии ежегодно производят 20 миллионов куб. м. древесины на специально выращенных энергетических плантациях.

В США первые энергетические плантации генетически модифицированных тополей появились в 1998 г. В 2008 г. Конгресс США принял закон «H.R.2419, the Food, Conservation, and energy Act of 2008», согласно которому предусмотрено ежегодное финансирование энергетического лесоводства в США в размере $15 млн. Там, как и в Канаде, действуют специальные правительственные программы поддержки развития энергии биотоплива.

На основании анализа зарубежной информации по энергетическому использованию древесной биомассы можно установить, что основными на-

правлениями проведения научно-исследовательских и проектно-конструкторских работ в этой области за рубежом являются:

1) поиски возможностей увеличения резервов древесной биомассы путем повышения производительности лесных массивов и оптимизации сроков рубки при заготовке энергетической биомассы;

2) повышение эффективности использования различных видов древесной биомассы при непосредственном сжигании ее в топках котельных установок для выработки тепловой и электрической энергии;

3) производство из древесной биомассы твердого топлива с повышенными показателями по транспортабельности, теплоте сгорания и водостойкости;

4) получение из древесной биомассы жидкого топлива для транспортных машин;

5) производство на основе древесной биомассы газообразного топлива, пригодного для использования в современных системах газоснабжения, а также в газобаллонных транспортных машинах.

В России из-за низкой плотности населения, высокой степени разбросанности сельских населенных пунктов и отсутствия централизованных систем теплоснабжения на большей части территории страны наиболее распространенным источником теплоэнергии в отдаленных районах служит биомасса деревьев. Это вызвано потребностью в источниках теплоэнергии, не требующих дорогостоящей транспортировки. Однако часто сбор древесины носит хаотичный порядок, зависит исключительно от природных условий, отмечается недоэксплуатация «порубочных остатков» (откомлевка, сучья, вершинник) на лесосеках. Экономическая эффективность такого вида топлива становится значительно выше, в тех районах, где уже проводятся специальные мероприятия, направленные на организацию выращивания древесины. В России только начинают закладывать опытные плантации.

Математическая модель оценки эффективности энергетических плантаций

Рассмотрим возможность создания специальных энергетических древесных плантаций вокруг населенных пунктов, для отопления жилищ которых используется получаемая биомасса в виде дров, топливных гранул или брикетов. Для анализа эффективности энергетической плантации построим математическую модель. Пусть величина Q означает норматив потребления теплоэнергии на отопление жилых помещений в поселке численностью n, который определяется как произведение годового удельного расхода теплоты на отопление 1 кв.м. площади и нормы жилищной обеспеченности (кв.м. общей площади на 1 человека). Она определяется по формуле

$$Q = Q_g \cdot n \cdot 18 . \quad (1)$$

Здесь Q_g – годовой удельный расход теплоты на отопление 1 кв.м. площади, 18 – норматив жилищной обеспеченности (количество кв.м. жилой площади на 1 человека), n – численность населения.

Годовой расход теплоты на отопление 1 кв.м. общей площади жилых зданий определяется по формуле

$$Q_g = g_h * \frac{t_{in} - t^m_{out}}{t_{in} - t_{out}} * 24 * d * 10^{-6} . \quad (2)$$

Здесь g_h – максимальный часовой расход теплоты на отопление 1 кв.м. общей площади жилых зданий, зависит от этажности (т.у.т./час), t_m – расчетная температура внутреннего воздуха отапливаемых зданий (^{0}C), t_{out} – расчетная температура наружного воздуха для проектирования отопления (^{0}C), t^m_{out} – средняя температура наружного воздуха за отопительный сезон (^{0}C), 24– продолжительность работы систем отопления в сутки (час), d – средняя продолжительность отопительного сезона (сутки).

Таким образом, на норматив потребления теплоэнергии значительное влияние оказывают этажность и теплотехнические характеристика зданий.

Для упрощения будем считать, что площадь плантаций представляет собой круг с радиусом *r* и с центром в населенном пункте. Круг разделен на *t* равных сегментов. Количество сегментов равно сроку созревания древесины (количеству лет). Каждый год на одном из сегментов весной осуществляются лесозаготовки, а осенью – посадки.

Обозначим *S(r,t)* – площадь сегмента, с которого в данном году осуществляются лесозаготовки (выраженная в квадратных километрах). Эта площадь определяется по формуле

$$S(r,t) = \frac{\pi r^2}{t}. \qquad (3)$$

Пусть $f_i(t)$ – функция прироста древесины на единицу площади (количество т. у. т., накопленного в биомассе древесины на одном квадратном километре по прошествии *t* лет после посадки), *i* – вид древесины, $i = 1, \ldots, k$, *k* – количество рассматриваемых видов древесины.

Обозначим $R(i, r, t)$ объем годового производства топлива, зависящий от радиуса *r*, количества сегментов плантации *t*, а также от вида древесины *i*. Этот объем годового производства топлива вычисляется по формуле

$$R(i, r, t) = S(r,t) \cdot f_i(t). \qquad (4)$$

Обозначим $C(i, r, t)$ приведенные затраты на производство древесного топлива в зависимости от вида древесины *i*, радиуса *r* и количества секторов *t*. Эти затраты включают в себя инвестиции на создание всей производственной инфраструктуры, затраты на непосредственное выращивание растений, затраты на заготовку и обработку древесины.

На основе перечисленных показателей предлагается, в частности, исследовать следующую задачу – определить оптимальный срок произрастания *t* культуры *i* и радиуса лесной площади *r* для обеспечения заданного потребления в топливе с минимальными издержками:

$$C(i, r, t) \to \min \qquad (5)$$

при условии, что объем годового производства топлива больше или равен потребности

$$R(i, r, t) \geq Q. \tag{6}$$

Следует отметить, что возраст заготавливаемой древесины не должен превышать такой момент времени t, при котором впервые не выполняется условие среднегодового увеличения биомассы

$$\frac{f(t)}{t} \leq \frac{1}{t-1} f(t-1). \tag{7}$$

Отсюда следует, что нужно ограничиться рассмотрением возраста древесины, не превышающей момент времени t, при котором впервые выполняется неравенство

$$(f(t) - f(t-1)) \leq \frac{f(t)}{t}. \tag{8}$$

Выбор оптимального объема использования древесного топлива в зависимости от цены альтернативного топлива и численности населения:

$$C(i, r, t) + pL \to \min, \tag{9}$$

при условии

$$K(i, r, t) + L = Q. \tag{10}$$

Здесь L – объем использования альтернативного топлива (угля), p – цена альтернативного топлива (включая расходы на транспортировку) [3].

Приведем некоторые конкретные данные. Для примера представим населенных пункт в Иркутской области численностью 5000 человек, состоящий из одноэтажных домов. Годовой расход теплоты на отопление 1 м^2 общей площади жилых зданий согласно формуле (2) равен 0.474 Гкал. или 0,0677 т.у.т. на 1 кв.м.

Климатические параметры, входящие в формулу (2) принимаются по данным местных метеорологических служб, либо по СНиП 2.01.01 – 82 «Строительная климатология и геофизика» (для Иркутской области $g_h = 158$ Ккал/час, $t_m = 18\ ^0$C, $t_{out} = -39\ ^0$C, $t_{out}^m = -10{,}3\ ^0$C, $d = 252$ суток).

Годовой норматив потребления теплоэнергии на отопление жилых зданий в поселке:

$$Q = 0.0677 \cdot 18 \cdot 5000 = 6093 \text{ (т.у.т.)},$$

где 18 – норматив жилой площади на 1 человека, а предполагаемая численность населения – 5000 человек.

Иными словами, для того, чтобы в течении одного года отапливать поселок, необходимо 6093 тонн условного топлива (или 1,21 тонны условного топлива на человека).

Согласно полученным данным специалистов Сибирского института физиологии и биохимии растений СО РАН, в зависимости от климата и микроклимата каждые 4 – 6 лет растения должны достигать высоты 4 – 5 метров и в пересчете на 1 гектар давать около 10 – 15 тонн древесины [8], что эквивалентно 6,85 т.у.т. с гектара. Основываясь на этих данных, применим $t = 4$. С помощью формул (3), (4) и (5) получаем:

$$6093 = \frac{\pi r^2}{4} * 1000,$$

$$r = 3{,}36.$$

Из уравнения мы получаем, что при сроке ротации 4 года, радиус плантации должен составлять 3,36 километра.

Оценим экономическую эффективность мероприятия. Для этого сделаем оценку приведенных затрат и сопоставим их с ценой на альтернативное топливо [5]. Эти затраты включают следующие составляющие.

1. Инвестиции на создание всей производственной инфраструктуры (строительство дорожной сети (21 километр грузосборочной магистрали и 5 км постоянных лесных дорог); строительство складов; установки для дробления и сушки древесины; автотранспорт (лесовоз и трактор) составят около 38,3 млн. руб.

2. Затраты на материалы для одного производственного цикла (посадочный материал (700 тыс.шт. саженцев); удобрения; дизельное топливо) составят около 3 млн. руб.

3. Затраты, связанные с работами по выращиванию растений и обработке биомассы для одного производственного цикла (посадка (50 тыс.часов); рекультивация (10 тыс.часов); заготовка (60 тыс.часов) составят около 3,5 млн. руб.

Выводы

На основании оценки экономической эффективности создание энергетических плантаций признано пригодными для теплоснабжения. Без учета первоначальных инвестиций стоимость 1 т. у. т. такого вида энергии обойдется в 1090 рубль, что ниже стоимости энергетического угля во многих районах Иркутской области (даже без учета стоимости транспортировки). Использование энергодревесины в теплоэнергетике региона позволит улучшить экологическую ситуацию, снизив выбросы углекислого газа, а также ввоз каменного угля на 8 тыс. тонн или мазута – на 4,5 тыс. тонн.

Эта оценка позволяет надеяться на возможность эффективного использования плантации. Особенно привлекательной идея использования энергетических плантаций должна быть для районов удаленных от основных дорожных магистралей.

Список литературы

[1] Биомасса // Официальный сайт электронного журнала энергосервисной компании "Экологические системы", 2009. – № 4. [Электронный ресурс]– Режим доступа: http://esco-ecosys.narod.ru/2009_4/art024.htm.

[2] Грошев Б.И., Синицын С.Г., Мороз П.И., Сеперович И.П. Лесотаксационный справочник. – М.: Лесн. пром-ть, 1980. – 288 с.

[3] Губий Е.В. Создание энергетических плантаций для развития биотеплоэнергетики в России // Охрана и рациональное использование животных и растительных ресурсов: Материалы международной научно-

практической конференции. Иркутск, 24 – 26 мая 2012 г. – Иркутск: Изд-во ИрГСХА, 2012. – С. 146 – 150.

[4] Жигунов А.В. Энергия с плантаций // Российская лесная газета, 2006. – № 29 – 30. – С. 159 – 160.

[5] Зоркальцев В.И., Губий Е.В. Об использовании энергетических плантаций для развития теплоэнергетики в Сибири // Материалы научно-практической конференции «Лесоприльзование в послепожарных древостоях, их реабилитация». – Иркутск: РИО САПЭУ, 2011. – С. 59 – 64.

[6] Леса и энергия. Основные вопросы. Отчет продовольственной и сельскохозяйственной организации объединенных наций, 2008. – Рим, Италия. [Электронный ресурс] – Режим доступа: ftp://ftp.fao.org/docrep/fao/010/i0139r/i0139r00.pdf.

[7] Рагулина И.Р. Биоэнергетический потенциал Калининградской области: автореф. дисс. канд. географ. наук. – Калининград, 2007. – 162 с.

[8] Цивенкова Н.М., Самылин А.А. Быстрорастущие плантации тополя – новая энергетическая сырьевая база // ЛесПромИнформ, 2005. – №8. – С. 58 – 63.

УДК 622.69.019.3+620.9.338.9

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЛЯ АНАЛИЗА И СИНТЕЗА НАДЕЖНОСТИ ПРИ МНОГОУРОВНЕВОМ МОДЕЛИРОВАНИИ СИСТЕМ ГАЗОСНАБЖЕНИЯ

Дзюбина Т. В.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: tvleo@isem.sei.irk.ru

Илькевич Н. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: lkev@isem.sei.irk.ru

Аннотация. В статье описаны математические модели анализа надежности: сложного действующего магистрального газопровода (МГ), газового месторождения и подземного хранилища газа (ПХГ); дано описание математической модели синтеза структурной надежности проектируемого МГ, более подробно ее модификация с учетом инвестиционных проектов МГ; приведена математическая модель анализа надежности функционирования сложной закольцованной системы газоснабжения; показаны алгоритмические особенности разработанных моделей по сравнению с существующими; даны описания соответствующих вычислительных программ и комплексов; а также предлагается двухэтапный методический подход для нахождения оптимальной надежности сложной газоснабжающей системы, основанный на определении эквивалентных характеристик надежности её объектов и оптимизации средств резервирования.

Ключевые слова. Надежность объекта, интенсивность отказов и восстановлений, закон распределения вероятностей работоспособного состояния, показатели надежности газоснабжения потребителей, заданная надежность.

Введение

В настоящее время в России развиваются процессы реформирования и реструктуризации энергетики, в том числе и газовой отрасли.

Несмотря на то, что газовая отрасль является «энергетическим компенсатором», подстраховывающим энергетику страны от возможных критических дисбалансов, перед ней также встают серьезные проблемы [1].

1. ***Старение оборудования.*** Почти 2/3 магистральных газопроводов (МГ) имеют срок эксплуатации более 20 лет, и около 20% установленной мощности газоперекачивающих агрегатов (ГПА) нуждается в замене. Все это

приводит к возрастанию числа отказов на действующих объектах газоснабжающих систем (ГСС).

2. ***Отрыв основных районов добычи газа от мест его интенсивного потребления.*** Это приводит к чрезмерной насыщенности газотранспортных поставок, и в случае крупной аварии может возникнуть дефицит газа для ряда регионов страны и существенное ограничение экспорта газа.

3. ***Продвижение в районы крайнего Севера и Заполярья****,* в том числе и на шельфы северных морей, при освоении ресурсов углеводородного сырья. Это приводит к существенному удорожанию освоения месторождений и требует разработки современных высокоэффективных и надежных способов добычи и транспорта природного газа.

4. ***Недостаточные маневренные возможности в Единой системе газоснабжения (ЕСГ).*** Значительная часть подземных хранилищ газа (ПХГ) осталось в бывших республиках Советского Союза. Это определяет низкие возможности в России по сезонному хранению газа и значительно затрудняет обеспечение надежности подачи газа потребителям в периоды прохождения зимних максимумов нагрузок.

В ИСЭМ СО РАН ведутся исследования основных проблем развития ГСС в современных условиях [2]: разрабатываются подходы, методы и программные средства для многоуровневого моделирования систем газоснабжения, принципы согласования решений, где особое место отводится учету свойств надежности этих систем на разных уровнях иерархии [3]. Разработки ориентированы на различные временные и территориальные уровни и позволяют рассматривать ГСС как: 1) подсистему топливно-энергетического комплекса (ТЭК); 2) самостоятельную газовую отрасль; 3) функционально-целостные системы: ЕСГ в целом и районные системы газоснабжения (РСГ); 4) совокупность производственно-технических объектов и подсистем.

Иерархия моделей оптимального развития системы газоснабжения показана на рис. 1.

На ***первом уровне*** система газоснабжения представляется в виде блока в общей модели ТЭК, отражающей баланс производства, передачи и потребления различных видов энергии по территории страны. Надежность работы отраслей в модели учитывается в виде нормативов народнохозяйственных резервов энергоносителей, которые закладываются в топливно-энергетический баланс, а также затрат самофинансирования отраслей, включающих в том числе затраты на поддерживание существующих и вновь создающихся резервов.

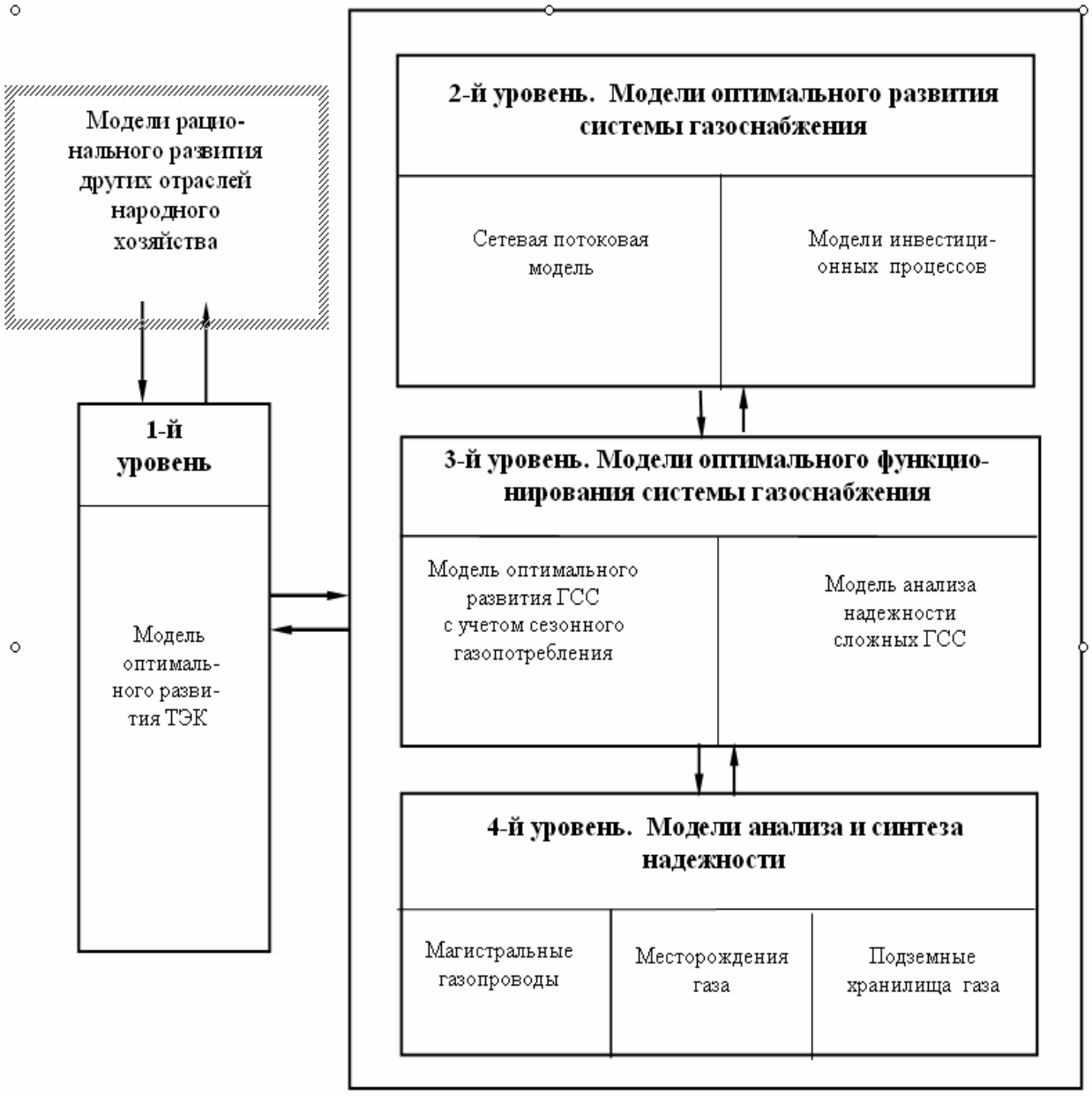


Рис. 1. Иерархия моделей оптимального развития ГСС

На ***первом уровне*** система газоснабжения представляется в виде блока в общей модели ТЭК, отражающей баланс производства, передачи и потребления различных видов энергии по территории страны. Надежность работы отраслей в модели учитывается в виде нормативов народнохозяйственных резервов энергоносителей, которые закладываются в топливно-энергетический баланс, а также затрат самофинансирования отраслей, включающих в том числе затраты на поддерживание существующих и вновь создающихся резервов.

На ***втором уровне*** с учетом интересов отрасли (как естественной монополии) с помощью сетевой потоковой модели детализируются решения, полученные на первом уровне. Надежность учитывается ценами самофинансирования, которые включают затраты на поддержание существующих и вновь создающихся резервов.

На ***третьем уровне*** с помощью квазидинамической модели детализируются за счет учета сезонного регулирования варианты развития ГСС, полученные на втором уровне. Надежность учитывается интегрально коэффициентами в технологической матрице, а также дисконтированными ежегодными затратами (включая затраты на поддержание требуемой надежности) действующих и новых объектов (с учетом создания необходимых резервов).

На ***четвертом уровне*** с позиций интересов инвесторов должны рассчитываться параметры объектов систем газоснабжения.

Детализированные решения, получаемые на 3-ем и 4-ом уровнях, требуют и более детального учета надежности.

В рамках исследований оптимального развития газоснабжающих систем с учетом многоуровневого моделирования была разработана система математических моделей (ММ), алгоритмов и вычислительных средств для анализа и синтеза надежности для низших уровней иерархии [4 – 7], показанная на рис 2.

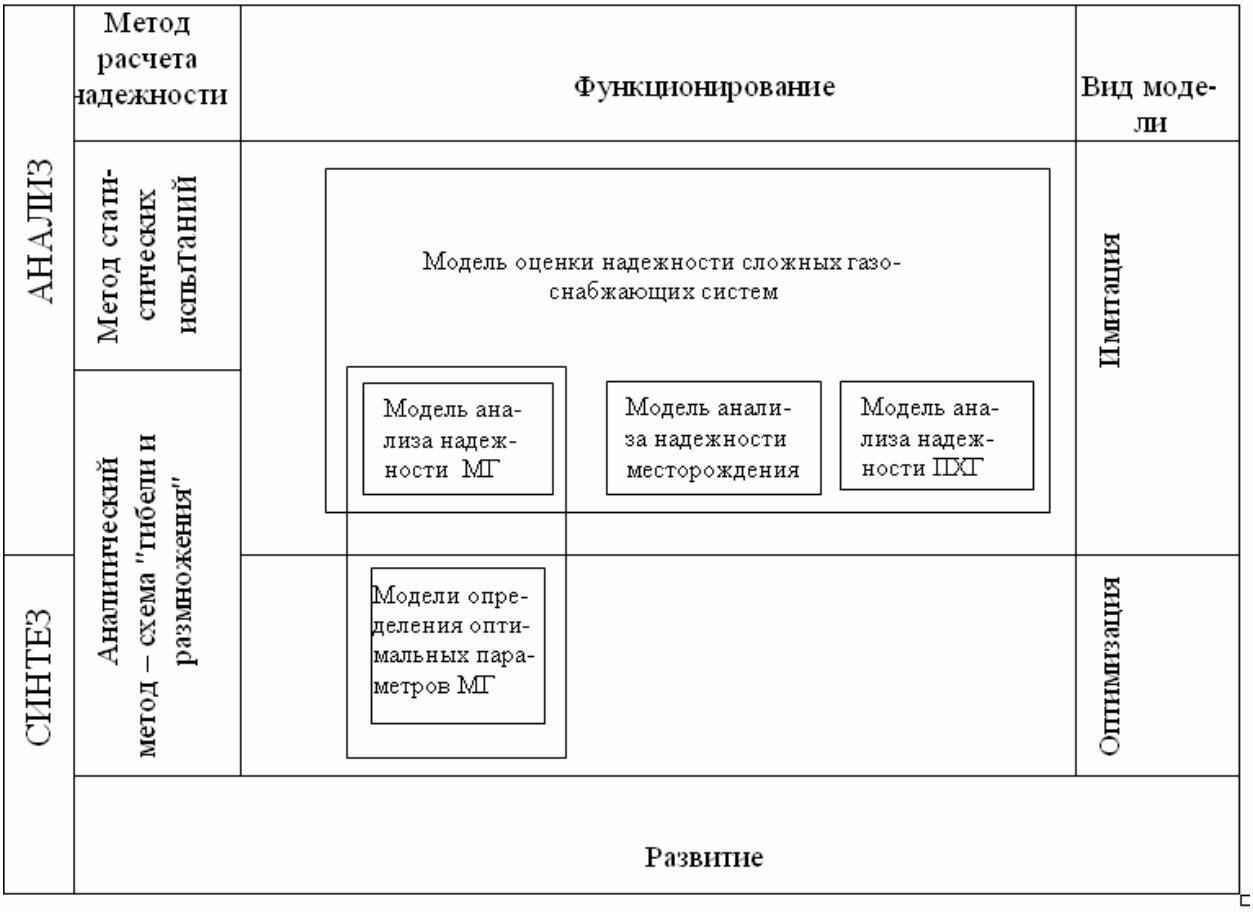


Рис. 2. Классификация разработанных моделей

Модели анализа надежности действующего магистрального газопровода, газового месторождения и подземного газохранилища

Модели анализа надежности действующих магистрального газопровода, газового месторождения и подземного газохранилища [3,8] используются при решении задач 4-ого иерархического уровня.

Объектом исследования является либо сложный многониточный газопровод, либо месторождение газа, либо подземное газохранилище. Каждый такой объект состоит в общем случае из нескольких наборов элементов.

Магистральный газопровод представляет собой совокупность нескольких параллельных ветвей, каждая из которых рассматривается как цепь последовательно соединенных звеньев – линейных участков трубопроводов разных диаметров и компрессорных станций (КС) с разными типоразмерами газоперекачивающих агрегатов, рис. 3.

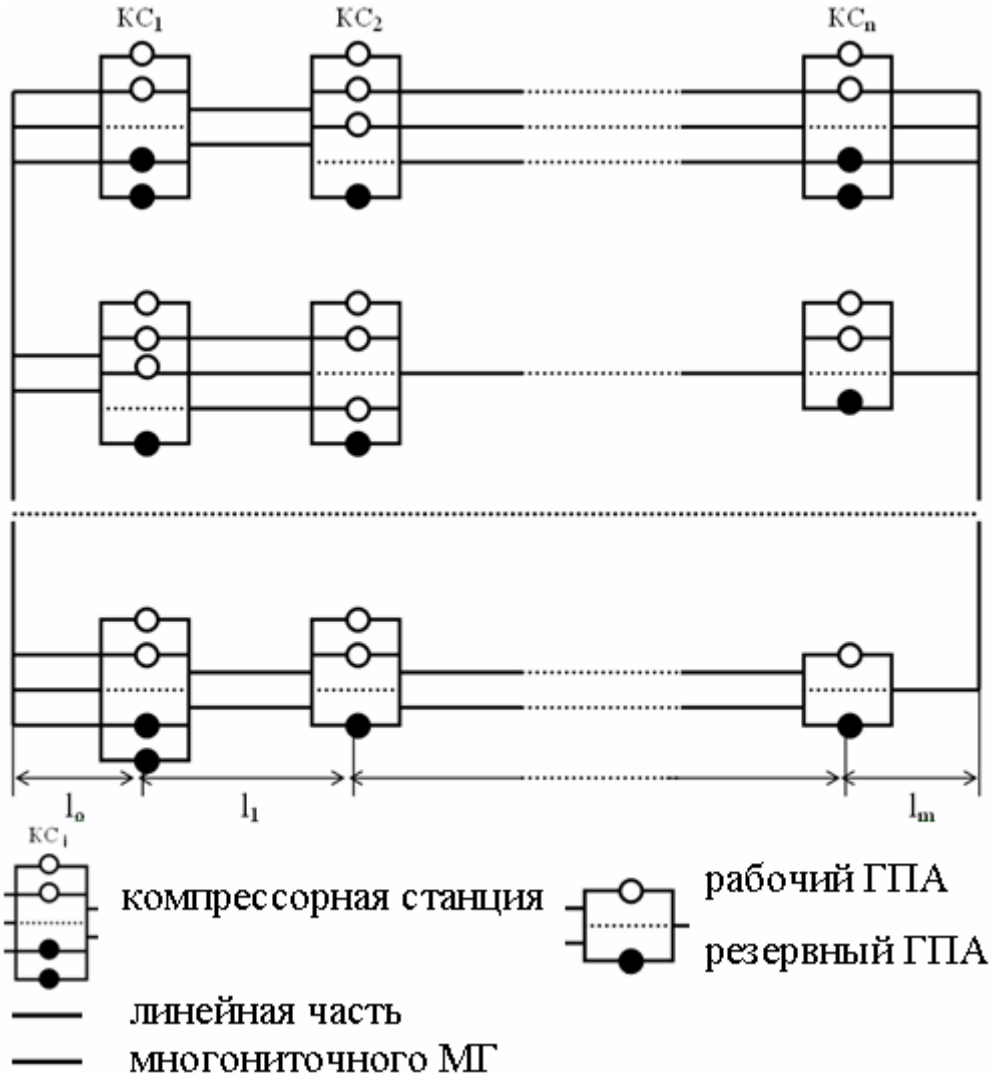


Рис.3. Расчетная схема многониточного МГ

Газовое месторождение и ПХГ, которое обычно создается в выработанных месторождениях, рассматриваются как совокупность параллельно соединенных кустов скважин, последовательно соединяющихся с газосборным коллектором, который подает газ на промысловую газораспределительную станцию, рис. 4.

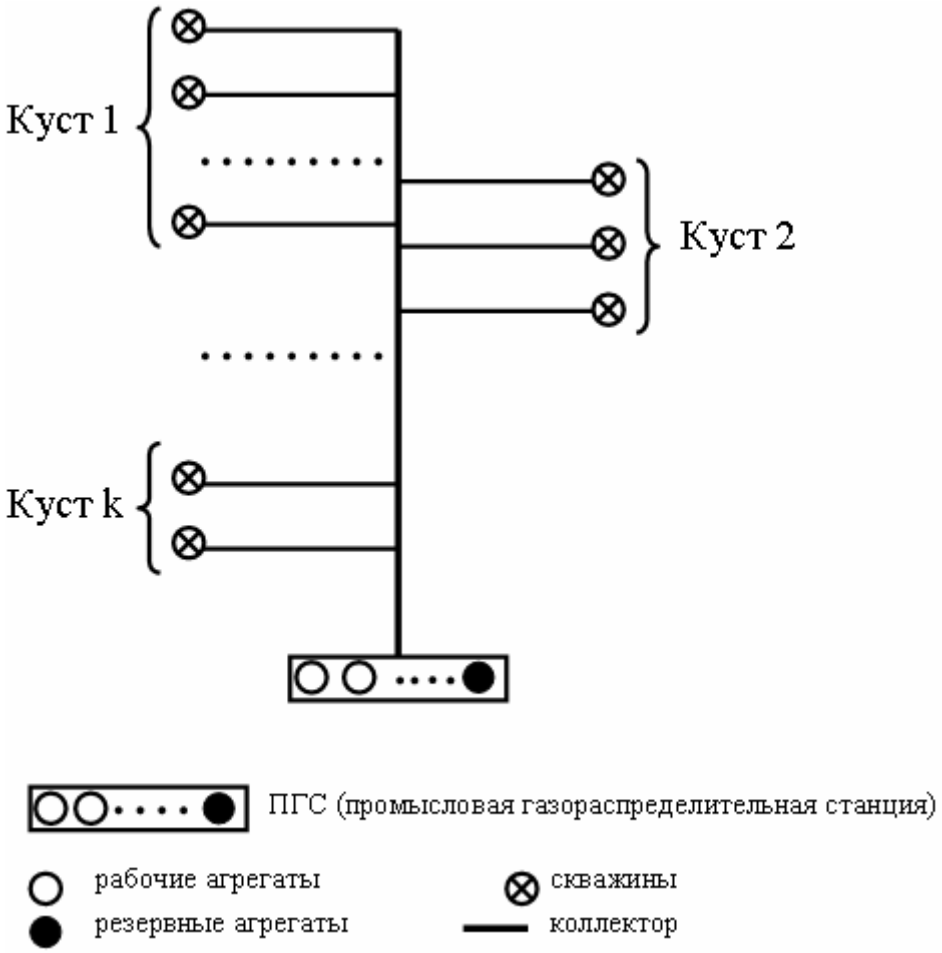


Рис.4. Расчетная схема газового месторождения.

При создании моделей приняты следующие допущения:

– объекты рассматриваются как отдельные элементы системы газоснабжения;

– рассматриваются только стационарные состояния объектов, переходные режимы не рассматриваются;

– не учитываются возможности резервирования за счет: аккумулирующей способности концевых участков МГ; форсированной работы ГПА в пределах технических ограничений; использования перемычек на линейной части (при этом следует учитывать, что указанные выше возможности могут приводить только к повышению надежности объектов, а не к ее снижению);

– среди свойств надежности объектов учитывается их безотказность и ремонтопригодность (причем потоки отказов и восстановлений оборудова-

ния (ГПА, ниток трубопроводов и скважин) являются простейшими потоками);

– учитываются только аварийные отключения оборудования и текущие ремонты в виде профилактики.

Анализ надежности осуществляется следующим образом:

1. Исходная система, состоящая из неоднородных элементов, путем эквивалентирования заменяется системой, состоящей из однородных элементов.

2. Для звеньев (линейных участков, КС или скважин и т.д.) определяются функции распределения вероятностей производительностей работоспособных состояний, при этом используется аналитический метод на уровне случайных марковских процессов, основанный на схеме «гибели и размножения».

3. Осуществляется композиция функций распределения производительности газотранспортной системы, месторождения или ПХГ в целом.

4. В результате анализа надежности МГ, месторождения газа или ПХГ находятся:

- конечные ряд и функция распределения вероятностей работоспособного состояния производительности объектов (кривая 1 и 2 рис. 5) $P(Q)$ и $F(Q)$, по которым определяются интегральные показатели надежности, такие как математическое ожидание, дисперсия и среднеквадратичное отклонение производительности за рассматриваемый интервал времени

$$M[Q] = \sum_{i=1}^{I} Q_i q_i[Q_i],$$

$$D[Q] = \sum_{i=1}^{I} (Q_i - M[Q])^2 \cdot q_i[Q_i],$$

$$\sigma[Q] = \sqrt{D[Q]},$$

где $q_i[Q_i]$ – вероятность работоспособного состояния производительности технических объектов Q_i, $i = \overline{1, I}$; I – количество членов в ряду распределения;

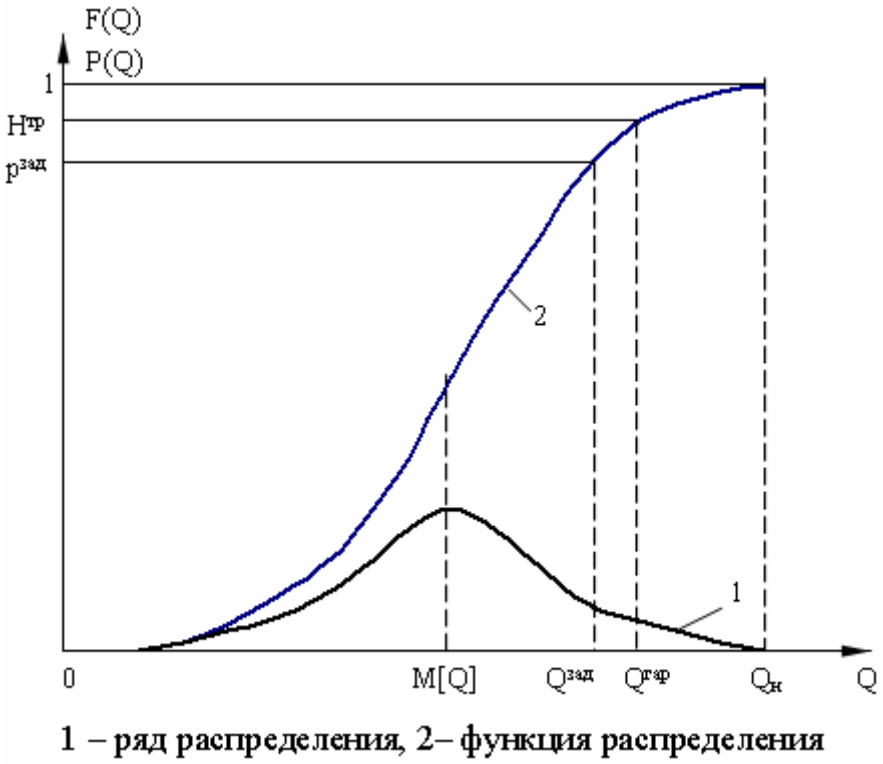


Рис. 5. Закон распределения вероятностей
пропускной способности МГ

- коэффициент надежности $K_{\text{н}}$

$$K_{\text{н}} = \frac{M[Q]}{Q_{\text{н}}}$$

где $Q_{\text{н}}$ – номинальное значение производительности МГ;

- средний процент потерь производительности, связанных с отказами основного оборудования

$$K_{\text{п}} = (1 - K_{\text{н}}) \cdot 100\%;$$

- вероятность обеспечения пропускной способности МГ или производительности месторождения (ПХГ) на уровне не ниже заданной $Q^{\text{зад}}$ (см. рис.5)

$$p^{\text{зад}} = F(Q = Q^{\text{зад}}) = \sum_{i=1}^{I} q_i[Q_i];$$

- величина производительности $Q^{\text{гар}}$, гарантирующая требуемую надежность газоснабжения потребителей $H^{\text{тр}}$ (см. рис. 5),

$$H^{\text{тр}} = F(Q^{\text{гар}}).$$

Алгоритмическими особенностями данных моделей анализа надежности МГ, месторождения газа и ПХГ являются: 1) учет разнородного состава элементов, формирующих рассматриваемые объекты; 2) учет всех видов резервов (нагруженного, ненагруженного и облегченного); 3) оригинальный алгоритм для сокращения числа расчетных состояний на стадии композиции функций распределения производительности.

Модель анализа надежности сложной системы газоснабжения

Модель анализа надежности сложной системы газоснабжения [5,6] используется для решения задач функционирования и развития систем газоснабжения на 3-ем иерархическом уровне.

Объектом исследования является многоузловая газоснабжающая система, охватывающая месторождения и другие источники газа, ПХГ и узлы потребления газа (с категориями потребителей), и связанная системой магистральных газопроводов. Рассматриваются следующие категории потребителей: 1) коммунально-бытовая нагрузка, зависящая в большой степени от отклонений температур наружного воздуха в отопительный период от среднемноголетних; 2) промышленность и котельные и 3) электростанции. В спросе на газ можно выделить случайную и детерминированную составляющие, куда условно можно отнести первую и две последние категории, соответственно.

Максимальный расчетный интервал – один сезон года (зима или лето). Это связано с тем, что при относительно кратковременных нарушениях в системе в силу ее инерционности переход ПХГ от режима отбора к режиму закачки газа практически невозможен, и подземные хранилища в зимний период работают как источники, а летом – как потребители газа.

При создании модели вводятся следующие допущения:

– рассматриваются только стационарные режимы;

– расчет режимов предусматривает центральный орган управления, главная цель которого максимальное удовлетворение спроса (это допущение

принято потому, что система газоснабжения является естественной монопольной структурой);

– поставка газа в систему от месторождений и ПХГ (в общем случае от источников газа) и его транспортировка по МГ зависят от состояний оборудования;

– спрос на газ задается для каждого узла потребления в соответствии с его категорией;

– спрос коммунально-бытовых потребителей (I категория) удовлетворяется полностью, затем обеспечиваются промышленные потребители и котельные (II категория), и в последнюю очередь электростанции (III категория);

– ПХГ в течение рассматриваемого периода работают либо на закачку, либо на отбор газа;

– учитывается запас газа в ПХГ на начало рассматриваемого периода;

– объем резервного топлива задается суммарно на каждый узел потребления.

В модели учитывается вероятностный характер спроса на газ коммунально-бытовых потребителей (I категория); пропускной способности МГ, объема поставки газа в систему от месторождений, объема отбора в систему или закачки газа в ПХГ. Три последних фактора зависят от состояния соответствующего оборудования и должны быть определены с учетом аварийных и плановых простоев.

Спрос коммунально-бытовых потребителей аппроксимируется нормальным законом распределения, остальные случайные величины могут иметь произвольный закон распределения, который определяется по описанным выше моделям оценки надежности магистрального газопровода, месторождения и подземного хранилища газа.

Начальный запас газа в ПХГ, объем резервного топлива в узлах потребления и спрос промышленности, котельных и электростанций (II и III категории), задаются детерминировано.

В содержательном плане модель позволяет по заданным в вероятностной форме спросу на газ и поставке газа в систему от месторождений, а также отбору в систему или закачке газа в подземные газохранилища и с учетом его запасов в ПХГ, а также с учетом пропускных способностей магистральных газопроводов, заданных также в вероятностной форме, и с учетом потерь газа на собственные нужды на месторождениях, при его хранении в ПХГ и при его транспортировке по МГ, а также с учетом взаимозаменяемости топлив определить: закон распределения небалансов газа по узлам потребления; основные показатели надежности, а именно: надежность газоснабжения как вероятность удовлетворения спроса, математическое ожидание недоотпуска газа и коэффициент относительной обеспеченности потребителей газом; мероприятия по уменьшению или увеличению надежности удовлетворения спроса; функции распределения потоков газа по МГ, работы месторождений и ПХГ.

В соответствии с принятыми допущениями и расчетными условиями задачи рассматривается газоснабжающая система в виде модельной сети, в которой узлы $i \in R$ соответствуют месторождениям $(i \in R_1)$, потребителям $(i \in R_2)$, подземным газохранилищам $(i \in R_3)$, точкам сопряжения трубопроводов $(i \in R_4)$. Дуги $(i, j) \in U$ сети отвечают газопроводам, соединяющим узлы i и j.

Рассматривается расчетный интервал (максимально – зима или лето), для которого и приведены все исходные характеристики.

Для каждого расчетного узла - месторождения $i \in R_1$ заданы: ряд распределения возможной поставки газа в систему $q^{\text{м}}[x^{\text{м}}]$; удельные затраты на добычу газа $C_i^{\text{м}}$; коэффициент, учитывающий расход газа на собственные нужды $\lambda_i^{\text{м}}$, $\lambda_i^{\text{м}} < 1$.

Для каждого расчетного узла - потребителя $i \in R_2$ заданы: спрос на газ коммунально-бытовых потребителей в виде м.о. – $M[X_i^I]$ и с.к.о. –

$\sigma[X_i^I]$ закона нормального распределения; спрос на газ промышленных потребителей и котельных X_i^{II}; спрос электростанций, работающих на газе, X_i^{III}; удельные стоимости соответственно для I, II и III категорий потребителей C_1, C_2 и C_3, причем $C_1 << C_2 << C_3$; максимальный объем резервного топлива B_i и затраты на его единицу C_{0i}.

Для каждого узла - ПХГ $i \in R_3$ заданы: ряд распределения либо максимально возможного отбора газа в систему $q^{\text{ПХГ}}[x_i^+]$ из ПХГ, либо максимально возможной закачки газа в ПХГ $q^{\text{ПХГ}}[x_i^-]$; удельные затраты на отбор C_i^+ или закачку газа C_i^-; вместимость хранилища V_i; запас газа на начало рассматриваемого периода S_i; коэффициент, учитывающий потери газа при его хранении λ_i^+, $\lambda_i^+ < 1$.

Для каждого магистрального газопровода $(i, j) \in U$ заданы: ряд распределения пропускной способности $q^{\text{МГ}}[x_{ij}]$; удельные затраты на транспорт газа от узла i до узла j C_{ij}; коэффициент, учитывающий потери газа при его транспортировке λ_{ij}, $\lambda_{ij} < 1$.

Заданной считается также требуемая (нормативная) надежность удовлетворения спроса потребителей на газ – $H^{\text{тр}}$.

Алгоритм решения задачи включает:

1. *Вероятностный блок.* Здесь используется метод статистических испытаний (метод Монте-Карло), с помощью которого по заданным рядам распределения разыгрываются различные значения случайных величин.

2. *Блок расчета режима функционирования системы*. Здесь используется метод оптимизации потокораспределения газотранспортных системах (модифицированный метод Басакера-Гоуэна), который позволяет учитывать первый закон Кирхгофа, соответствующие ограничения на производительность трубопроводов, месторождений, хранилищ и на подачу газа потребите-

лям, которая не должна превышать расчетный спрос, а также потери газа на стадии добычи, хранения и транспортировки и взаимозаменяемость топлив.

3. *Блок вычисления показателей надежности*. Здесь используются методы теории вероятностей: теоремы о сложении и умножении вероятностей различных событий и методы вычисления функций распределения небалансов газа, по которым и определяются основные показатели надежности.

Структурная схема алгоритма задачи оценки надежности функционирования газоснабжающей системы приведена на рис. 6, где блоки 1 – 3 содержат исходную информацию, а остальные - промежуточные и итоговые результаты расчета.

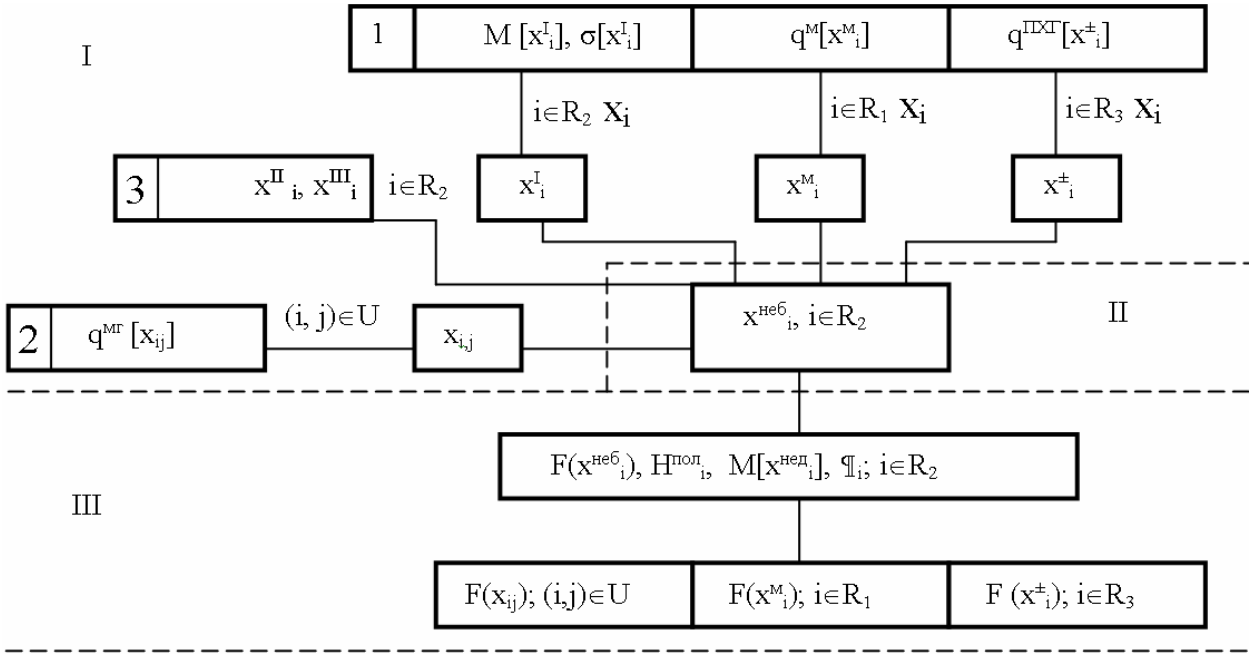


Рис. 6. Структурная схема задачи оценки надежности функционирования системы газоснабжения

Для построения функции распределения небалансов газа $F[x_i^{\text{неб}}]$ по узлам-потребителям используется метод статистических испытаний, т.е. по заданным для данного расчетного интервала рядам распределения спроса на газ коммунально-бытовых потребителей, возможной поставки газа от месторождений и ПХГ в систему (или закачки в ПХГ), пропускной способности МГ статистически разыгрываются случайные состояния этих значений, а именно X_i^I, $X_i^{\text{м}}$, X_i^-, $i \in R$, а также X_{ij}, $(i,j) \in U$.

Расчет оптимального режима сводится к определению следующих величин: $x_i^{\text{м}}$ – объем добычи газа на i-ом месторождении (источнике); x_{0i} – объем поставки резервного топлива i-ому потребителю; $x_i^{\text{уд}}$ – удовлетворенный объем поставки газа, который складывается из объемов поставки всем категориям i-ого узла потребления, $x_i^{\text{уд}} = x_i^{I} + x_i^{II} + x_i^{III}$; x_i^{-}, x_i^{+} – объемы газа при закачке или отборе газа из i-ого ПХГ; x_{ij} – потоки газа по соответствующему газопроводу (i, j).

Формализованная постановка задачи расчета оптимального режима приобретает вид

$$\sum_{i\in R_1} C_i^{\text{м}} x_i^{\text{м}} + \sum_{i\in R_2} C_{oi} x_{oi} + \sum_{i\in R_3} C_i^{-} x_i^{-} + \sum_{(i,j)\in U} C_{ij} x_{ij} + \sum_{i\in R_2} y_{oi} x_i^{\text{деф}} \Rightarrow \min$$

при ограничениях

$$\left.\begin{array}{l} \sum\limits_j \lambda_{ji} x_{ji} - \sum\limits_j x_{ij} + \lambda_i^{\text{м}} x_i^{\text{м}} = 0, \\ 0 \le x_i \le X_i \end{array}\right\},\ i \in R_1;$$

$$\left.\begin{array}{l} \sum\limits_j \lambda_{ji} x_{ji} - \sum\limits_j x_{ji} + x_{oi} - x_i^{\text{уд}} = 0, \\ \text{где} \quad x_i^{\text{уд}} = X_i^{\text{I}} + X_i^{\text{II}} + X_i^{\text{III}} - x_i^{\text{деф}}, \\ 0 \le x_i^{\text{I(II,III)}} \le X_i^{\text{I(II,III)}}, \\ 0 \le x_{oi} \le B_i, \\ 0 \le x_i^{\text{деф}} \le X_i^{\text{I}} + X_i^{\text{II}} + X_i^{\text{III}} \end{array}\right\},\ i \in R_2$$

$$\left.\begin{array}{l} \sum\limits_j \lambda_{ji} x_{ji} - \sum\limits_j x_{ji} + \lambda_i^{+} x_i^{+} = 0, \\ 0 \le x_i^{+} \le \min\{X_i^{+}, S_i\}, \\ 0 \le x_i^{-} \le \min\{X_i^{-}, V_i - S_i\}, \end{array}\right\}, i \in R_3;$$

$$\sum_j \lambda_{ji} x_{ji} - \sum_j x_{ji} = 0,\ i \in R_4;$$

$$0 \le x_{ij} \le X_{ij},\ (i, j) \in U,$$

где $x_i^{\text{деф}}$ – суммарные дефициты газа по всем категориям i-ого узла - потребителям, y_{oi} - удельные ущербы от недоотпуска газа по узлам.

В качестве критерия рассматривается минимум затрат на доставку газа потребителям и м.о. ущерба от недоотпуска газа по узлам. Ограничения в виде равенств представляют собой балансы газа соответствующих узлов, а остальные ограничения заданы в виде двухсторонних неравенств. Потери газа при хранении учитываются только при его отборе из ПХГ, а потери газа в газопроводах учитываются один раз в конце связи.

Состояния системы вычисляются последовательно по одному, при этом каждое следующее расчетное состояние формируется после того, как в блоке расчета режима функционирования системы проанализировано предыдущее состояние.

Окончание расчетов осуществляется либо когда величина относительной погрешности статистического моделирования станет ниже заданной, либо число испытаний достигнет заданного предельного значения.

Особенностью построения рядов распределения небалансов газа является то, что получаемые после расчета каждого состояния системы значения вероятностей $q[x_i^{\text{неб}}]$ суммируются с соответствующими значениями предыдущих состояний. Для того, чтобы получить ряды распределения в "чистом виде", необходимо полученные значения разделить на число проведенных испытаний, т.е. вычисляются "статистические вероятности".

Результирующие ряды распределения небалансов газа содержат в себе отрицательные значения, представляющие собой избытки газа, и положительные значения, представляющие дефициты газа.

Имея для каждого расчетного узла-потребителя $i \in R_2$ ряд распределения вероятностей небаланса газа, можно построить и соответствующую функцию распределения $F[x_i^{\text{неб}}]$ (рис. 7), с помощью которой определяются:

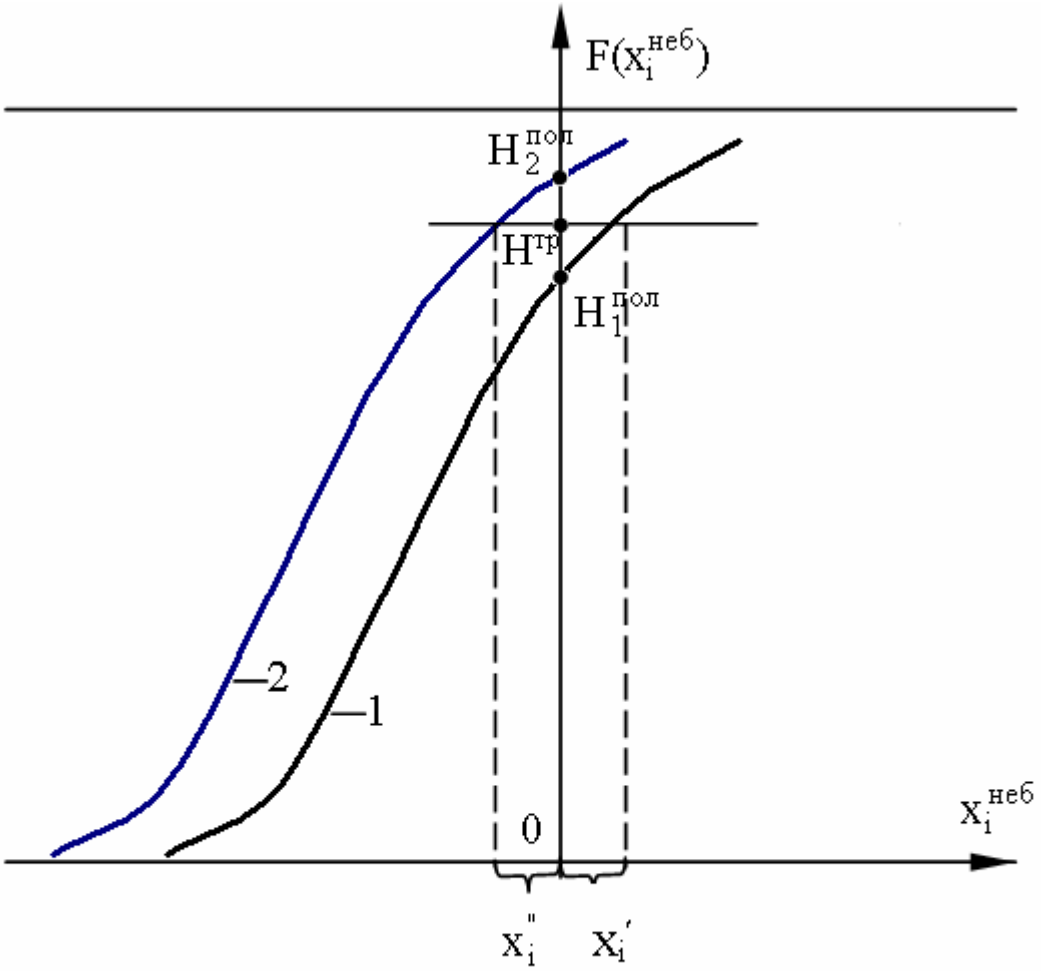
- надежность газоснабжения потребителей как вероятность удовлетворения спроса

$$H_i^{\text{пол}} = F(x_i^{\text{неб}} = 0);$$

- математическое ожидание недоотпуска газа

$$M[X_i^{\text{нед}}] = \sum_{k=1}^{K} x_k^{\text{неб}} q_k[x_k^{\text{неб}}]$$

при $x_k^{\text{неб}} > 0$ или $F(x_i^{\text{неб}}) > H_i^{\text{пол}}$,



1 – при дефиците газа, 2 – при избытке газа

Рис. 7. Функции распределения небалансов газа

где $q_k[x_k^{\text{неб}}]$ – вероятность небаланса газа $x_k^{\text{неб}}, k = \overline{1,K}$; K - количество положительных членов в ряду распределения;

- коэффициент относительного обеспечения потребителей газом

$$\pi_1 = 1 - M[X_i^{\text{нед}}] / M[X_i^{\text{спр}}],$$

где

$$M[X_i^{\text{спр}}] = M[X_i^{\text{I}}] + X_i^{\text{II}} + X_i^{\text{III}};$$

- в зависимости от соотношения полученной надежности $H_i^{\text{пол}}$ и требуемой $H^{\text{тр}}$ предлагаются различные мероприятия по ее регулированию.

При $H^{\text{тр}} > H_i^{\text{пол}}$ (рис. 7) возможно уменьшение спроса на потребление газа, т.е. введение плановых ограничений (лимита $\Delta x_i^{\text{спр}}$).

На практике плановые ограничения реализуются в виде лимитов для потребления газа электростанциями и промышленными предприятиями. Как уже отмечалось выше, это потребление условно составляет детерминированную часть спроса. При введении лимитов можно считать, что исходная суммарная функция распределения спроса $X_i^{\text{спр}}$, включающая в себя функцию распределения коммунально-бытовых потребителей X_i^{I} и детерминированные значения спроса промышленных предприятий и котельных X_i^{II} и электростанций X_i^{III}, не меняется, а лишь сдвигается влево.

При этом, естественно, не меняется и вид функции распределения небалансов газа, а происходит лишь сдвиг ее влево на величину изменения (сдвига) спроса.

Отсюда, имея функцию распределения небалансов газа $F(x_i^{\text{неб}})$, можно определить величину планируемого лимита на спрос, удовлетворяющего требуемой надежности

$$\Delta x_i^{\text{спр}} = x_i^{'} \quad \text{при} \quad F(x_i^{'}) = H^{\text{тр}}.$$

Можно также рассмотреть возможность увеличения резервов потребителя на величину $\Delta x_i^{\text{спр}}$.

При $H_2^{\text{пол}} > H^{\text{тр}}$ возможно снижение поставки газа в систему от месторождений или ПХГ до величины, обеспечивающей требуемую надежность. Из рис. 6 очевидно, что такое снижение определяется как

$$\Delta x_i = -x_i^{''} \quad \text{при} \quad F(x_i^{''}) = H^{mp}.$$

Здесь также можно рассматривать возможность создания запасов газа в ПХГ, либо уменьшения резервов потребителя на соответствующую величину.

При реализации различных мероприятий по уменьшению или увеличению надежности соответствующего узла - потребителя может измениться надежность соседних узлов и, чтобы ее оценить, требуется провести следующий расчет по модели.

Алгоритмическими особенностями данной модели анализа надежности сложных газоснабжающих систем по сравнению с известными являются: 1) использование комбинации аналитических методов (схемы "гибели и размножения") для получения соответствующих исходных рядов распределения и метода статистических испытаний для моделирования различных состояний системы на их основе; 2) вероятностный учет большего числа случайных факторов: помимо вероятностного учета работы трубопроводной сети, здесь учитывается аварийность оборудования на месторождениях и ПХГ, а также неравномерность спроса на газ.

Модель определения оптимальных параметров газотранспортной системы

Модель определения оптимальных параметров газотранспортной системы [7] является моделью синтеза структурной надежности и используется при решении задач развития систем газоснабжения 4-ого уровня иерархии.

Процесс выбора оптимальных параметров МГ предполагает:

1. Вариантное рассмотрение способов развития на перспективу газотранспортной системы.
2. Анализ надежности данной системы.
3. Оптимальный выбор рационального варианта на основе расчета технико-экономических характеристик и интегральных показателей надежности.

В общем виде задача формулируется следующим образом: исходя из среднесуточной производительности $МГ(Q)$, его технико-технологических

(T), надежностных (H) и технико-экономических показателей $(Э)$, принципиальной схемы МГ и избыточных конечных способов резервирования (r) определить оптимальные технические параметры, которые соответствовали бы экстремуму (максимуму или минимуму) критерия экономической эффективности при условии удовлетворения заданной нормы надежности газоснабжения

$$З = f(T, H, Э, r) \to ext,$$

$$П = (Q, H, r) \geq П_{н.}$$

В качестве основных критериев экономической эффективности сопоставляемых вариантов может использоваться либо *минимум приведенных затрат*, либо *максимум чистого дисконтированного дохода* проекта и внутренняя норма возврата, которые позволяют детально учитывать динамику движения денежных средств в течение всего периода реализации проекта.

Сформулированная задача относится к классу задач целочисленного программирования и может рассматриваться как комбинаторная оптимизационная задача. Инженерный опыт исследований показывает, что число вариантов развития газопровода сравнительно небольшое, и все они могут быть просмотрены обычным перебором.

Более подробно рассмотрим вторую модификации модели, где в качестве критерия рассматривается чистый дисконтированный доход проектируемого МГ. Он определяется с учетом следующих допущений.

1. Доход предприятия образуется от продажи газа на выходе МГ.

2. Текущие издержки включают в себя эксплуатационные расходы: в линейную часть МГ; пропорциональные установленные мощности КС и ГПА; не зависящие от установленной мощности КС.

3. Капиталовложения обеспечиваются за счет кредитов. За их счет могут оплачиваться текущие расходы, если величина прибыли не покрывает их. Погашение кредитов начинается только тогда, когда появляется прибыль.

4. Налоги начисляются только на прибыль. Добавочная стоимость не рассматривается. Налог на себестоимость учитывается в постоянной составляющей эксплуатационных издержек.

5. Амортизационные отчисления не облагаются налогом на прибыль и поступают на счет предприятия. Если счет предприятия положительный, то он рассматривается как вклад в банке, на который производятся ежегодные начисления. Если счет предприятия отрицательный, т.е. имеется долг, то этот долг в конце каждого года увеличивается на процент за кредит.

6. Возврат кредитов и процентов по ним производится из чистой прибыли.

С учетом принятых допущений чистый дисконтированный доход NPV проектируемого газопровода определяется следующим образом:

$$NPV = \sum_{t=T_c}^{T_в} C_t Q_{гt} D_t - \sum_{t=T_c}^{T_в} K_{лт} l_o I_k M\, D_t - \sum_{t=T_c}^{T_в} (K_t + K_{ct} I_{уст}) N\, D_t -$$

$$\sum_{t=T_c}^{T_в} [K_{лт}^{эк} l_o I_k M + (K_t^{эк} + K_{ct}^{эк} I_{уст}) N] a_n D_t - \sum_{t=T_c}^{T_в} H_t D_t - \sum_{t=T_c}^{T_в} P_t D_t + \sum_{t=T_c}^{T_в} S_t D_t - \sum_{t=T_c}^{T_в} R_t D_t$$

где индексом t обозначены величины, относящиеся к t-ому году, $t = \overline{T_c, T_в}$; T_c – год начала строительства МГ, T_1 – срок строительства; $T_в$ – последний год эксплуатации МГ, T_2 – срок эксплуатации; $T_1 + T_2 = T_в - T_c + 1$; $Q_{гt}$ – годовая расчетная производительность МГ, млн.м^3; C_t – цена газа на выходе МГ, тыс. долл. за 1 млн.м^3 газа; $D_t = \dfrac{1}{(1+d_t)^t}$ – индекс дисконтирования; d_t – коэффициент дисконтирования; $K_{лт}$ – удельные годовые капиталовложения в линейную часть МГ, тыс.долл./км; K_{ct} – удельные ежегодные капиталовложения, пропорциональные установленной мощности КС тыс. долл./ГПА; K_t – годовые капиталовложения, не зависящие от установленной мощности КС; $K_{лт}^{эк}$ – суммарные капиталовложения во введенную в эксплуатацию к началу t-го года линейную часть МГ (основные фонды), тыс.долл./км; $K_{ct}^{эк}$ – суммарные удельные капиталовложения во введенное в эксплуатацию к началу t-го года

оборудование компрессорных станций, пропорциональные их установленной мощности (основные фонды), тыс. долл./ГПА; $K_t^{эк}$ – суммарные капиталовложения во введенное в эксплуатацию к началу t-го года оборудование КС, не зависящие от их установленной мощности (основные фонды), тыс. долл.; a_n – доля постоянной части эксплуатационных издержек от капиталовложений, учитывающая амортизацию оборудования, капитальные и текущие ремонты и общемагистральные издержки, а также включающая в себя оплату труда обслуживающего персонала, энергии, материалов и налог на себестоимость; H_t – годовые налоги; P_t – годовые расходы на погашение кредита; S_t – годовые начисления на размещенный в банке положительный счет предприятия (МГ); R_t – годовые начисления за полученный кредит.

Первое слагаемое в выражении (1) представляет собой дисконтированный доход от продажи газа на выходе МГ за все годы его функционирования.

Второе слагаемое представляет собой дисконтированные капитальные вложения в линейную часть МГ. Здесь l_o – длина линейного участка, I_k – число ниток трубопроводов проектируемого МГ, M – число линейных участков.

Третье слагаемое представляет собой дисконтированные капитальные вложения в КС МГ, где $I_{уст}$ – число установленных ГПА; N – количество КС по трассе МГ.

Четвертое слагаемое - дисконтированные текущие издержки по МГ.

Последующие четыре слагаемые представляют собой дисконтированные налоги, расходы на погашение кредита, начисления на размещенный в банке положительный счет предприятия и начисления за полученный кредит за все годы строительства и функционирования МГ соответственно.

Годовые налоги определяются из выражения

$$H_t = \{C_t Q_{гt} - [K_{лt}^{эк} l_o I_к M + (K_t^{эк} + K_{ct}^{эк} I_{уст})N]a_n\}a_{пр},$$

где $a_{пр}$ – доля налога на прибыль.

Годовые расходы на погашение кредита, если он еще не выплачен, определяются из выражения:

$$P_t = \begin{cases} O, \ если\, \Phi_t \le 0, \\ \Phi_t, если\, \Phi_t < \delta_t, \\ \delta_t, если\, \Phi_t \ge \delta_t, \end{cases}$$

где Φ_t – финансовое состояние текущего счета предприятия на начало года t, δ_t – максимальная сумма, которая может быть направлена на погашение долгов по кредиту, определяется из выражения

$$\delta_t = \{C_t Q_{гt} - [K_{лт} l_o I_к M + (K_t^{эк} + K_{ct}^{эк} P_{уст})N]a_n - H_t\}a_{пк},$$

$a_{пк}$ – максимальная доля чистой прибыли, которая может быть направлена на погашение долгов по кредиту.

Финансовое состояния Φ_t на начало года t и определяется следующим образом

$$\Phi_1 = -K_1,$$

$$\Phi_t = \Phi_{t\text{-}1} + C_{t\text{-}1} Q_{г(t\text{-}1)} - K_{лт} l_o I_k M\text{-}(K_t + K_{ct} P_{уст})N +$$

$$+ [K_{л(t\text{-}1)}^{эк} \cdot l_o I_k M + (K_{t\text{-}1}^{эк} + Kc_{(t\text{-}1)}^{эк} P_{уст})N](a_a - a_n) - H_{t\text{-}1} - P_{t\text{-}1} + S_{t\text{-}1} - R_{t\text{-}1},$$

где K_1 - капиталовложения в МГ в 1-й год строительства; a_a – доля амортизационных отчислений от стоимости действующего оборудования (основных фондов).

Начисления на вклад в банке определяются как

$$S_t = \begin{cases} 0, eесл\, \Phi_t \le 0 \\ \Phi_t \beta_{1t}, eесл\, \Phi_t > 0, \end{cases}$$

где β_{1t} – доля начислений на депозит.

Начисления за полученный кредит определяются как

$$R_t = \begin{cases} 0, \ если\ \Phi_t \ge 0 \\ \Phi_t \beta_{2t}, \ если\ \Phi_t < 0 \end{cases}$$

где β_{2t} – доля годовых начислений за кредит.

Внутренняя норма возврата IRR определяется подбором коэффициента дисконтирования d_t, одинакового для всех лет расчетного периода, при котором $NPV = 0$, т.е. при $NPV = 0$ (с какой-то погрешностью) $IRR = d_t$.

После того как будут проанализированы все возможные варианты конструктивного оформления проектируемого МГ, определяется величина либо минимальных дисконтированных затрат, либо максимального чистого дисконтированного дохода и выдаются оптимальные параметры МГ, соответствующие заданному критерию экономической эффективности.

В результате решения данной задачи определяются следующие параметры МГ: число ниток трубопроводов; соответствующие оптимальные диаметры; число КС; число линейных участков; длина линейного участка; число рабочих и резервных ГПА на каждой КС; оптимально-номинальная мощность ГПА; металловложения в линейную часть МГ.

Алгоритмической особенностью данной модели синтеза структурной надежности является определение рациональных параметров проектируемого магистрального газопровода на основе сочетания оценки надежности и использования новых экономических критериев эффективности сопоставляемых вариантов, таких как чистый дисконтированный доход и внутренняя норма возврата.

На основе предложенных математических моделей были разработаны соответствующие программные средства. Для расчета интегральных показателей надежности сложного магистрального газопровода, газового месторождения и подземного хранилища газа на основе схемы “гибели и размножения” были разработаны вычислительные программы **«ANAS», «ANAM»** и **«ANAX»,** соответственно.

Для определения оптимальных параметров проектируемых МГ были разработаны вычислительные программы «**SINTEZ**» и для второй модификации – «**SINT**».

Для анализа надежности функционирования сложной закольцованной системы газоснабжения была разработана ВП **«ANES-1»,** а также единый программно-вычислительный комплекс **«ANES-2»,** объединяющий выше перечисленные программы **«ANAS», «ANAM»** и **«ANAX»,** с помощью которых определяются законы распределения вероятностей производительностей ма-

гистральных газопроводов, месторождений газа и подземных хранилищ газа, являющимися исходными данными для ВП **«ANES-1»,** с последней.

Эти средства использовались для оптимизации технических параметров и участков газотранспортной сети на Востоке России [9]. Расчеты проводились для первого этапа формирования газотранспортной системы на период до 2020 г. на базе освоения Ковыктинского газоконденсатного месторождения, в частности для экспортного МГ Иркутск – Пекин (через Улан-Батор).

Однако следует отметить, что разработанные программно-вычислительные средства были созданы в 90-х гг., к настоящему моменту морально устарели (не реализовано деление программ на независимые модули, не проводился рефакторинг, не составлена документация разработчика и т.д.).

Модель синтеза надежности сложной газоснабжающей системы

Модель синтеза надежности сложной газоснабжающей системы является постановочной, предлагается использовать при решении задач развития систем газоснабжения на 3-ем иерархическом уровне.

Для нахождения оптимальной надежности газоснабжающей системы предлагается двухэтапный методический подход на основе анализа и синтеза её надежности, при котором решаются следующие задачи:

1. Определение эквивалентных характеристик надежности объектов системы газоснабжения.

2. Оптимизация средств резервирования газоснабжающей системы на уровне оптимальной надежности.

При этом исходим из условия, что к этому моменту должна быть решена сетевая потоковая задача и определенны рациональные объёмы добычи газа в газодобывающих центрах, а также объёмы и направления межрайонных потоков газа. Данное решение должно быть детализировано в модели оптимизации сезонного газопотребления, и оно должно являться основной исходной информацией двухэтапного подхода для решения задач модели анализа и синтеза надежности разветвленной системы газоснабжения.

Место модели синтеза надежности ГСС в иерархии моделей оптимального развития газоснабжающей системы и их взаимосвязи показаны на рис. 8.

1 этап. Эквивалентные характеристики надежности для газотранспортных и газодобывающих предприятий представляют собой зависимости интегральных показателей надежности (математических ожиданий фактической производительности) (рис. 9, а) и дисконтированных издержек (рис. 9, б) от установленной производительности существующих объектов с имеющимся резервированием и установленной производительности дополнительных средств резервирования в эти объекты.

Рассматриваемые функции нелинейные, однако, практический опыт расчетов показывает возможность их линеаризации с относительно небольшой погрешностью в следующих границах

$$o \le x \le d; \quad d \le y \le d^r,$$

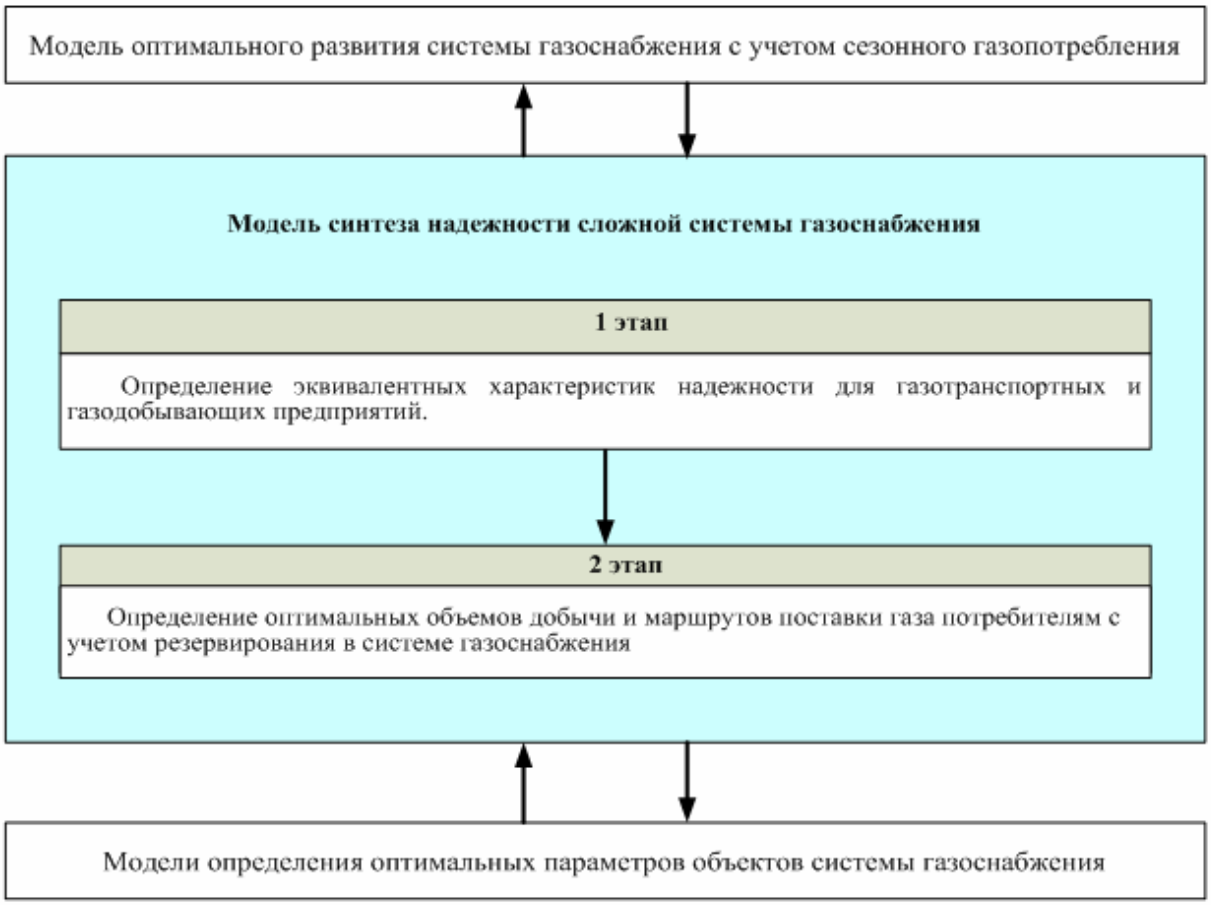


Рис. 8. Модель синтеза надежности в иерархии моделей оптимального развития ГСС.

где d – верхняя граница производительности объектов (взята из результатов решения задачи развития системы газоснабжения с учетом сезонности газопотребления); d^r – верхняя граница производительности с учетом заданных резервов упомянутых объектов (величина резервов задается предварительно).

Тангенсы углов наклона ломаной линии (λ) и (π) показывают коэффициенты надежности объекта с учетом расхода газа на собственные нужды и потери из-за ненадежности и с учетом дополнительного резервирования.

Для расчета указанных показателей используется описанные выше математические модели анализа и синтеза надежности объектов ГСС, а именно: газодобывающих или газотранспортных предприятий. На их основе определяется конечный закон распределения вероятностей работоспособного состояния объекта, по которому и рассчитываются искомые показатели надёжности, в частности, математические ожидания фактической производительности. Эти показатели определяются при разных способах резервирования объектов.

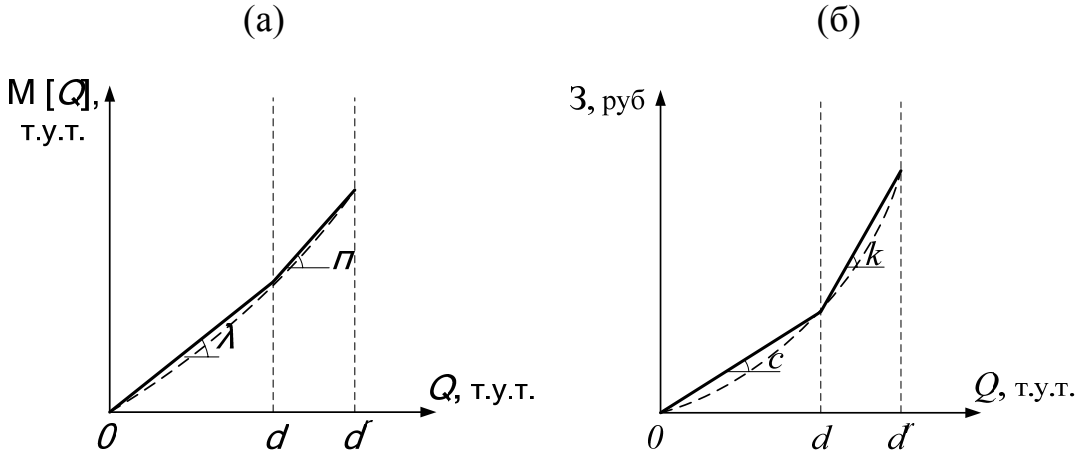


Рис. 9. Эквивалентные характеристики надежности

а) зависимость математического ожидания фактической производительности объекта от его установленной производительности;

б) зависимость дисконтированных издержек от установленной производительности объекта.

Нелинейная зависимость приведенных затрат от производительности существующих объектов с имеющимся резервированием и производительно-

сти дополнительных средств резервирования в эти объекты (рис. 2, б) по своей сути является функцией предложения с учетом резервирования, и ее также возможно линеаризовать ломаной линией в тех же самых границах. Тангенс угла наклона этой линии (c) показывает удельную величину затрат (цену) добычи или транспорта для существующих объектов с уже имеющимся резервированием, а тангенс угла наклона (k) на участке от d до d^r – удельную величину затрат (цену) на создание дополнительной резервной производительности этого объекта.

Таким образом, полученные показатели λ, π, c, k являются основной исходной информацией для задачи второго этапа, они определяются с помощью эквивалентных характеристик и соответствуют оптимальному уровню надежности.

2 этап. В общем виде задачу второго этапа можно сформулировать следующим образом: исходя из критерия минимума издержек в существующие объекты системы газоснабжения, в создание новых резервных производительностей для этих объектов и в создание запасов резервного топлива (допускающего замену газа у потребителей), найти оптимальное сочетание этих способов резервирования, обеспечивающих в каждом узле расчетной схемы балансы приходящих и уходящих производительностей объектов и другого топлива, с оптимальной надежностью и при заданных ограничениях.

Эта задача может быть записана в следующей форме

$$\sum_{(i,j)\in U}\left(c_{ij}x_{ij}+k_{ij}y_{ij}\right)+p_j z_j \to \min,$$

$$\sum_{i\in\Gamma_j^+}\left(\lambda_{ij}x_{ij}+\pi_{ij}y_{ij}\right)+\alpha_j z_j-\sum_{j\in\Gamma_j^-}x_{ji}=\begin{cases}-Q, & j=s;\\ 0, & j\neq s,t;\\ B, & j=t,\end{cases}$$

$$0\le x_{ij}\le d_{ij};\quad 0\le y_{ij}\le d_{ij}^r-d_{ij};\ 0\le z_j\le Z_j.$$

Первое выражение– это минимизируемая целевая функция издержек. *Второе* условие показывает балансы прихода и ухода производительностей объектов с существующим резервированием (x) и с дополнительными ре-

зервными средствами для этих объектов (y), а также с учетом запасов резервного топлива (z). Для каждого узла j должен соблюдаться баланс приходящих и уходящих производительностей (первый закон Кирхгофа, рис. 10), причем Γ_j^+ – множество узлов, из которых дуги заходят в узел j, а Γ_j^- – множество узлов, в которые заходят дуги, исходящие из узла j. *Третье* выражение показывает двухсторонние ограничения производительности объектов.

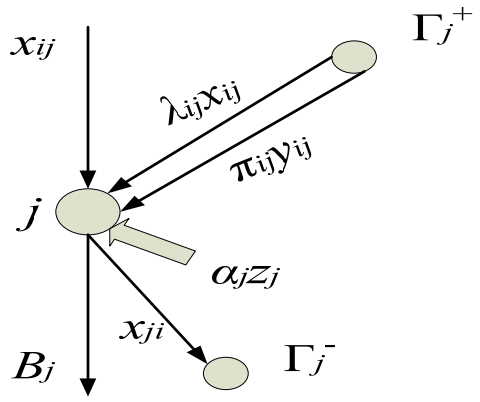


Рис. 10. Узел расчетной схемы.

Каждой дуге (i, j) соответствуют величины: x_{ij} – поток производительности дуги графа, отражающий объект-источник или объект транспорта, с существующим резервированием; y_{ij} – поток дополнительной резервной производительности дуги графа, отражающей объект источника или транспорта; z_j – объем резервного топлива; c_{ij} – удельная величина стоимости с существующим резервированием объекта источника или транспорта газа; k_{ij} – удельная величина дополнительной резервной стоимости объекта источника или транспорта газа; p_j – удельная величина стоимости резервного топлива; λ_{ij} – коэффициент, учитывающий расход газа на собственные нужды и потери из-за ненадежности; π_{ij} – коэффициент надежности дуги графа с учетом дополнительной резервной производительности; α_j – коэффициент надежности топливоснабжения; Q – суммарное значение добываемого газа

всеми его источниками; B – суммарное значение используемого газа всеми потребителями.

Удельные величины стоимости резервного топлива (p_j) для j-х узлов задаются из решения задач верхнего иерархического уровня и показывают стоимость т у.т. мазута или угля, способных компенсировать дефицит газа.

Решение сформулированной линейной задачи позволяет при заданной расчетной схеме, в рассматриваемом сезоне года, найти такой план поставки газа от источников до потребителей, который бы обеспечивал минимум издержек в систему с учетом резервирования. Сформулированная задача может решаться методами линейного программирования.

Выводы

1. Разработана система математических моделей, алгоритмов и вычислительных средств для анализа и синтеза надежности газоснабжения при многоуровневом моделировании, а именно: 1) модель анализа надежности магистрального газопровода; 2) модель анализа надежности месторождения газа; 3) модель анализа надежности подземного хранилища газа; 4) модель оценки надежности сложных газоснабжающих систем; 5) модель определения оптимальных параметров газотранспортной системы.

2. Особенность системы ММ анализа надежности заключается в использовании комбинации аналитических методов на основе схемы "гибели и размножения" и метода статистических испытаний. При таком сочетании методов стал возможен в сравнении с известными моделями вероятностный учет большего числа случайных факторов.

3. Особенность модели определения оптимальных параметров газотранспортной системы заключается в сочетании учета надежности и использования новых экономических критериев эффективности.

4. Разработанная система ММ может использоваться при решении задач развития газоснабжающих систем при их многоуровневом моделировании, однако, для практического использования системы ММ требуется её модернизация на базе современных вычислительных средств.

5. Предложен двухэтапный методический подход для нахождения оптимальной надежности сложной газоснабжающей системы, основанный на определении эквивалентных характеристик надежности её объектов и оптимизации средств резервирования.

Список литературы

[1] Генеральная схема развития газовой отрасли на период до 2030 года. – М.:офиц. изд. Правительства РФ, 2009.– 180 с.

[2] Илькевич Н.И., Меренков А.П. Многоуровневое моделирование и согласование задач развития систем газоснабжения // Методы управления физ.-техн. системами энергетики в новых условиях. – Новосибирск: Наука, Сиб. изд. фирма РАН, 1995. – С. 41 – 45.

[3] Дзюбина Т.В., Илькевич Н.И. Методические вопросы многоуровневого исследования надежности газоснабжения регионов России в условиях приоритетного развития газовой промышленности в Восточном геополитическом направлении // Методические вопросы исследования надежности больших систем энергетики. Вып. № 48. Мурманск, 1996. – С. 61 – 65.

[4] Дзюбина Т.В., Илькевич Н.И. Методические подходы и математические модели для анализа и синтеза надежности при многоуровневом исследовании газоснабжающих систем // Надежность систем энергетики. – Новосибирск: Наука, Сибирское предприятие РАН, 1999. – С. 325 – 333.

[5] Дзюбина Т.В., Илькевич Н.И. Комплексная математическая модель анализа надежности функционирования ЕСГ // Методы управления физ.-техн. системами энергетики в новых условиях. – Новосибирск: Наука, Сиб. изд. фирма РАН, 1995. – С. 181 – 186.

[6] Дзюбина Т.В., Илькевич Н.И. Модель оценки надежности сложных газоснабжающих систем // Известия РАН. Энергетика. 1998. –№6. – С.84–91.

[7] Дзюбина Т.В., Илькевич Н.И. Определение рациональных параметров МГ с учетом надежности и эффективности их инвестиционных проектов // Российский научный симпозиум по энергетике. Методические вопро-

сы исследования надежности БСЭ. – Казань: Казан. Гос. энерг. унив-т, 2001. – Т. IV. – С. 132 – 135.

[8] Дзюбина Т.В., Илькевич Н.И., Ковалев Г.Ф., Лебедева Л.М. Сравнительный анализ методов вычислений функций распределения состояний оборудования в ГСС и ЭЭС // Известия РАН. Энергетика, 1997. – № 1. – С. 169 – 173.

[9] Дзюбина Т.В., Илькевич Н.И. Оптимизация параметров МГ в рыночных условиях с учетом надежности // Известия АН. Энергетика, 2002. – № 3. – С. 93 – 101.

УДК 622.69.019.3+620.9.338.9

ИНФОРМАЦИОННО – ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ЗАДАЧ АНАЛИЗА И СИНТЕЗА НАДЕЖНОСТИ ПРИ МНОГОУРОВНЕВОМ МОДЕЛИРОВАНИИ ГАЗОСНАБЖАЮЩИХ СИСТЕМ

Дзюбина Т. В.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: tvleo@isem.sei.irk.ru

Метелкин А. М.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: AlexMetelkin@narod.ru

Аннотация. Приведено описание существующих программных средств; информационных моделей, используемых при решении задач анализа и синтеза надёжности газоснабжающих систем; обоснована необходимость их модернизации. Описана объектная модель газоснабжающей системы с точки зрения задач надежности. Приведена усовершенствованная модель организации данных, призванная обеспечить универсальность и расширяемость создаваемых средств с учетом интеграции задач при многоуровневом моделировании. Дано описание вновь создаваемых и модернизируемых программных средств.

Ключевые слова. Газоснабжение, моделирование, надежность, модель данных, программное средство.

Задачи исследования надежности. Методическое и информационное обеспечение

В ИСЭМ СО РАН ведутся исследования основных проблем развития газоснабжающих систем (ГСС) в современных условиях: разрабатываются подходы, методы и программные средства для многоуровневого моделирования систем газоснабжения, принципы согласования решений, где особое место отводится учёту свойств надёжности на разных уровнях иерархии [1]. Разработки ориентированы на различные временные и территориальные уровни и позволяют рассматривать ГСС в виде: подсистемы топливно-энергетического комплекса; самостоятельной газовой отрасли; функционально-целостных систем; ЕСГ в целом и региональных систем

газоснабжения; совокупности производственно-технических объектов и подсистем. Общая схема иерархии задач, методов и моделей представлена на рис. 1.

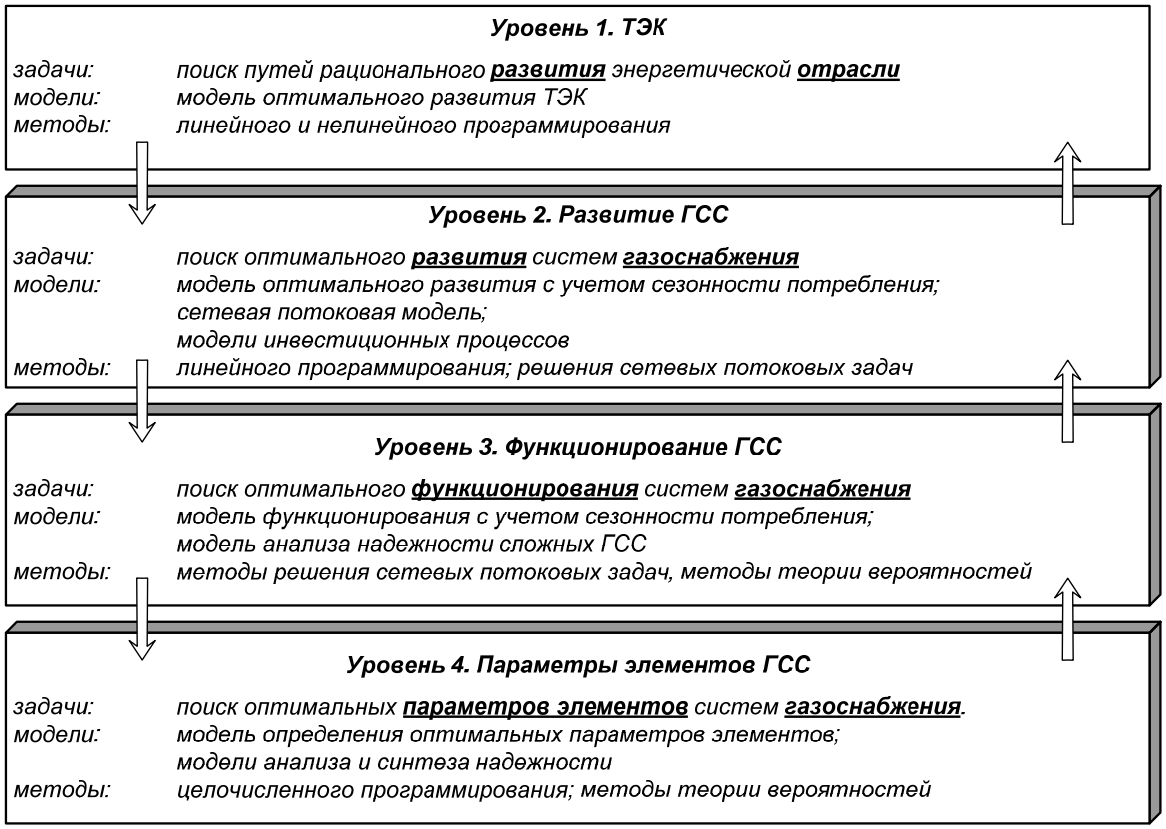


Рис. 1. Иерархия задач, моделей и методов,
применяемых в исследовании ГСС

На первом уровне система газоснабжения представляется в виде блока в общей модели ТЭК, отражающей баланс производства, передачи и потребления различных видов энергии по территории страны. Надежность работы отраслей в модели учитывается в виде нормативов различного рода резервов энергоносителей, которые закладываются в топливно-энергетический баланс, а также затрат самофинансирования отраслей, включающих в том числе затраты на поддержание существующих и вновь создающихся резервов. На втором уровне с учетом интересов отрасли (как естественной монополии) определяются возможные пути развития ЕСГ. Надежность учитывается ценами самофинансирования, которые включают затраты на поддержание существующих и вновь создаваемых резервов. На

третьем уровне с помощью квазидинамической модели варианты развития ЕСГ детализируются за счет учета сезонного регулирования и надежностных характеристик подсистем. На четвертом уровне с позиций интересов инвесторов рассчитываются параметры объектов систем газоснабжения.

Задачи исследования надежности рассматриваются на третьем и четвертом уровнях предлагаемой иерархии.

В рамках исследований оптимального развития газоснабжающих систем с учетом многоуровневого моделирования была разработана система математических моделей для задач анализа и синтеза надежности для низших уровней иерархии [2]. Методическое обеспечение указанных задач включает аналитические и статистистические методы теории вероятностей и математической статистики (теоремы о сложении и умножении вероятностей, методы построения законов распределения и их композиции, схему "гибели и размножения", метод статистических испытаний и т.п.); аналитические методы решения оптимизационных задач (метод Басакера-Гоуэна).

Объектом исследования на третьем уровне иерархии является сложная закольцованная газоснабжающая система, охватывающая месторождения и другие источники газа, подземные хранилища газа (ПХГ) и узлы потребления газа с различными категориями потребителей, и связанная системой магистральных газопроводов (МГ). Месторождения, а также трубопроводы, осуществляющие импорт газа, рассматриваются как внешние по отношению к системе источники, а потребители и экспортные трубопроводы – как внешние потребители. ПХГ трактуются либо как источники, либо как потребители газа в зависимости от их режима работы на рассматриваемый период. Поскольку расчетный интервал достаточно большой (сезон или год), то рассматриваются только стационарные режимы, то есть совокупность состояний системы, без учета последовательности, в которой эти состояния появляются во времени.

Отличительной особенностью задач анализа и синтеза надежности сложных энергетических систем, и в частности ГСС, является огромное

число возможных состояний. При наличии очень большого числа элементов, одинаково влияющих на общую надежность, целесообразным становится использование методов математической статистики, в данном случае метода статистических испытаний (метода "Монте-Карло").

С его помощью разыгрываются различные значения случайных величин, которые характеризуют конкретное состояние системы. Далее для каждого состояния рассчитывается режим функционирования системы, для этого используется метод оптимизации потокораспределения газотранспортных систем (модифицированный метод Басакера-Гоуэна). Он позволяет учитывать первый закон Кирхгофа, соответствующие ограничения на производительность трубопроводов, месторождений, хранилищ и на подачу газа потребителям, которая не должна превышать расчетный спрос, а также потери газа на стадии добычи, хранения и транспортировки и взаимозаменяемость топлив. В результате вычисляются функции распределения небалансов газа (дефицита или избытка) для каждого потребителя, по которым и определяются основные показатели надежности, а именно: обеспеченность потребителя газом как вероятность удовлетворения спроса, коэффициент относительного обеспечения потребителей газом при существующих (или планируемых) мощностях, резервах и запасах, и другие характеристики системы.

На четвертом уровне иерархии объектами исследования являются либо сложный многониточный газопровод, либо месторождение газа, либо подземное газохранилище. Каждый такой объект в общем случае представляет собой совокупность разнородных последовательно-параллельно соединенных групп элементов. Путем эквивалентирования они заменяются группами из однородных элементов, что дает возможность использовать аналитические методы расчета, в данном случае схему «гибели и размножения». Составляется система математических уравнений, из которой находятся вероятности возможных состояний однородной группы элементов. Зная вероятностные характеристики для каждой такой однородной группы,

можно осуществить их композицию и получить результирующий закон распределения вероятностей для всего объекта, по которому рассчитываются соответствующие показатели надежности.

Программные средства. Необходимость их модернизации.

На основе системы математических моделей для задач анализа и синтеза надёжности ГСС и её элементов в 90-е годы прошлого века на алгоритмическом языке FORTRAN были разработаны алгоритмы и программные средства, предназначенные для анализа надежности сложного многониточного магистрального газопровода, газового месторождения и подземного хранилища газа; вычислительный комплекс для анализа надежности сложной закольцованной системы газоснабжения; а также вычислительные средства синтеза надежности МГ, т.е. средства определения оптимальных параметров проектируемого МГ (2 модификации) [2 – 5].

Основные недостатки имеющихся средств обусловлены их моральным старением. Перечисленные программные средства с момента их создания практически не подвергались модернизации, большая их часть эксплуатируется в исходном виде. Средства работают в пакетном режиме, имеют примитивный интерфейс командной строки либо не имеют его вовсе. Большинство решателей написано на языке FORTRAN, не использующем концепции объектно – ориентированного программирования. Обмен данными осуществляется с помощью специально формируемых текстовых файлов, из которых программа считывает исходные данные и в которые записывает результаты вычислений. Файлы имеют собственный уникальный формат; программы не обеспечивают поддержку общеупотребительных форматов файлов. Алгоритмы выполнения расчета тесно интегрированы с алгоритмами ввода-вывода и не позволяют разделить программный продукт на составные части (модули).

Отдельно следует остановиться на механизмах работы с данными указанных вычислительных средств. Все рассматриваемые программные средства используют "внутреннюю модель", т.е. в процессе работы создаются

промежуточные структуры данных, инкапсулированные внутри решающего модуля (решателя) и, как правило, недоступные извне. Это лучший вариант с точки зрения эффективности работы вычислительных алгоритмов, однако, при таком подходе возникает дублирование информации; соответственно появляется необходимость поддержки целостности и непротиворечивости данных. При больших размерностях решаемой задачи и использовании сложного алгоритма, требующего комбинированной работы нескольких решателей [3], существенными становятся затраты на хранение и преобразование информации, связанные с созданием значительного массива дублирующих данных. Как пример можно привести вычислительный комплекс для оценки надежности сложной закольцованной ГСС на 3-ем уровне иерархии (рис. 1), который в качестве исходных данных для вероятностно заданных величин использует ряды распределения, полученные в результате решения задач анализа надежности элементов ГСС на 4-ом уровне иерархии.

Относительно информационного обеспечения: при создании программных средств недостаточно внимания уделялось составлению описаний структур данных и алгоритмов работы рассматриваемых программ; информационному моделированию предметной области. Вследствие этого, при модернизации программных средств, требуется восстановление алгоритмов работы и структур данных; такого рода анализ зачастую сопоставим с переписыванием программного средства на произвольном языке программирования.

Большая часть данных на сегодняшний момент хранится в файлах пользовательских форматов; ведутся работы над внедрением реляционных СУБД. Активно используются существующие универсальные программные комплексы (MATLAB, OpenOffice, Microsoft Visio и т.п.). Многие задачи требуют согласования данных из различных источников. Существующие постановки задач изменяются, возникают новые постановки задач; соответственно требуются изменения структуры моделируемой системы при

минимальных трудозатратах. Структура и параметры моделируемой системы зависят от многочисленных факторов, среди которых фигурируют расчетные схемы, варианты и серии расчетов, собственное мнение исследователя и т. п. При этом требуется обеспечить минимальную избыточность информации, возможность сопоставления различных вариантов и однозначность отображения на реальные объекты моделируемой системы.

Все вышесказанное обусловило необходимость модернизации имеющегося программно – информационного обеспечения. Требуется формализация информационных моделей предметной области, документирование существующих и вновь создаваемых программных средств. Необходим переход на современные универсальные средства разработки, обеспечивающие легкость в модификации программного продукта. Требуется отделить средства выполнения расчетов (решатели) от средств, осуществляющих диалог с пользователем. Одновременно необходим поиск более современных решателей и разработка соглашений, обеспечивающих легкую замену одного решателя другим, а также совместное использование различных решателей. Программное обеспечение должно обеспечивать более высокую переносимость как между различными версиями одной операционной системы, так и между различными системами. Требуется разработка современного пользовательского интерфейса, обеспечивающего легкость в использовании программного продукта.

Указанные условия требуют значительных изменений, как в способах реализации программных средств, так и в базовых концепциях. В этом случае наиболее целесообразно выполнить восстановление алгоритмов и структур данных и последующей их реализации с использованием более современных средств. Достоинством такого подхода является возможность совершенствования самих расчетных алгоритмов, учёта новых факторов, влияющих на надежность и т.п. С другой стороны, повторное кодирование неизбежно приводит к появлению ошибок, необходимости тестирования и т.п., что является недостатком предлагаемого подхода.

Усовершенствованная модель представления данных.

С целью разработки информационной модели предметной области была выполнена формализация структур данных, используемых при анализе надежности объектов ГСС. Опираясь на полученную информацию проводится объектное моделирование области исследований. Фрагмент структуры ГСС с точки зрения задач надежности в виде UML-диаграммы представлен на рис. 2.

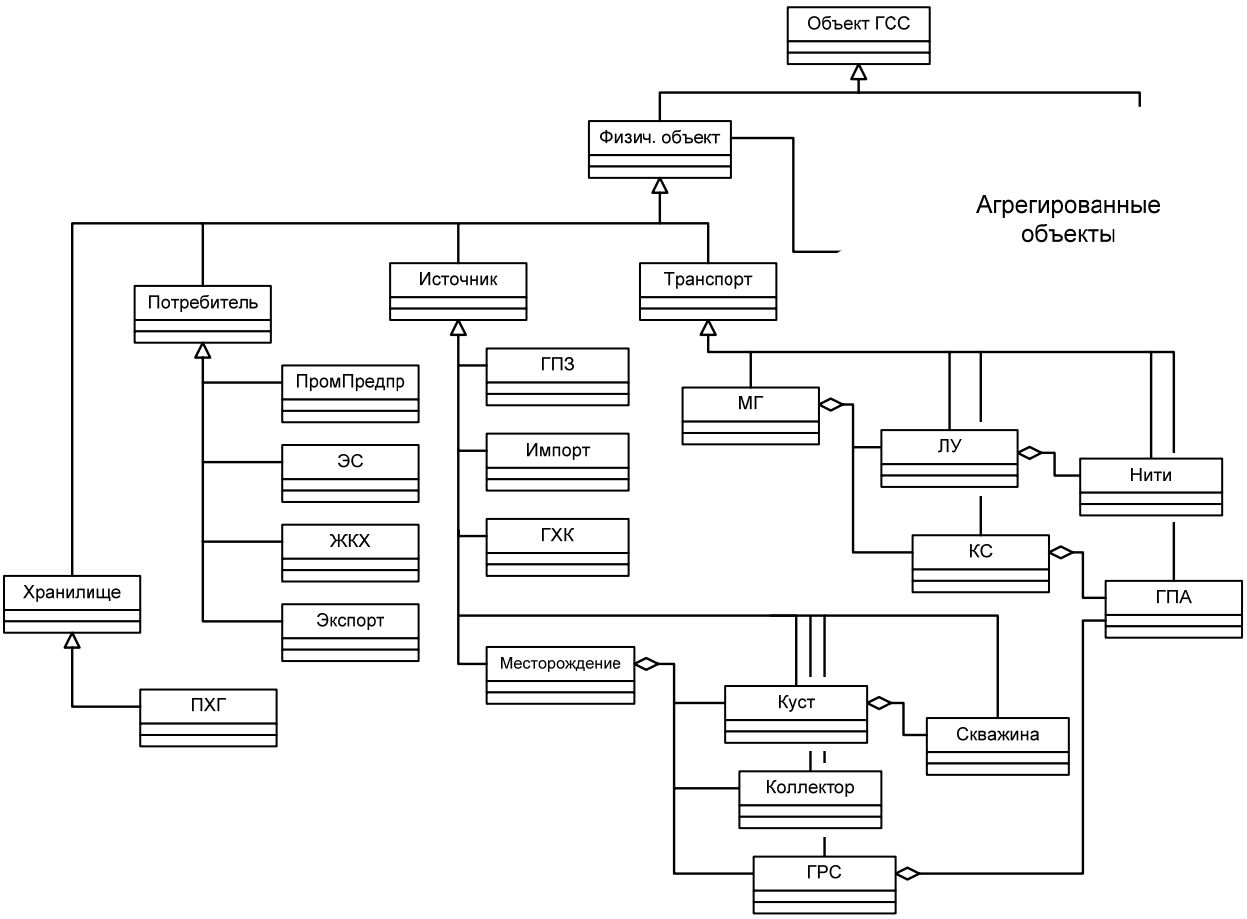


Рис. 2. UML – диаграмма иерархии и связей составляющих объектов ГСС с точки зрения задач надежности.

Модель, приведенная на рис. 2, отражает иерархию и связи основных типов объектов газоснабжающей отрасли, обладающих свойством однозначного отображения (т.е. каждый объект модели соответствует некоему физическому объекту отрасли, рассматриваемому отдельно от других). В дополнение к этому, в задачах развития и функционирования (второй и третий уровень на рис. 1) рассматриваются агрегированные объекты (на

рисунке не показаны), представляющие собой группы разнородных физических объектов, объединяемых по тому или иному признаку. Все объекты подразделяются на четыре основных типа:

- источники газа: месторождения (наиболее часто рассматриваемый источник), газохимические комбинаты (ГХК), газоперерабатывающие заводы (ГПЗ) и др.;
- потребители газа: жилищно-коммунальное хозяйство (ЖКХ), газовые электростанции (ЭС), промышленные предприятия и др.;
- подземные хранилища газа (ПХГ);
- транспорт газа – магистральные газопроводы.

Требуется отметить, что к настоящему моменту не все из объектов газоснабжающей отрасли нашли свое однозначное выражение в рассматриваемой модели.

Магистральный газопровод (МГ), как уже отмечалось выше, является основным звеном системы, достаточно сложным и содержащим большое число разнородных элементов, таких как головная и линейные компрессорные станции (КС) для перекачки газа, линейные участки трубопроводов (ЛУ), газораспределительные станции, и т.п. Линейный участок газопровода состоит из одной или более нитей (параллельных трубопроводов различного диаметра) с перемычками между ними. Компрессорная станция рассматривается как совокупность газоперекачивающих агрегатов (ГПА), находящихся в различных режимах работы: рабочие агрегаты, нагруженный и облегченный резервы и агрегаты ненагруженного резерва.

Основные объекты, определяющие состав месторождений, а также подземных хранилищ газа (поскольку хранилища создаются на базе выработанных месторождений газа), следующие: скважины, прискважинные сооружения, выкидные линии, групповые сборные пункты, газосборный коллектор, промысловая газораспределительная станция (ГРС) с устройствами очистки, осушки и охлаждения газа, дожимная компрессорная

станция. Основным элементом, имеющим наибольший удельный вес по влиянию, как на стоимость, так и на общую надежность месторождения, является скважина. Газ, полученный от скважин, направляется в газосборный коллектор, подающий газ на газораспределительную станцию (ГРС), оснащенную газоперекачивающими агрегатами. Из ГРС газ поступает в головную компрессорную станцию или непосредственно в магистральный газопровод.

Среди потребителей обычно выделяют следующие категории: коммунально-бытовые (жилые и общественные здания, малые котельные и др.); промышленные предприятия, осуществляющие переработку газа; промышленные котельные и электростанции, работающие на газе; другие категории потребителей (например, приобретающий все большее значение транспорт, работающий на газе). С точки зрения задач надежности можно выделить детерминированную составляющую, определяемую топливно-энергетическими и промышленными потребителями, и случайную составляющую, куда, в частности, входят коммунально-бытовые потребители.

На основании анализа постановок и процессов решения задач, существующих программных средств, используемых моделей ГСС, современных средств и технологий, были сформулированы следующие требования к разрабатываемой модели представления данных:

- использование различных источников данных: форматированные файлы (включая стандартные и уникальные форматы), реляционные базы данных, средства работы с данными, предоставляемые существующими программными комплексами, данные непосредственно в памяти компьютера;
- возможность изменения структуры моделируемой системы с минимальным влиянием на структуру самой модели представления данных;
- предоставление средств работы с данными для различных языков, в том

числе не использующих понятия объектно-ориентированного программирования;

- обеспечение вариативности структуры системы и параметров ее объектов при сохранении однозначности самих объектов;
- представление данных в виде массивов, актуальное при выполнении интенсивных вычислений.

Существующие модели представления данных можно разделить на два класса: с эксплицитной (заложенной в структуру самой модели) и имплицитной (выводимой на основании дополнительных данных, т.е. метаданных) структурой моделируемой системы. К первым относятся модели "сущность – связь", объектная модель и т.п. Модели такого типа изменяют собственную структуру при изменении структуры моделируемой системы. Ко вторым относятся модели типа Resource Description Framework (RDF), DOM-модель, используемая для представления XML-документов, иерархический формат представления данных HDF, и т.п.; они сохраняют собственную структуру, но требуют дополнительных средств для хранения формальных описаний моделируемых систем (метаданных), развитых средств верификации и поиска. Наиболее удобной оказалась модель представления данных EAV (точнее, ее вариант EAV/CR), часто используемая в системах медицинского назначения [6].

Разработанная модель представления данных в виде диаграммы "сущность – связь" (Entity – Relationship, ER) приведена на рис. 3. Здесь следует отметить, что ER – нотация применена исключительно для удобства описания; разработанная модель одинаково успешно может быть представлена и в других стандартах формального описания моделей данных.

Модель по принципу своей организации близка к моделям с имплицитной (выводимой на основании метаданных) структурой. Отличительной чертой является то, что основные метаданные заложены в самой модели в виде элементов (типы объектов описываемой системы, таблица "element") и их параметров (числовых, строковых, структурных и т.п.

характеристик объектов, таблица "parameter"), описывая структуру моделируемых систем. Значения параметров для каждого элемента конкретизируются его экземпляром (таблица "instance"), который можно сопоставить с экземпляром класса – объектом в традиционном объектно – ориентированном программировании. Сами значения находятся в таблице "context_value", реализуя подход, согласно которому все значения одного типа находятся в одной строке одной таблицы. Кроме ссылок на параметр и экземпляр элемента, значение конкретизируется еще т.н. "контекстом представления" (таблица "context").

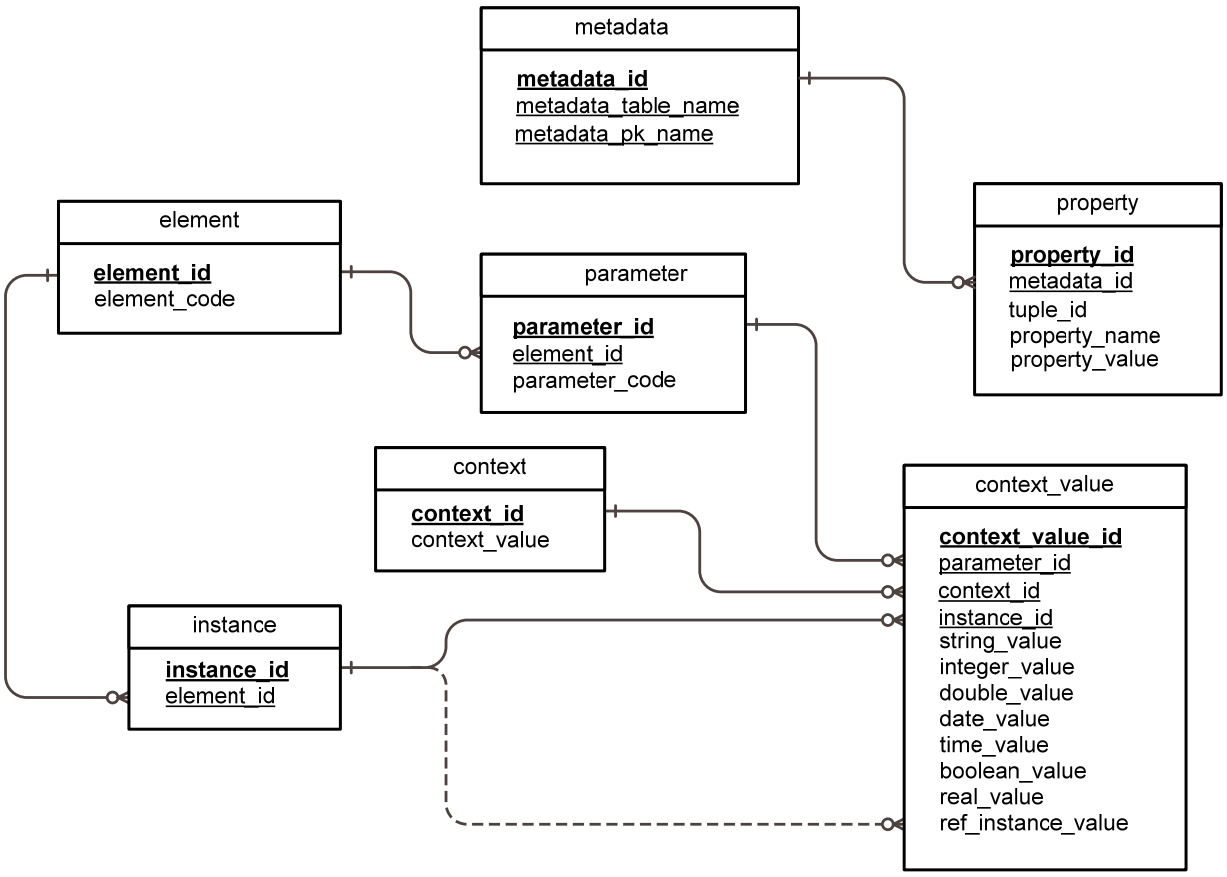


Рис. 3. Модель представления данных для задач анализа и синтеза ГСС.

Понятие "контекста представления" является уникальной отличительной чертой описываемой модели и отражает точку зрения на систему, от которой зависит структура и значения параметров рассматриваемой системы при сохранении однозначности составляющих ее объектов. Точка зрения определяется как произвольная комбинация одного

или более существующих в модели экземпляров различных элементов. В реализации инструментальных средств, описываемых в следующем разделе, принято вычисление хэш-значения указанной комбинации и создание записи в таблице "context" представляющей вычисленное значение.

Таблицы "metadata" и "property" введены для обеспечения дополнительных возможностей модели при ее реализации в реляционных СУБД. Такими возможностями являются хранение информации о составляющих модель таблицах и возможность добавления пользовательских комментариев к любым записям модели данных.

Описанная модель обладает неизменностью собственной структуры. Это означает, что инструментальные средства работы с данными по описанной модели не потребуют изменений при изменениях структуры моделируемых систем. В то же время модель позволяет вывести структуру систем в терминах объектного (класс-объект, поле-значение) моделирования, что несомненно является значимым достоинством. Введение контекста позволило избежать трудностей, связанных с необходимостью создания и сопоставления многочисленных вариантов расчетов, выполняемых по одной схеме ГСС; позволило обеспечить однозначное сопоставление с объектами моделируемой системы.

При таком подходе, однако, возникла задача, связанная с необходимостью создания средств поиска по модели. Для большинства стандартных моделей (включая RDF) уже имеются разработанные формальные языки запросов и реализации поисковых машин. Реализаций для моделей EAV и EAV/CR к настоящему моменту найдено не было; кроме того, значительные изменения исходной структуры делают применения стандартных средств поиска невозможным. Таким образом, для полноценной работы в будущем потребуется разработка собственного формального языка запросов к данным и создание программной реализации поисковой машины.

Вновь создаваемые и модернизируемые программные средства.

В качестве базовой модели было предложено использовать деление

программных модулей на пять типов: средства хранения данных, их визуального представления, выполнения расчетов (решатели), средства преобразования данных, призванные согласовать различные представления данных между программными единицами.

Кроме того, выделены средства организации взаимодействия, которые используются для организации работы различных модулей программы или различных программ между собой. На указанные средства ложится задача взаимодействия с удаленным компьютером, в качестве которого может рассматриваться сервер БД с находящейся на нем общей базой данных, более мощный вычислительный сервер для выполнения интенсивных расчетов, сервер бизнес-логики (для доступа через Web-интерфейс) и т. п.

Взаимодействие программных компонент в предлагаемой модели разбивается на два относительно независимых типа: взаимодействие по управлению, подразумевающее алгоритмы вызова функций и возврата управления (собственно бизнес-логика) и взаимодействие по данным, подразумевающее передачу и согласование данных.

Для обеспечения взаимодействия по данным предлагается отделить интерфейсы работы с данными от интерфейсов бизнес – логики и провести их формализацию, основываясь на описанной выше модели организации данных. Поскольку модель обладает неизменной структурой, появляется возможность объявления интерфейсов работы с данными для конкретных языков и сред программирования. Дальнейшая их реализация будет зависеть от используемых средств доступа к данным, среди которых планируются к применению реляционные БД, текстовые файлы различных форматов и внутреннее представление данных в программах OOo Calc, MATLAB и MS Visio.

Для осуществления взаимодействия по управлению имеется достаточно обширный арсенал средств, начиная от мощных универсальных технологий (CORBA, EJB, ActiveX, …) и заканчивая простейшими решениями наподобие использования класса ServiceLoader в Java или создания dll – библиотек.

Задача состоит лишь в проведении бизнес – анализа, объявлении интерфейсов и кодировании их реализаций. В зависимости от того, какими средствами организуется взаимодействие, его можно разделить на четыре уровня, представленных на рис. 4.

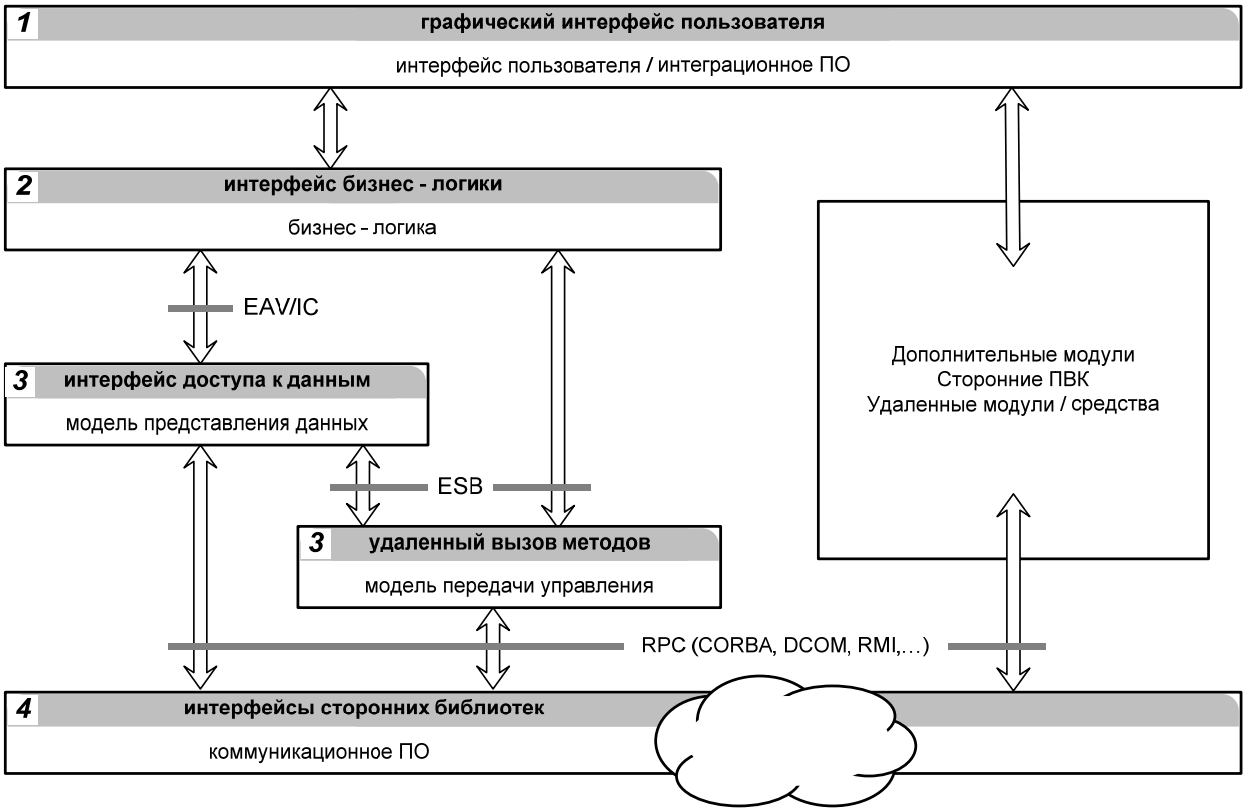


Рис. 4. Уровни организации взаимодействия программных компонент.

Необходимо создать уровень абстракции "3", который объединит механизмы передачи управления, пересылки и согласования данных. В дальнейшем такой подход позволит использовать любую удобную технологию удаленного взаимодействия (уровень "4"). Базовым уровнем в данной модели является уровень реализации программного кода (уровень "2"), в котором располагаются средства выполнения расчетов, работы с данными и т.п. Программные компоненты и модули этого уровня имеют два "выхода"; они могут использоваться непосредственно как библиотеки/модули на определенном языке программирования, либо использовать средства, предоставляемые уровнем "3", для создания абстрактного механизма обращений. Библиотеки и модули, реализованные на уровне "2", в свою

очередь могут служить основой для создания программных средств, интегрирующих в себе различный набор необходимых модулей согласно решаемым задачам. Такие средства могут быть оформлены как самостоятельные программы для решения частных постановок задач, как дополнения для иных программных комплексов, либо как интегрированный программный комплекс.

Для использования описанной ранее модели организации данных выполнена экспериментальная реализация программной библиотеки, предоставляющей инструментальные средства работы с данными. Платформой реализации был выбран язык Java и платформа JavaSE. Выбор был обусловлен наличием опыта разработки на данном языке, низким порогом вхождения и существованием большого количества бесплатных библиотек и компонентов. Главный недостаток платформы — относительно невысокая скорость исполнения программного кода — на данном этапе не является значимой величиной, т.к. размерности решаемых задач очень малы (менее 10^3 элементов расчетной схемы).

Реализация первого варианта указанной библиотеки с использованием хранилища данных в виде многомерных массивов простейших типов в памяти компьютера была выполнена при создании программы поиска оптимального потокораспределения газа в сети ГСС, оформленной как дополнение для табличного процессора Calc, входящего в программный пакет OpenOffice.org версии 3. Требуется заметить, что в отличие от большинства дополнений, реализуемых на внутреннем языке пакета OOo Basic, данная программа написана на языке Java, что предполагает возможности ее свободного развития вплоть до преобразования в самостоятельный программный комплекс.

Основные положения разработки и выводы по итогам эксплуатации описаны в [7]. Библиотека обеспечивает: более быстрое чтение/запись за счет прямого обращения к массивам; удобство представления значений параметра для всех экземпляров элемента; быстроту замены значений

параметров при смене контекста представления.

В процессе эксплуатации были также выявлены недостатки как самой модели, так и реализации программной библиотеки, а именно: слабая наглядность (при этом возрастает вероятность появления и сложность локализации ошибок); отсутствие средств реализации наследования, аналогичных средствам ООП; отсутствие транзакционности и потокобезопасности; сложность представления единичных экземпляров; отсутствие возможности формирования выборок по критериям пользователя (формальный язык запросов). По результатам экспериментальной эксплуатации сделано заключение о том, что требуется существенная модернизация созданной библиотеки.

В данный момент разрабатывается усовершенствованный вариант библиотеки взаимодействия по разработанной модели представления данных, призванный устранить часть сформулированных выше недостатков. Средством хранения данных (в дополнение к многомерным массивам в памяти компьютера, развитие которых было временно остановлено) были выбраны реляционные БД, что позволит эффективно тестировать создаваемый код (энергонезависимое хранение данных, возможность модифицировать данные внешними средствами), осуществить пробную реализацию языка формальных запросов, и в будущем обеспечит программным модулям доступ к СУБД.

Следующим этапом модернизации программно – вычислительных средств определена модернизация программы анализа надежности магистрального газопровода. В настоящий момент:

- восстановлены и формально описаны структуры входных и выходных данных наследуемой программы;
- восстановлены алгоритмы работы процедур;
- созданы структуры данных, описывающих моделируемый газопровод;
- расчетные алгоритмы реализованы в системе MATLAB, обеспечена возможность выполнения расчетов средствами системы;

- разрабатываются средства графического пользовательского интерфейса;
- ведется тестирование и документирование создаваемых средств.

Для обеспечения следующего этапа модернизации программное средство разделено на относительно независимые уровни реализации, приведенные на рис. 5.

Рис. 5. Уровни взаимодействия компонент программы анализа надежности МГ.

В качестве приложения (уровень 1) выступают средства диалога с пользователем (графический интерфейс приложения) и организации взаимодействия, разрабатываемые на языке Java. Приложение можно будет рассматривать в качестве одного из компонент при дальнейшей их интеграции в единый комплекс. К указанному уровню на данном этапе также относятся разрабатываемые инструментальные средства работы с данными и средства взаимодействия компонент (уровни 5 и 6).

Средства взаимодействия на данном этапе реализуются в простейшем варианте с использованием средств загрузки сервисов платформы JavaSE 6. В дальнейшем средства будут модернизированы с использованием технологий удаленного доступа и реализованы в виде библиотек для их использования в платформо – зависимых языках (Fortran, Pascal, ...), что обеспечит возможность их непосредственного применения на уровнях 3 и 4.

Платформо – независимые сервисы (уровень 2) на данном этапе

представляют собой объявления интерфейсов на языке Java, организующих пакетное выполнение расчета. После реализации на уровне 5 средств удаленного взаимодействия здесь будут находиться объявления сервисов бизнес-логики приложения.

Уровень платформо – зависимых сервисов (уровень 3) определился специфическим интерфейсом среды MATLAB Compiler Runtime (MCR), использующей собственную модель представления данных. Для ее инкапсуляции на языке “C” был создан набор функций, предоставляющий выполнение операций в терминах модели магистрального газопровода. На данном этапе осуществляется реализация взаимодействия с интерфейсом среды MCR. Поскольку в системе MATLAB имеются средства экспорта расчетных функций непосредственно в классы языка Java, а также поскольку реализация средств работы с данными существует только в языке Java, было решено осуществить прямой переход на уровень 2, создав сервис, напрямую использующий средства, предоставляемые средой времени исполнения (MCR). В дальнейшем средства для обмена данными должны быть реализованы в виде библиотек для их использования в платформо – зависимых языках (Fortran, Pascal и т.д.), что обеспечит возможность их непосредственного применения на уровнях 3 и 4.

Заключение

1. Исследование функционирования и развития ГСС является сложной задачей, требующей наличия средств автоматизации расчетов, организации эффективного хранения данных и их представления для удобства анализа. Существующее программное обеспечение морально устарело, и требует модернизации для соответствия современным требованиям.
2. Разработана модель организации данных, призванная обеспечить их универсальное представление вне зависимости от языка и средств реализации; создаются соответствующие инструментальные средства, проводится их тестирование с целью обеспечения удобства использования.
3. Осуществляется модернизация программ для расчета надежностных

характеристик объектов ГСС; предусматривается восстановление структур данных и алгоритмов работы программ, формальное описание моделей ГСС и составляющих ее объектов с точки зрения задач надежности, а также реализация алгоритмов более современными средствами и разработка графических интерфейсов пользователя.

Список литературы

[1] Илькевич Н. И., Дзюбина Т. В., Калинина Ж. В. Обоснование развития систем газоснабжения // Системные исследования в энергетике: ретроспектива научных направлений СЭИ – ИСЭМ // Под. ред. Н. И. Воропая. – Новосибирск: Наука, 2010. – С. 546 – 556.

[2] Дзюбина Т.В., Илькевич Н.И. Методические подходы и математические модели для анализа и синтеза надежности при многоуровневом исследовании газоснабжающих систем // Надежность систем энергетики // Под ред. Н.И. Воропая. – Новосибирск: Наука, Сибирское предприятие РАН, 1999. – С. 325 – 333.

[3] Дзюбина Т.В., Илькевич Н.И. Модель оценки надежности сложных газоснабжающих систем // Известия РАН. Энергетика, 1998. – № 6. – С. 84 – 91.

[4] Дзюбина Т.В., Илькевич Н.И. Определение рациональных параметров МГ с учетом надежности и эффективности их инвестиционных проектов // Российский научный симпозиум по энергетике. Методические вопросы ис-следования надежности БСЭ. – Казань: Казан. Гос. энерг. унив-т, 2001. – Т. IV. – С. 132 – 135.

[5] Дзюбина Т.В., Илькевич Н.И. Оптимизация параметров МГ в рыночных условиях с учетом надежности // Известия АН. Энергетика, 2002. – №3.– С. 93 – 101.

[6] Dinu, Valentin; Nadkarni, Prakash. Guidelines for the effective use of entity-attribute-value modeling for biomedical databases // International journal of medical informatics – vol. 76, iss. 11–12. – 2007. – С. 769 – 779.

[7] Метелкин А. М. Разработка формата данных для межпрограммного

обмена массивами информации // Системные исследования в энергетике. – Иркутск: ИСЭМ СО РАН, 2011. – С. 191 – 196.

УДК 330.150

ИМИТАЦИОННАЯ МОДЕЛЬ ПОВЕДЕНИЯ ОДНОПРОДУКТОВОГО РЫНКА

Зоркальцев В. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: zork@isem.sei.irk.ru

Мокрый И. В.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: ygr@isem.sei.irk.ru

Аннотация. Рассматривается имитационная модель поведения в непрерывном времени олигопольного рынка. Потребители в модели представляют рынки, которые описываются обратной функцией спроса. Для поставщиков описаны два правила поведения (правило «Вальрас» и правило «Курно»), задаваемые системами обыкновенных дифференциальных уравнений. Сценарий определяет по каким правилам действует каждый из поставщиков. Из расчетов видно, что одному из поставщиков в точке олигопольного равновесия выгодно перейти к другому правилу поведения из-за чего его прибыль возрастёт при снижении прибыли остальных участников.

Ключевые слова. Рынок, олигополия, поставщик, потребитель, модель, обратная функция спроса, издержки, прибыль.

Введение

Представленные в данной статье результаты исследования поведения олигопольных рынков осуществлены в связи с проведенными в электроэнергетике России рыночными преобразованиями.

Одной из особенностей олигопольных рынков является их неустойчивость. В частности, у отдельных олигополистов всегда может возникнуть желание к смене правил своего поведения на рынке. Из-за чего он может временно выигрывать за счёт других.

Рассмотрим граф (рис. 1), состоящий из m узлов поставщиков и n узлов потребителей (в дальнейшем поставщиков и потребителей). Пусть $i \in \{1,\ldots,m\}$ – номера поставщиков, $j \in \{1,\ldots,n\}$ – номера потребителей.

Поставки осуществляются от каждого поставщика каждому потребителю. Переменные (зависящие от времени) x_{ij} – объемы поставок продукта поставщика i потребителю j для $i \in \{1,\ldots,m\}$, $j \in \{1,\ldots,n\}$.

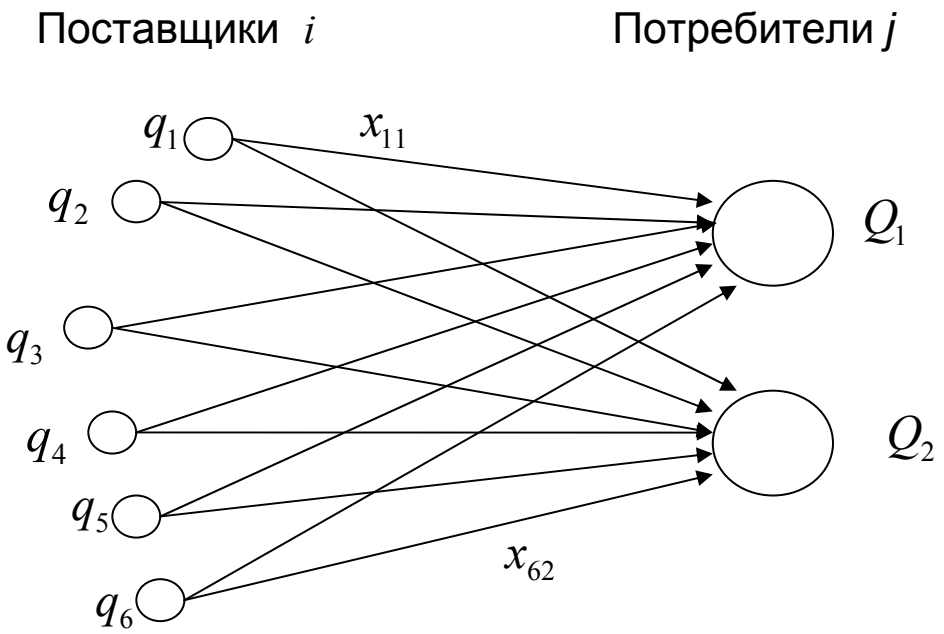


Рис.1. Поставщики и потребители модели: q_i – объем производимого продукта поставщиком i, Q_j – объем потребляемого продукта потребителем j, x_{ij} – объемы поставок продукта поставщика i потребителю j

Потребители

Считаем, что выделенный в модели «потребитель» с номером j представляет некий рынок с большим количеством реальных потребителей, имеющих малый удельный вес на этом рынке, не может в своих интересах воздействовать на цену, складывающуюся на этом рынке. Эта цена зависит только от общего объема товара Q_j, поставляемого на этот рынок.

Состояния потребителей определяются следующими переменными:

$Q_j = \sum_{i=1}^{m} x_{ij}$ – объем потребляемого продукта потребителя j;

$P_j = \pi_j(Q_j)$ – уровень цены потребителя j (убывающая функция от Q_j);

$\pi_j(Q_j)$ – обратная функция спроса, зависимость уровня цены потребителя j от объема потребляемого им продукта Q_j.

В рассматриваемых здесь примерах используется линеаризованная функция спроса

$$\pi_j(Q_j) = A_j - B_j Q_j$$

при заданных $A_j > 0$, $B_j > 0$.

Поставщики

Поставщики характеризуются следующими переменными:

$q_i = \sum_{j=1}^{n} x_{ij}$ – объем производимого продукта поставщика i;

$C_i = \frac{k_i \cdot q_i^2}{2}$ – переменные издержки поставщика i при заданном $k_i > 0$;

$MC_i = \frac{\partial C_i}{\partial q_i} = k_i \cdot q_i$ – предельные издержки поставщика i;

$R_{ij} = P_j x_{ij}$ – доходы поставщика i в результате поставок потребителю j;

$\overline{R}_i = \sum_{j=1}^{n} R_{ij} = \sum_{j=1}^{n} P_j x_{ij}$ – суммарные доходы поставщика i в результате поставок всем потребителям.

Отсюда прибыль поставщика i

$$S_i = \overline{R}_i - C_i,$$

тогда общая прибыль всех поставщиков будет

$$S = \sum_{i=1}^{m} S_i .$$

Во всех рассматриваемых далее моделях величины x_{ij}, Q_j, q_i, P_j, C_i, MC_i, R_{ij}, $\overline{R}_i$, S_i, S являются переменными, зависящими от времени, для которых в каждый момент времени $t \geq 0$ выполняются указанные здесь соотношения.

Рассмотрим три варианта моделей.

Модель «Вальрас»

Поставщик предполагает, что не может повлиять на установившиеся цены, и при этом стремится осуществить такие поставки, которые максимизируют его прибыль. Общее равновесие достигается установлением равенства в спросе на продукт и предельными издержками на его производство, включающими в себя и нормальную прибыль.

Продукты производятся и перераспределяются так, чтобы разность цена продукта минус предельные издержки на производство стремилась к нулевому значению, что в частности выражает следующее правило поведения *i*–го поставщика (правило «Вальрас»)

$$\frac{dx_{ij}}{dt} = x_{ij}(P_j - MC_i),\ \ j \in \{1,\ldots,n\},$$

при выбранных начальных значениях переменных $x_{ij} > 0$.

Модель «Вальрас» реализуется, если все поставщики в модели действуют по этому правилу. Данная модель приводит к стационарным решениям, которых может быть много по переменным x_{ij}, но имеющих единственные значения по переменным Q_j, P_j, q_i. Происходит выравнивание цен во всех узлах потребителях. В стационарном состоянии выполняется система равенств

$$P_j - MC_i = 0,\ i \in \{1,\ldots,m\},\ j \in \{1,\ldots,n\}.$$

Алгоритм. Вычисляем для каждого момента времени

$$P_j = A_j - Q_j B_j,\ \ j \in \{1,\ldots,n\},$$

$$MC_i = \frac{dC_i}{dq_i} = k \cdot q_i,\ \ i \in \{1,\ldots,m\}.$$

Динамика изменения объемов поставок определяется системой уравнений

$$\frac{dx_{ij}}{dt} = x_{ij}(P_j - MC_i),\ \ i \in \{1,...,m\},\ \ \ j \in \{1,\ldots,n\}.$$

Модель «Курно»

Поставщик предполагает, что объемы поставок всех других поставщиков зафиксированы и максимизирует свою прибыль учитывая изменение цен, вызванных перераспределением собственных поставок. Изменение потоков во времени зависит от разницы между предельной выручкой в узле потребления

$$MR_{ij} = \frac{dR_{ij}}{dt}$$

и предельными издержками, что выражает следующее правило поведения *i*–го поставщика (правило «Курно»)

$$\frac{dx_{ij}}{dt} = x_{ij}(MR_{ij}(x_{ij}) - MC_i(q_i)),\ \ j \in \{1,\ldots,n\}.$$

Модель «Курно» реализуется, если все поставщики в модели действуют по правилу «Курно». Данная модель всегда приводит к одному стационарному решению по всем переменным. В стационарном состоянии выполняется система равенств

$$MR_{ij} - MC_i = 0,\ i \in \{1,\ldots,m\},\ j \in \{1,\ldots,n\}.$$

Отметим, если при регулировании перетоков в модели “Вальрас” кроме предельных издержек поставщику требуется знать только складывающиеся в смежных узлах цены, то в модели «Курно» требуется еще располагать информацией о значении производной обратной функции спроса, так как

$$MR_{ij} = P_j + x_{ij}\frac{dP_j(Q_j)}{dQ_j}.$$

В этом случае при распределении поставок между потребителями кроме цен учитывается также насколько изменится доход от всего объема x_{ij} поставок данному потребителю из-за изменения цен $dP_j(Q_j)/dQ_j$ вследствие прироста поставок.

Алгоритм. Вычисляем для каждого момента времени

$$MR_{ij} = A_j - Q_j B_j - x_{ij} B_j,\ i \in \{1,\ldots,m\},\ j \in \{1,\ldots,n\}\ ,$$

$$MC_i = \frac{dC_i}{dq_i} = k \cdot q_i,\ i \in \{1,\ldots,m\}.$$

И затем динамика изменения объемов поставок узла i определяется системой уравнений

$$\frac{dx_{ij}}{dt} = x_{ij}(MR_{ij} - MC_i),\ \ j \in \{1,\ldots,n\}.$$

Смешанная модель «Вальрас – Курно»

Один поставщик с индексом i=1 действует по правилу «Вальрас», другие поставщики с индексом $i \in \{2,\ldots,m\}$ действуют по правилу «Курно». Цель построения таких вариантов модели состоит в том, чтобы рассмотреть вопрос не могут ли отдельные поставщики олигопольного рынка выигрывать при смене правил своего поведения на рынке.

Параметры рассматриваемых моделей

Все, рассматриваемые далее модели включают два узла потребителя ($n = 2$), для которых заданы следующие обратные функции спроса:

$$P_1 = 25 - Q_1,$$

$$P_2 = 20 - Q_2.$$

Число узлов-поставщиков m может варьироваться от 1 до 6. Все поставщики имеют одинаковые издержки

$$C_i(q_i) = \frac{0{,}5 \cdot q_i^2}{2}.$$

Соответственно, предельные издержки поставщика i будут

$$MC_i = 0{,}5 \cdot q_i.$$

Результаты расчетов

Рассмотрим значения основных переменных для стационарных состояний трех вариантов моделей - «Вальрас», «Курно» и «Вальрас – Курно» с различным числом поставщиков.

Если все производители действуют по правилу «Вальрас», то независимо от объемов начальных поставок (табл. 1), модель переходит в стацио-

нарное состояние с предопределенными значениями следующих переменных:

- для потребителей (цены, объемы потребления товара),
- для поставщиков (объемы производимых продуктов, прибыль).

Различаются только доли поставляемых продуктов от одного поставщика к различным потребителям, которые зависят от начальных поставок. Устанавливаются равные для всех потребителей цены.

Таблица 1

Все поставщики действуют по правилу «Вальрас»

m	P_1	P_2	Q_1	Q_2	MC_i	q_i	S_i	S
1	11,25	11,25	13,75	8,75	11,25	22,5	126,6	126,6
2	7,5	7,5	17,5	12,5	7,5	15	56,25	112,5
3	5,625	5,625	19,38	14,37	5,625	11,25	31,64	94,92
4	4,5	4,5	20,5	15,5	4,5	9	20,25	81
5	3,75	3,75	21,25	16,25	3,75	7,5	14,06	70,31
6	3,214	3,214	21,79	16,79	3,214	6,429	10,33	61,99

Обратим внимание, что по мере увеличения числа поставщиков их суммарный доход монотонно убывает. Если все производители действуют по правилу «Курно», то независимо от объемов начальных поставок, модель переходит к одному стационарному решению по всем переменным (табл. 2). В модели «Курно» цены для разных потребителей различаются.

Таблица 2

Все поставщики действуют по правилу «Курно»

m	P_1	P_2	Q_1	Q_2	MC_i	q_i	S_i	S
1	16,25	13,75	8,75	6,25	7,5	15	171,9	171,9
2	12,08	10,42	12,92	9,583	5,625	11,25	96,31	192,6
3	9,625	8,375	15,38	11,62	4,5	9	61,53	184,6
4	8	7	17	13	3,75	7,5	42,69	170,8
5	6,845	6,012	18,15	13,99	3,214	6,429	31,34	156,7
6	5,982	5,268	19,02	14,73	2,813	5,625	23,99	143,9

Обратим внимание:

- по мере увеличения числа поставщиков происходит сближение цен потребителей,
- суммарная прибыль поставщиков по мере увеличения их числа вначале возрастает, принимает максимальное значение $S = 192{,}6$ при $m = 2$ и далее, монотонно убывает.

Один поставщик с индексом i=1 действует по правилу «Вальрас», другие поставщики с индексом $i \in \{2, \ldots, m\}$ действуют по правилу «Курно».

В этом варианте моделей по мере увеличения числа поставщиков их суммарная прибыль монотонно убывает. Происходит выравнивание цен во всех узлах – потребителях. При небольшом числе поставщиков $m < 5$ прибыль поставщика, действующего по правилу «Вальрас», ниже прибыли поставщика, действующего в тех же условиях по правилу «Курно». Однако начиная с общего числа поставщиков $m \geq 5$ его прибыль оказывается выше чем у поставщика действующего по правилу «Курно» в тех же условиях (табл. 3).

Таблица 3

Один поставщик действует по правилу «Вальрас», другие по правилу «Курно»

m	P_1	P_2	Q_1	Q_2	MC_1	MC_i	q_1	q_i	S_1	S_i	S
2	9	9	16	11	9	4,5	18	9	81	60,75	141,8
3	7,5	7,5	17,5	12,5	7,5	3,75	15	7,5	56,25	42,19	140,6
4	6,429	6,429	18,57	13,57	6,429	3,214	12,86	6,429	41,33	30,99	134,3
5	5,625	5,625	19,37	14,38	5,625	2,812	11,25	5,625	31,64	23,73	126,6
6	5	5	20	15	5	2,5	10	5	25	18,75	118,7

Для случая m=6 рассмотрим эту ситуацию на графике (рис. 2). Пусть шесть поставщиков действуют по правилу «Курно». В момент времени t = 10.0 поставщик 1 переходит к правилу «Вальрас» и в результате модель переходит в стационарное состояние, где поставщик 1 получает большую прибыль.

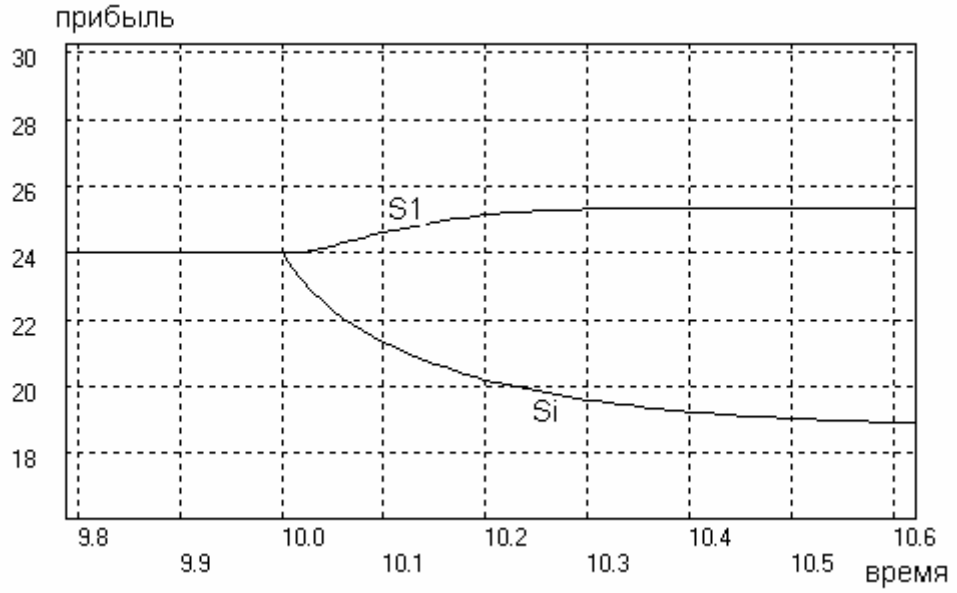


Рис.2. Переход поставщика 1 от правила «Курно» к правилу «Вальрас»

Кривая S_1 – прибыль поставщика с индексом $i = 1$,

кривая S_i – прибыль одного из поставщиков с индексом $i = 2, \ldots, 6$

Обратный переход (рис. 3) показывает, что на определенном интервале времени, прежде чем модель перейдет в стационарное состояние, поставщик 1 получает прибыль существенно выше, чем действуя по правилу «Вальрас» в стационарном режиме.

Из этих графиков видно, что одному из олигополистов в точке олигопольного равновесия выгодно перейти к другому правилу поведения – на схему «Вальрас». Из-за чего его прибыль возрастёт при снижении прибыли остальных участников. После этого ему опять выгодно перейти на схему «Курно», ведущую к олигопольному равновесию. Здесь после его временного выигрыша наступает период снижения его прибыли. Затем ему опять выгодно перейти на схему «Вальрас» и т.д.

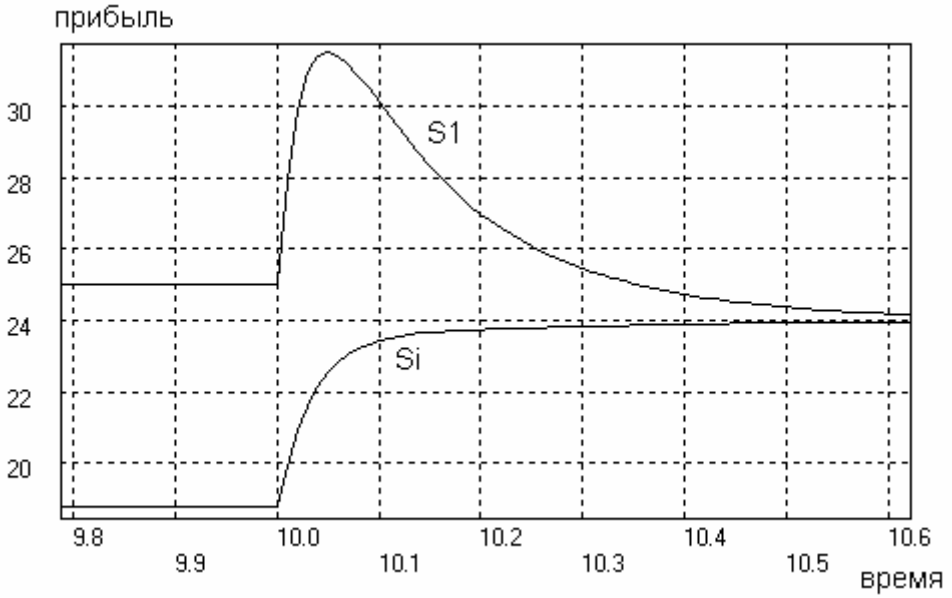


Рис.3. Переход поставщика 1 от правила «Вальрас» к правилу «Курно»

Кривая S_1 – прибыль поставщика с индексом $i = 1$,

кривая S_i – прибыль одного из поставщиков с индексом $i = 2, \ldots, 6$

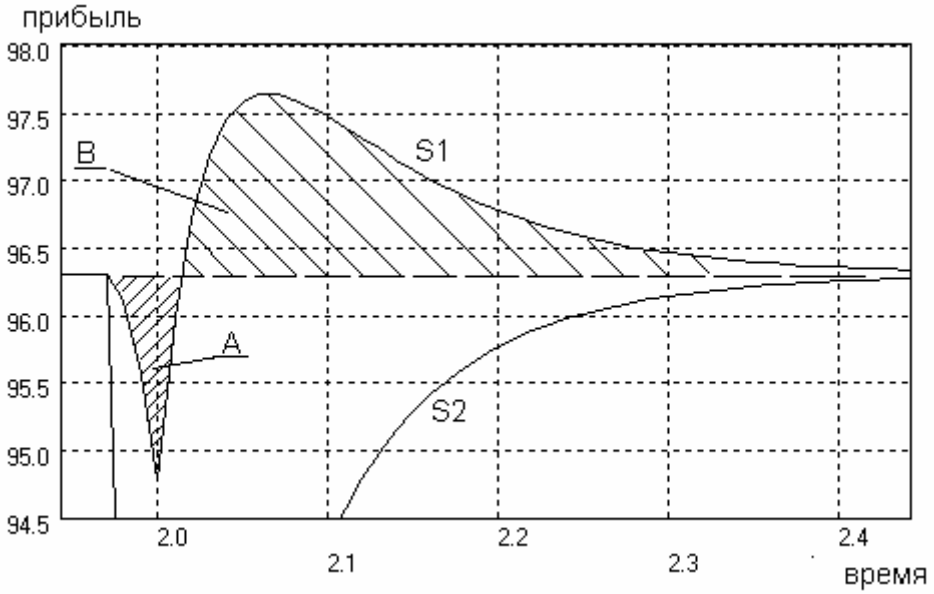


Рис.4. Переход поставщика 1 от правила «Курно» к «Вальрас» и обратно

Кривая S_1 – прибыль 1-го поставщика,

кривая S_2 – прибыль 2-го поставщика

На рис. 4 представлена ситуация, когда участвуют всего 2 поставщика. Поставщик 2 действует по правилу «Курно», а поставщик 1 до момента времени $t = 3.97$ действует по правилу Курно, затем переходит на правило «Вальрас» до момента времени $t = 4.0$, после чего опять переключается на

правило «Курно». Заштрихованные на графике области А и В показывают потерю прибыли и дополнительную прибыль соответственно, по отношению к случаю, если бы поставщик 1 все время действовал по правилу «Курно».

Заключение

Представленные в данной статье результаты имитационного поведения поставщиков - производителей товара на олигопольных рынках иллюстрируют возможные эффекты в изменении ситуации на рынке при смене правил поведения отдельных поставщиков. В некоторых ситуациях отдельные поставщики могут получать долгосрочные преимущества при выходе из состояния олигопольного равновесия. Расширение объемов поставок у таких поставщиков может дать им увеличение прибыли, хотя это будет сопровождаться снижением цен. Само снижение цен приведет к падению прибыли у других поставщиков. Данный факт, если он будет осознан и другими поставщиками, будет стимулировать и их выход из состояния олигопольного равновесия, что в конечном итоге приведет к снижению прибыли у всех поставщиков.

Приведенные расчеты иллюстрируют также возможности появления краткосрочных преимуществ у отдельных поставщиков при смене их правил поведения за счет инерционности в изменениях состояний других поставщиков. Данный факт означает возможность получения дополнительных преимуществ за счет частой смены правил поведения отдельных поставщиков. Этот факт, осознанный многими поставщиками, будет стимулировать к частой смене их правил поведения, что в конечном итоге приведет к хаосу на рынках.

Следует подчеркнуть, что рассмотренными примерами не исчерпывается все многообразие правил поведения и вариантов возможной смены этих правил. Многие другие варианты изменений правил поведения очевидно также могут сопровождаться долгосрочными им временными выигрышами отдельных поставщиков, а также их общим проигрышем и неустойчивой ситуацией на рынках в конечном счете.

УДК 519.83+621.311:51.001.57

МОДЕЛИ АНАЛИЗА СОСТОЯНИЙ ЭЛЕКТРОЭНЕРГЕТИЧЕСКОЙ СИСТЕМЫ

Зоркальцев В. И.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: zork@isem.sei.irk.ru

Пержабинский С. М.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: smper@isem.sei.irk.ru

Аннотация. Рассматриваются варианты модели оценки дефицита мощности, разработанные для вычислительных комплексов анализа надежности электроэнергетических систем. Обсуждаются математические свойства и особенности моделей. Особое внимание уделено модели, в которой учитывается нелинейный характер потерь мощности в линиях электропередачи. Указан способ сведения ее к задаче выпуклого программирования.

Ключевые слова. Электроэнергетическая система, дефицит мощности, надежность электроэнергетических систем, модель оценки дефицита мощности.

Введение

В начале 70-х годов прошлого века была разработана методика анализа надежности электроэнергетических систем (ЭЭС) [1, 2], основывающаяся на использовании метода статистических испытаний (метод Монте-Карло). Согласно данной методике сначала формируются случайным образом возможные состояния ЭЭС, затем для сформированных состояний в результате расчетов на модели оценки дефицита мощности определяются дефициты мощности в узлах ЭЭС. На основе многократных статистических испытаний вычисляются показатели вероятности и математического ожидания дефицита мощности в узлах системы.

В методике анализа надежности ЭЭС представляется в виде сложной многоузловой расчетной схемы. Узел схемы ЭЭС – энергорайон, который характеризуется суммарной мощностью генерирующего оборудования и суммарной нагрузкой потребления электроэнергии в нем. Связь между узлами – это совокупность линий электропередачи, обладающая некоторой ограни-

ченной пропускной способностью. Пропускные способности определяются либо электромеханической устойчивостью, либо термической устойчивостью линий электропередачи.

Модель оценки дефицита мощности (ОДМ) в данной методике занимает центральное место. От ее реализации зависит не только качество результатов, но и время проведения всего цикла расчетов. Поэтому к модели минимизации дефицита мощности предъявляются особые требования. Она должна быть агрегированной, максимально адекватной действительности, легко реализуемой, рассчитываемой за минимальное время. Чем меньше время расчетов на модели ОДМ, тем большее количество случайных состояний можно «проиграть» и тем самым увеличить точность оценки показателей.

Исходная модель оценки дефицита мощности ЭЭС

Рассматривается схема электроэнергетической системы, состоящая из n узлов и некоторого набора связей между ними. Заданы располагаемая мощность $\bar{x}_i$, нагрузка $\bar{y}_i$ в $i-$ом узле ЭЭС, предел пропускной способности $\bar{z}_{ij}$ линии электропередачи (ЛЭП), по которой передается мощность из узла i в узел j, $i=1,\ldots,n$, $j=1,\ldots,n$. Считается, что для всех i и j, $\bar{x}_i \geq 0$, $\bar{y}_i \geq 0$, $\bar{z}_{ij} \geq 0$. Если $\bar{z}_{ij}=0$ при некоторых i и j, то это означает, что фактически поток мощности из узла i в узел j не возможен.

Переменными задачи являются: используемая мощность x_i, покрываемая нагрузка y_i в узле i, поток мощности z_{ij} из узла i в узел j, $i=1,\ldots,n$, $j=1,\ldots,n$. Дефицит мощности в узле i соответствует величине $d_i=\bar{y}_i-y_i$.

В исходном варианте модели для определения расчетных величин решалась задача минимизации суммарного дефицита мощности в ЭЭС:

$$\sum_{i=1}^{n}(\bar{y}_i-y_i)\to\min, \qquad (1)$$

при линейных ограничениях-равенствах, выражающих баланс мощности в узлах:

$$x_i - y_i + \sum_{j=1}^{n} z_{ji} - \sum_{j=1}^{n} z_{ij} = 0,\ i = 1, \ldots, n, \qquad (2)$$

и двусторонних линейных ограничениях-неравенствах на переменные:

$$0 \le y_i \le \bar{y}_i,\ i = 1, \ldots, n, \qquad (3)$$

$$0 \le x_i \le \bar{x}_i,\ i = 1, \ldots, n, \qquad (4)$$

$$0 \le z_{ij} \le \bar{z}_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, n. \qquad (5)$$

При описании вариантов моделей ОДМ в целях упрощения записи используется избыточный состав переменных. В некоторых узлах может не быть генерации ($\bar{x}_i = 0$) или нагрузки ($\bar{y}_i = 0$). Поэтому соответствующие переменные x_i и y_i тождественно равны нулю и могут быть исключены из рассмотрения. Если между узлами i и j нет связи по ЛЭП, то $\bar{z}_{ij} = 0$ и переменная z_{ij} должна быть исключена из рассмотрения при расчетах на модели, в частности для всех узлов $\bar{z}_{ii} = 0$. При реализации модели потоки из узла i в узел j можно объединить в одну переменную, имея в виду, что потоки в обратном направлении к выбранному за положительное имеют отрицательный знак. Тогда вместо (5) должно использоваться интервальное ограничение на потоки мощности от некоторого отрицательного числа до некоторого положительного.

Определение 1. Узел i назовем потенциально дефицитным, если для одного из оптимальных решений задачи (1) – (5) величина дефицита для этого узла положительная.

Для реализации модели (1) – (5) используется алгоритм Форда-Фалкерсона. На основе расчетов было подмечено, что алгоритм Форда-Фалкерсона нередко приводит к показателям надежности ЭЭС, зависящим от порядка нумерации узлов. Это обусловлено тем, что в случае неединственности решения у задачи (1) – (5) алгоритм Форда-Фалкерсона дает «крайнюю» точку многогранного множества решений, в которой не все потенциально дефицитные узлы являются дефицитными.

Модель, в которой осуществляется распределение минимального дефицита мощности по всем потенциально дефицитным узлам

Из-за неоднозначности показателей надежности, получаемых в результате использования алгоритма Форда-Фалкерсона, в [3] предложено распределять суммарный дефицит пропорционально нагрузкам потенциально дефицитных узлов. Эта идея была реализована в виде двухкритериальной модели с лексикографически упорядоченными целевыми функциями. На первом этапе решалась задача минимизации дефицита мощности с одновременным выявлением всех потенциально дефицитных узлов. Для этого стал использоваться при решении задачи (1) – (5) метод внутренних точек [3], приводящий к относительно внутренним точкам множества оптимальных решений. Это свойство метода внутренних точек означает, что в получаемом оптимальном решении задачи (1) – (5) дефицитными окажутся все потенциально дефицитные узлы. На втором этапе решалась задача распределения полученного минимального суммарного дефицита между всеми потенциально дефицитными узлами [3].

Напомним, что в многокритериальных задачах с лексикографически упорядоченными целевыми функциями оптимизация по предыдущему (в порядке нумерации) критерию во всех отношениях считается приоритетнее, чем оптимизация по следующему за ним критерию. Для двухкритериальной задачи лексикографическая оптимизация означает, что оптимальное по второму критерию решение ищется в множестве оптимальных по первому критерию решений.

Стандартный путь решения таких двухкритериальных задач состоит в том, что сначала определяется оптимальное по первому критерию решение (на исходном множестве допустимых по условиям задачи решений). Затем фиксируется значение первой целевой функции на установленном оптимальном уровне. Это условие в виде ограничения добавляется в исходный набор ограничений. Наконец, решается задача оптимизации по второму критерию при расширенном составе ограничений (и при том же наборе переменных).

Ситуация существенно меняется, если задачей первого этапа является задача линейного программирования, а используемый алгоритм всегда приводит к относительно внутренней точке оптимальных решений. В [3] показано, что таковыми алгоритмами являются методы внутренних точек.

Относительной внутренностью выпуклого множества называется внутренность этого множества относительно минимального линейного многообразия, его содержащего [4]. Относительная внутренность многогранного множества решений системы линейных неравенств совпадает с множеством решений этой системы, имеющих минимальный набор активных ограничений. Активными для данного решения системы называются такие ограничения-неравенства, которые выполняются в виде равенства. Обсуждаемые здесь определение и свойства более подробно рассматриваются в [5].

Пусть $\overline{X}$ множество оптимальных решений задачи линейного программирования

$$c^T x \to \min,\ Ax = b,\ 0 \le x \le \bar{x}, \tag{6}$$

относительно вектора переменных $x \in R^n$. Заданными являются векторы $c \in R^n$, $\bar{x} \in R^n$, $b \in R^m$ и матрица A размера $m \times n$. Относительную внутренность $\overline{X}$ обозначим $ri\ \overline{X}$.

Пусть $f(x)$ – целевая функция второго этапа оптимизации. Если получено какое-то оптимальное решение $\tilde{x} \in \overline{X}$, то задача второго этапа имеет вид

$$f(x) \to \min,\ Ax = b,\ c^T x = c^T \tilde{x},\ 0 \le x \le \bar{x}. \tag{7}$$

В задаче (7) тот же состав переменных. При этом сохраняются все ограничения задачи первого этапа. Добавлено еще одно ограничение-равенство.

Если же известно, что полученное решение находится в относительной внутренности множества оптимальных решений $\hat{x} \in ri\ \overline{X}$, то множество оптимальных решений определяется как совокупность решений системы

$$Ax = b;\ x_j = 0,\ j \in J^1;\ x_j = \bar{x}_j,\ j \in J^2;\ 0 \le x_j \le \bar{x}_j,\ j \in J^3. \tag{8}$$

Здесь $J^1 = \{j : \hat{x}_j = 0\}$, $J^2 = \{j : \hat{x}_j = \bar{x}_j\}$, $J^3 = \{j : 0 < \hat{x}_j < \bar{x}_j\}$.

Задача второго этапа должна в этом случае решаться при ограничениях (8). Заметим, что в отличие от (7) в описании множества оптимальных решений (8) не вводится ограничение на значение первой целевой функции. Это ограничение учтено здесь неявно. Более того, условия $x_j = 0$ при $j \in J^1$ и $x_j = \bar{x}_j$ при $j \in J^2$ означают, что переменные с номерами из J^1 и J^2 зафиксированы и искомыми являются только переменные x_j с номерами $j \in J^3$, т.е. состав переменных у задачи второго этапа сокращается.

Важно также, что, вычислив относительно внутреннюю точку оптимальных решений исходной модели ОДМ (1) – (5), мы автоматически определяем все потенциально дефицитные узлы. Ими будут узлы, которые дефицитны при таком решении.

Итак, пусть известно, что в результате решения задачи (1) – (5) получена относительно внутренняя точка множества оптимальных решений. Это решение обозначим $\hat{x}_i$, $\hat{y}_i$, $\hat{d}_i$, $\hat{z}_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, n$. Тогда множество номеров потенциально дефицитных узлов составляют узлы, дефицитные для данного решения. Обозначим это множество $L = \{i : \hat{y}_j < \bar{y}_j\}$.

На втором этапе в рассматриваемой модели при условиях (2) – (5) и зафиксированных переменных, вышедших на границы ограничений-неравенств (3) – (5), осуществляется минимизация суммы квадратов отклонений дефицитов по узлам от дефицитов, получаемых в результате распределения минимального дефицита пропорционально нагрузкам:

$$\sum_{i \in L} (d_i^* - d_i)^2 \to \min, \tag{9}$$

где $d_i^* = D\bar{y}_i / \sum_{l \in L} \bar{y}_l$, $D = \sum_{i \in I} \hat{d}_i$ – минимальный суммарный дефицит, $\hat{d}_i = \bar{y}_i - \hat{y}_i$ – дефицит в узле i, полученный в результате решения задачи первого этапа.

В тех случаях, когда для оптимального решения задачи второго этапа (9) целевая функция равна нулю, т. е. $d_i = d_i^*$ при всех $i \in L$, полученное решение точно отражает задачу распределения суммарного дефицита между всеми потенциально дефицитными узлами пропорционально их нагрузкам. В иных ситуациях, обусловленных невозможностью перераспределения дефицита из-за ограничений на пропускную способность межузловых связей, распределение дефицита строго пропорционально нагрузкам в результате решения задачи (9) не реализуется.

Модель, в которой осуществляется распределение минимального суммарного дефицита пропорционально нагрузкам всех потенциально дефицитных узлов

В [6] была предложена одноэтапная схема расчетов, в которой используется квадратичный критерий, обеспечивающий одновременную минимизацию дефицита мощности с распределением его по узлам пропорционально нагрузкам в них. В этой модификации модели (1) – (5) распределение дефицита всегда получается пропорциональным нагрузкам потенциально дефицитных узлов. В результате проведенных экспериментальных расчетов [7] оказалось, что такая постановка более эффективна и в вычислительном отношении.

Определим решение задачи (1) – (5) с распределением минимального суммарного дефицита пропорционально нагрузкам.

Определение 2. Допустимое решение системы (2) – (5) x_i, y_i, z_{ij} и $d_i = \bar{y}_i - y_i$, $i = 1, \ldots, n$, $j = 1, \ldots, n$, является решением с распределением дефицита пропорционально нагрузкам, если при выполнении неравенства

$$d_l / \bar{y}_l > d_k / \bar{y}_k \tag{10}$$

для каких-либо узлов l, k таких, что $\bar{y}_l > 0$, $\bar{y}_k > 0$, не существует допустимого решения $\tilde{x}_i$, $\tilde{y}_i$, $\tilde{z}_{ij}$, $\tilde{d}_i$, при котором

$$\tilde{d}_l < d_l,\ \tilde{d}_k > d_k, \tag{11}$$

$$\widetilde{d}_i = d_i \text{ при } i \neq l, k. \qquad (12)$$

Согласно этому определению, дефициты распределены непропорционально нагрузкам, если для двух узлов выполняется неравенство (10), означающее указанную непропорциональность, которую при этом можно уменьшить согласно (11), (12), передав часть мощности из узла l в узел k.

Теорема 1 [6]. *У задачи* (1) – (5) *существует решение с распределенным минимальным дефицитом пропорционально нагрузкам узлов. У всех таких решений значения покрываемых нагрузок* y_i *и величин дефицитов* d_i *по узлам являются единственными, причем дефицитными для данного решения будут все потенциально дефицитные узлы.*

В изложенной выше двухэтапной схеме вычислений можно на втором этапе при тех же ограничениях вместо (9) использовать критерий

$$\sum_{i \in L} (d_i)^2 / \bar{y}_i \to \min, \qquad (13)$$

В вычислительном отношении этот критерий явно не хуже (9). При этом гарантируется получение решения с распределением минимального суммарного дефицита пропорционально нагрузкам.

Теорема 2 [6]**.** *При замене в двухэтапном алгоритме задачи* (9) *второго этапа на задачу* (13) *будет получено решение с минимальным суммарным дефицитом, распределенным по узлам пропорционально нагрузкам.*

На первом этапе вычислений вместо критерия (13) можно воспользоваться его обобщением. В [7] было доказано для случая, когда во всех узлах имеются нагрузки ($\bar{y}_i > 0$ для $i = 1, \ldots, n$), утверждение о том, что если решать задачу

$$\sum_{i=1}^{n} (d_i)^2 / \bar{y}_i \to \min \qquad (14)$$

при условиях (2) – (5), то получим решение с минимальным суммарным дефицитом мощности, распределенным пропорционально нагрузкам в узлах. Более того, в данной одноэтапной задаче можно без ущерба исключить из со-

става ограничений условие (3). Оно будет выполняться автоматически для получаемых оптимальных решений.

Введенный новый критерий (14) корректно определен только для случая, когда во всех узлах $\bar{y}_i > 0$, $i = 1, \ldots, n$. Если $\bar{y}_i = 0$, то для составляющей суммы (14) с данным номером i возникает неопределенность вида $0/0$. Эту составляющую резонно будет исключить из рассмотрения. Введем множество номеров с положительными нагрузками. Пусть $I = \{i : \bar{y}_i > 0\}$.

Рассмотрим задачу

$$\sum_{i \in I} (\bar{y}_i - y_i)^2 / \bar{y}_i \to \min \qquad (15)$$

при ограничениях (2), (4), (5) и условии $\bar{y}_i = 0$ для $i \notin I$. Отметим, что в этой постановке ограничения-неравенства (3) для $i \in I$ не вводятся.

Теорема 3 [7]**.** *Множество оптимальных решений задачи* (15) *и множество решений исходной модели ОДМ с минимальным суммарным дефицитом, распределенным пропорционально нагрузкам узлов, совпадают.*

Итак, решая сразу на первом (и единственном) этапе вычислений задачу (15), получим требуемое решение с минимальным суммарным дефицитом, распределенным среди потенциально дефицитных узлов пропорционально их нагрузкам.

Для решения задачи (15) могут использоваться многие эффективные методы. В частности, ее можно решать методом внутренних точек [3]. В таблице представлены результаты расчетов по нескольким схемам ЭЭС методом внутренних точек [7]. В эксперименте сравнивались одно- и двухэтапные варианты модели ОДМ. Поскольку объем вычислений на одной итерации в алгоритмах внутренних точек, предназначенных для реализации модели в данных постановках, примерно одинаков, то можно считать, что соотношение количества итераций характеризует соотношение времени счета.

Таблица 1

Количество итераций метода внутренних точек для реализации двух постановок модели оценки дефицита мощности ЭЭС

Номер схемы	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Число узлов	2	5	5	6	10	11	23
Число связей	1	5	4	6	9	13	39
Количество итераций в исходной постановке	9	11	35	40	44	45	43
В том числе:							
ввод в допустимую область	2	1	2	2	1	1	2
минимизация дефицита	6	8	2	7	12	13	10
распределение дефицита по узлам	1	2	31*	31*	31*	31*	31*
Количество итераций для реализуемой модели в новой постановке	11	9	7	12	23	19	10

* В данных случаях окончание расчетов произошло не по критерию получения оптимального решения, а по заданному числу итераций.

Как видно из приведенных в таблице данных, одноэтапный вариант модели ОДМ эффективен в вычислительном отношении, требует меньшего объема и времени вычислений. При этом за счет исключения условий (3) в одноэтапном варианте модели упрощается процедура выбора допустимой стартовой точки, необходимой для начала работы алгоритма внутренних точек. Согласно табл. 1 для многих примеров окончание расчетов по двухэтапной модели ОДМ осуществлялось не в результате получения оптимального решения задачи (9), а по априори заданному числу итераций решения этой задачи. То есть полученное решение только из-за этого могло существенно отличаться от решения с распределением суммарного дефицита пропорционально нагрузкам узлов.

Модель с линейными потерями мощности в ЛЭП.

Одним из направлений повышения адекватности модели оценки дефицита мощности является учет потерь мощности при передаче ее по межузловым связям. При этом модель получает более реальный физический смысл.

В [8] предложена модель ОДМ, в которой учитываются потери мощности при ее передаче между узлами. Потери задаются в виде линейных функ-

ций от объема передаваемой мощности. Это повышает адекватность модели реальности и позволяет во многих случаях иметь однозначное распределение дефицита мощности между узлами.

Такая модель отличается от исходной тем, что вместо (2) используются балансовые ограничения вида

$$x_i - y_i + \sum_{j=1}^{n}(1-\beta_{ji})z_{ji} - \sum_{j=1}^{n} z_{ij} = 0,\ i = 1, \ldots, n. \qquad (16)$$

Здесь β_{ji} (коэффициент потерь на потоки мощности из узла j в узел i) – заданная величина из интервала $(0, 1)$.

Полученная модель (1), (3) – (5), (16) сводится к задаче о максимальном потоке с потерями, что является частным случаем задачи линейного программирования. Для ее решения существует много эффективных алгоритмов. В [8] для реализации модели используется один из вариантов двойственного симплекс-метода.

Если удельные потери по разным связям различаются, то решения по данной модификации модели ОДМ являются однозначными по распределению дефицита мощности. Этот дефицит концентрируется в тех узлах, поставки в которые сопряжены с большими потерями. Если удельные потери потоков мощности равны, в модели (1), (3) – (5), (16) возникает неоднозначность распределения дефицита мощности по узлам системы.

Модель с квадратичными потерями мощности в ЛЭП

В [7] предложена модель ОДМ, в которой потери мощности заданы в виде квадратичной функции от объема передаваемой мощности по ЛЭП. В этой модели балансовые ограничения (2) заменены на соотношение

$$x_i - y_i + \sum_{j=1}^{n}(1-\alpha_{ji}z_{ji})z_{ji} - \sum_{j=1}^{n} z_{ij} = 0,\ i = 1, \ldots, n. \qquad (17)$$

Коэффициенты α_{ij}, используемые для описания потерь при передаче электроэнергии из узла i в узел j, заданы и удовлетворяют условию

$$2\alpha_{ij}z_{ij} \le 1 - \mu, \text{ для всех } i \text{ и } j. \qquad (18)$$

Здесь μ – некоторая величина из интервала $(0, 1)$. Это условие означает, что дополнительная единица мощности, передаваемая из узла *i* в узел *j*, достигает узла *j* с положительным значением: при любом $z_{ij} \in [0, \bar{z}_{ij}]$.

Тройку векторов x, y, z, удовлетворяющих балансовым уравнениям (17) и неравенствам (3) – (5), будем называть допустимым решением задачи (1), (3) – (5), (17). Множество таких троек векторов образует множество допустимых решений.

Множество допустимых решений задачи (1), (3) – (5), (17), как правило, невыпукло (за исключением некоторых тривиальных случаев). Действительно, пусть существуют такие допустимые решения $(\tilde{y}, \tilde{x}, \tilde{z})$, $(\hat{y}, \hat{x}, \hat{z})$, что $\tilde{z} \neq \hat{z}$, тогда их выпуклая комбинация $(\lambda\tilde{y} + (1-\lambda)\hat{y}, \quad \lambda\tilde{x} + (1-\lambda)\hat{x}, \quad \lambda\tilde{z} + (1-\lambda)\hat{z})$ не удовлетворяет балансовым ограничениям (17). Значения функций в левой части ограничений (17) равны $\lambda(1-\lambda)\sum_{j=1}^{n}\alpha_{ji}\left(\tilde{z}_{ji} - \hat{z}_{ji}\right)^2$. Эти величины будут положительны для $\lambda \in (0, 1)$ и всех *i* таких, что существует номер *j*, при котором $\tilde{z}_{ji} \neq \hat{z}_{ji}$.

Для представления задачи (1), (3) – (5), (17) в виде эквивалентной задачи выпуклого программирования заменим ограничения (17) на следующие:

$$x_i - y_i + \sum_{j=1}^{n}(1-\alpha_{ji}z_{ji})z_{ji} - \sum_{j=1}^{n}z_{ij} \geq 0, \ i = 1, \ldots, n. \qquad (19)$$

Множество векторов, соответствующих ограничениям (3) – (5), (19), является выпуклым, для любой выпуклой комбинации двух допустимых решений $(\tilde{y}, \tilde{x}, \tilde{z})$, $(\hat{y}, \hat{x}, \hat{z})$ выполняются ограничения (1), (3) – (5), (19).

Задачу (1), (3) – (5), (17) будем называть исходной задачей минимизации дефицита мощности. Задачу (1), (3) – (5), (19) будем называть расширенной задачей минимизации дефицита мощности. Множество допустимых решений расширенной задачи содержит множество решений исходной задачи минимизации дефицита мощности. Покажем, что распределение дефицита

мощности по узлам системы, найденное в результате решения расширенной задачи, будет оптимальным и в исходной задаче.

Теорема 4 [9]**.** *Пусть значения переменных* $\tilde{y}_i,\ \tilde{x}_i, \tilde{z}_{ij},\quad i=1,\ldots,n,$ $j=1,\ldots,n,$ *являются оптимальным решением расширенной задачи, тогда существуют такие* $\hat{x}_i,\ \hat{z}_{ij}$, *что значения переменных* $\tilde{y}_i,\ \hat{x}_i,\ \hat{z}_{ij},\ i=1,\ldots,n,$ $j=1,\ldots,n,$ *составляют оптимальное решение исходной задачи минимизации дефицита мощности.*

Поскольку множество допустимых решений расширенной задачи содержит множество решений исходной задачи минимизации дефицита мощности, то из теоремы 4 следует справедливость обратного утверждения. Если $\tilde{y}_i,\ \tilde{x}_i, \tilde{z}_{ij},\ i=1,\ldots,n,\ j=1,\ldots,n$, являются оптимальным решением исходной задачи минимизации дефицита мощности (1), (3) – (5), (17), то они составляют оптимальное решение расширенной задачи (1), (3) – (5), (19).

Задача (1), (3) – (5), (19) имеет решение, так как множество допустимых решений не пусто (значения $y=0,\ x=0,\ z=0$ составляют допустимое решение), выпукло и ограничено (в силу условий (3) – (5)), целевая функция линейна. Из существования решения в задаче (1), (3) – (5), (19) следует (согласно теореме 4) существование решения в задаче (1), (3) – (5), (17). В [10] доказано, что решение в задаче (1), (3) – (5), (19) единственно по переменным, составляющим компоненты вектора y. Единственность распределения минимального суммарного дефицита по узлам системы в рассматриваемой модели гарантирует однозначность оценок вероятности и математических ожиданий дефицита в узлах системы после проведения расчетов.

Теорема 5 [10]**.** *Решение задачи* (1)*,* (3) – (5)*,* (18) *единственно по переменным* $y_i,\ i=1,\ldots,n.$

Ранее задача (1), (3) – (5), (17) решалась на базе алгоритма внутренних точек, основывающегося на итеративной линеаризации [3]. В [9] было предложено для решения задачи (1), (3) – (5), (19) использовать алгоритм внут-

ренних точек, в котором учитываются квадратичные аппроксимации ограничений. Результаты расчетов тестовых примеров, основанных на схемах реальных ЭЭС, показали [9], что с использованием алгоритма внутренних точек с квадратичными аппроксимациями решение получается за меньшее количество итераций (в среднем в полтора раза), чем алгоритмом, базирующемся только на линеаризации.

Заключение

Представлено описание математических свойств моделей оценки дефицита мощности, предназначенных для анализа надежности ЭЭС. Обсуждены особенности данных моделей, их достоинства и недостатки.

Для модели оценки дефицита мощности ЭЭС с квадратичными потерями мощности в ЛЭП указан способ представления модели в виде задачи выпуклого программирования. В [10] доказано, что в данной модели распределение дефицита мощности по узлам электроэнергетической системы единственно. Единственность распределения дефицита гарантирует однозначность оценок рассчитанных показателей надежности ЭЭС. Для реализации данной модели в [9] предложен алгоритм внутренних точек с квадратичными аппроксимациями. Алгоритм показал высокую вычислительную эффективность.

Возможные направления развития моделей ОДМ и программно-вычислительных комплексов анализа надежности:

а) распараллеливание вычислений в программных комплексах анализа надежности позволит одновременно «проигрывать» большее количество случайных ситуаций и тем самым получать более точные показатели надежности;

б) внедрение более эффективных алгоритмов внутренних точек, в том числе двойственных алгоритмов. В частности, как показали многие экспериментальные и некоторые теоретические исследования [11], использование двойственных алгоритмов внутренних точек позволяет сократить время решения исходной задачи оптимизации.

Список литературы

[1] Руденко Ю.Н., Чельцов М.Б. Надежность и резервирование в электроэнергетических системах. – Новосибирск: Наука, 1974. – 263 с.

[2] Александров И.А., Кузнецов Ю.А., Руденко Ю.Н. Общее и отличительное в исследовании надежности электроэнергетических и газоснабжающих систем // Методические вопросы исследования надежности больших систем энергетики. 1974. – Вып. 1. – С. 6 – 19.

[3] Дикин И.И., Зоркальцев В.И. Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). – Новосибирск: Наука, 1980. – 144 с.

[4] Рокафеллар Р. Выпуклый анализ. – М.: Мир, 1973. – 469 с.

[5] Зоркальцев В.И., Киселева М.А. Системы линейных неравенств. – Иркутск: Изд-во Иркут. гос. ун-та , 2007. – 129 с.

[6] Зоркальцев В.И. Относительно внутренняя точка оптимальных решений. – Сыктывкар: Коми фил. АН СССР, 1984. – 48 с.

[7] Зоркальцев В.И., Ковалев Г.Ф., Лебедева Л.М. Модели оценки дефицита мощности электроэнергетических систем. – Иркутск: ИСЭМ СО РАН, 2000. – 25 с.

[8] Чукреев Ю.Я. Модели обеспечения надежности электроэнергетических систем. – Сыктывкар: Коми НЦ УрО РАН, 1995. – 173 с.

[9] Зоркальцев В.И., Лебедева Л.М., Пержабинский С.М. Модель оценки дефицита мощности электроэнергетической системы с учетом квадратичных потерь мощности в линиях электропередач // Сиб. журн. вычислит. математики. – 2010. – Т. 13. – №3. – С. 285 – 295.

[10] Зоркальцев В.И., Пержабинский С.М. Модель оптимизации дефицита мощности электроэнергетической системы // Управление большими системами. – 2010. – Специальный выпуск 30.1 «Сетевые модели в управлении». – С. 300 – 318.

[11] Зоркальцев В.И. Об одном классе алгоритмов внутренних точек // Журн. вычислит. математики и мат. физики. – 2009. – Т. 49. – №12. – С. 2114 – 2130.

УДК621.311.1

ДЕКОМПОЗИЦИЯ И СИНТЕЗ СИСТЕМНОЙ НАДЁЖНОСТИ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ

Крупенёв Д. С.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: krupenev@isem.sei.irk.ru

Аннотация. В статье рассматриваются вопросы, связанные с обеспечением системной надёжности при долгосрочном проектировании развития электроэнергетических систем. Представлены методика декомпозиции и синтеза системной надёжности, методика определения математического ожидания недоотпуска электроэнергии в результате ненадежной работы связей электроэнергетической системы.

Ключевые слова. Электроэнергетическая система, системная (балансовая) надёжность, звено генерации, сетевое звено, декомпозиция.

Введение

Задача обеспечения системной надёжности электроэнергетических систем (ЭЭС) является одной из основных составляющих при планировании развития ЭЭС. Как правило, при долгосрочном планировании развития ЭЭС из системной надёжности выделяют балансовую составляющую (балансовую надёжность). Балансовая надёжность – это способность ЭЭС обеспечивать совокупную потребность в электрической мощности и энергии потребителей с учетом ограничений в виде плановых и неплановых отключений элементов ЭЭС, ограничений на поставку первичных энергоресурсов[1]. Именно обеспечению данного вида надёжности пойдёт речь в данной работе.

Среди множества задач обеспечения надёжности важной является задача оценки вклада в системную (балансовую) надёжность каждого из технологических звеньев ЭЭС: обеспечения электростанций первичными энергоресурсами, генерирования мощности, транспорта электроэнергии и её распределения по потребителям (рис 1.). Решение этой задачи даёт возможность определить «узкие» места в цепочке технологических звеньев ЭЭС и провести синтез системной надёжности путем гармонизации уровней надёжности исследуемых звеньев. Также при декомпозиции балансовой

надёжности есть возможность определить ответственность субъектов электроэнергетического рынка за низкую надёжность.

Рис. 1. Схема взаимосвязи технологических звеньев ЭЭС

Декомпозиции балансовой надёжности

Существующая практика оценки балансовой надёжности ЭЭС основанной на методе статистического моделирования (Монте-Карло) включает в себя три основных этапа [2].

1. Вероятностное моделирование множества расчетных состояний электроэнергетического оборудования и нагрузок потребителей в ЭЭС.
2. Оптимизация режимов расчетных состояний.
3. Вычисление показателей надёжности.

Данная методология оценки балансовой надёжности реализована в ИСЭМ СО РАН в виде программно-вычислительного комплекса «Янтарь». Опираясь на представленную методику оценки системной (балансовой) надёжности разработана методика декомпозиции балансовой надёжности. Для описания алгоритма декомпозиции балансовой надёжности рассмотрим свойства электроэнергетической системы, которые имеют важное значение для решения данной задачи:

- все элементы ЭЭС взаимосвязаны и находятся в непрерывном взаимодействии друг с другом;
- схема взаимосвязи технологических звеньев ЭЭС в плане надёжности соответствует рис.1;
- при «абсолютно» надёжном сетевом звене, ЭЭС становится концентрированным узлом, и любой генераторный агрегат влияет на надёжность любого потребителя одинаково.

С учётом изложенного, предлагаются следующие основные этапы алгоритма выявления вклада генерирующего и сетевого звеньев в системную надёжность, т.е. определение показателей надёжности для каждого из звеньев основной структуры ЭЭС.

1. *Опыт 1*: оценка надёжности основной структуры исследуемой ЭЭС для фактического состава и параметров элементов системы. Данный расчёт показывает уровень системной надёжности исследуемой ЭЭС. Полученные показатели надёжности отражают ситуацию по системе в целом, без расстановки акцентов на вклад в надёжность звеньев основной структуры ЭЭС.
2. *Опыт 2*: оценка надёжности исследуемой ЭЭС в предположении абсолютной надёжности сетевого звена с целью оценки вклада в системную надёжность ЭЭС надёжности генерирующего звена. В данном опыте, по сути, система работает в режиме концентрированного узла.
3. Расчет надёжности сетевого звена в предположении о последовательности соединения звеньев основной структуры ЭЭС (см. рис.1). На этом этапе предлагаемого алгоритма по формулам, приведённым ниже, вычисляются показатели надёжности сетевого звена, или, другими словами, доля вклада сетевого звена исследуемой системы в показатели системной надёжности, определённые в опыте 1.
4. Анализ полученных вкладов и выявление узких мест в звене генерации и сетевом звене ЭЭС. Гармонизация надёжности этих звеньев с целью

получения достаточной системной надёжности, то есть синтез сетевой надёжности.

Представление абсолютно надёжным сетевого звена в опыте 2 можно провести несколькими способами.

А. Для обеспечения абсолютной надёжности сетевого звена принимается аварийность линий равной нулю и увеличиваются пропускные способности связей до значений, не препятствующих пропуску требуемых потоков мощности в различных режимах. Увеличение пропускных способностей связей можно осуществлять различными способами: увеличивать пропускную способность каждой линии или увеличивать число линий в связи. Признаком достаточности выбираемых пропускных способностей является равенство во всех нагрузочных узлах системе в целом вероятностей бездефицитной работы, так как распределение дефицита мощности в системе происходит пропорционально нагрузкам в узлах.

Б. Ещё одним способом обеспечения абсолютной надёжности сетевого звена можно считать представление всей системы в виде одного концентрированного узла, т.е. узла, характеризуемого суммарной генерацией и нагрузкой по системе в целом в предположении, что сеть системы обеспечивает любые перетоки в любых режимах. Результатом расчёта являются только общесистемные показатели, из которых вероятность бездефицитной работы будет соответствовать и показателям в узлах. Суммарный недоотпуск электроэнергии, в этом случае, возможно, распределить по узлам пропорционально нагрузкам.

При декомпозиции балансовой надёжности показатели балансовой надёжности, вычисляемые в ПВК «Янтарь», представляются соответственно (рис.1) следующим образом.

Вероятность безотказной (бездефицитной) работы системы – вероятность того, что в пределах заданной наработки отказ системы не возникнет. Вероятность безотказной работы за заданное время t определяется по следующему выражению:

$$P(t) = 1 - F(t), \quad (1)$$

где: $F(t)$ – вероятность отказа системы за время *t*.

Цепочка звеньев системы: звена генерации, звена транспорта электроэнергии (сетевое звено) и звена потребления, с позиции теории надёжности, представляет последовательно соединённую цепочку. Таким образом, вероятность безотказной работы $P_{\text{сист}}$ при декомпозиции системной надёжности будет иметь следующий вид:

$$P_{\text{сист}} = P_{\text{ген}} P_{\text{сети}}, \quad (2)$$

где $P_{\text{ген}}$ – вероятность безотказного функционирования генерирующей части системы; $P_{\text{сети}}$ – вероятность безотказного функционирования сетевой части системы.

Данная формула согласно теории вероятностей соответствует случаю, когда два последовательно работающих звена системы обеспечивают работоспособность системы только при работоспособном состоянии обоих звеньев.

Математическое ожидание (м.о.) недоотпуска электроэнергии – это величина снижения потребления электроэнергии в результате ограничения или отключения нагрузки. М.о. недоотпуска электроэнергии является показателем надёжности, который позволяет перейти к экономическим показателя, т.е. зная величину недоотпуска электроэнергии и удельный ущерб от недоотпуска электроэнергии у потребителей можно найти ущерб у потребителей в денежном эквиваленте.

При декомпозиции системной надёжности недоотпуск электроэнергии в системе будет равен сумме недоотпуска электроэнергии, возникшего в результате отказов и дефицита в генерирующей части системы, и недоотпуска электроэнергии, возникшего в результате отказов и недостаточной пропускной способности сетевой части системы. М.о. полного недоотпуска электроэнергии в системе при декомпозиции системной надёжности примет следующий вид:

$$\Delta E_{\text{сист}} \approx \Delta E_{\text{ген}} + \Delta E_{\text{сети}}, \quad (3)$$

где $\Delta E_{\text{ген}}$ – недоотпуск электроэнергии, возникший в результате отказов и дефицита в генерирующей части системы, кВт·ч; $\Delta E_{\text{сети}}$ – недоотпуск электроэнергии, возникший в результате отказов и малой пропускной способности сетевой части системы, кВт·ч.

Приближённость данной формулы объясняется возможным появлением в системе недоотпуска из-за одновременных отказов генерирующего и сетевого оборудования, что в ЭЭС является редким событием.

Коэффициент обеспеченности электроэнергией потребителей – это отношение математического ожидания количества электроэнергии, отпущенной потребителям за заданный период времени, к требуемому её количеству за этот же период времени.

Коэффициент обеспеченности электроэнергией определяется из следующего выражения:

$$\pi = 1 - \Delta E \big/ E_{\text{треб}}, \quad (4)$$

где ΔE – м.о. недоотпуска электроэнергии потребителям, кВт·ч; $E_{\text{треб}}$ – требуемая выработка электроэнергии, кВт·ч.

При декомпозиции системной надёжности коэффициент π примет свою конечную форму исходя из следующих преобразований:

$$\Delta E = (1 - \pi) E_{\text{треб}}, \quad (5)$$

Запишем вместо ΔE значения $\Delta E_{\text{сист}}$, $\Delta E_{\text{ген}}$, $\Delta E_{\text{сети}}$, а вместо π значения $\pi_{\text{сист}}$, $\pi_{\text{ген}}$, $\pi_{\text{сети}}$:

$$\Delta E_{\text{сист}} = (1 - \pi_{\text{сист}}) E_{\text{треб}}, \quad (6)$$

$$\Delta E_{\text{ген}} = (1 - \pi_{\text{ген}}) E_{\text{треб}}, \quad (7)$$

$$\Delta E_{\text{сети}} = (1 - \pi_{\text{сети}}) E_{\text{треб}}, \quad (8)$$

Подставив выражение (6), (7), (8) в выражение (3), получим:

$$(1-\pi_{\text{сист}})E_{\text{треб}} = (1-\pi_{\text{ген}})E_{\text{треб}} + (1-\pi_{\text{сети}})E_{\text{треб}}, \quad (9)$$

После преобразований получим:

$$\pi_{\text{сист}} = \pi_{\text{ген}} + \pi_{\text{сети}} - 1, \quad (10)$$

При $\pi_{\text{сист}} > 0{,}99$ имеется возможность пользоваться более простым выражением:

$$\pi_{\text{сист}} = \pi_{\text{ген}}\pi_{\text{сети}}. \quad (11)$$

Для этого выразим $\pi_{\text{сист}}$, вычисляемого по (11), через $\pi_{\text{сист}}$, вычисляемого по (10), проведя необходимые преобразования:

$$\pi_{\text{сист}} = \pi_{\text{ген}}\pi_{\text{сети}} = \left(1-\frac{E_{\text{ген}}}{E_{\text{треб}}}\right)\left(1-\frac{E_{\text{сети}}}{E_{\text{треб}}}\right) = 1-\frac{E_{\text{ген}}}{E_{\text{треб}}}-\frac{E_{\text{сети}}}{E_{\text{треб}}}+\frac{E_{\text{ген}}E_{\text{сети}}}{(E_{\text{треб}})^2} =$$

$$= \pi_{\text{ген}} + \pi_{\text{сети}} - 1 + (1-\pi_{\text{ген}})(1-\pi_{\text{сети}}) = \pi_{\text{сист}} + \Delta\pi_{\text{сист}}.$$

Таким образом, погрешность значения $\pi_{\text{сист}}$, вычисленной по (11), по сравнению со значением $\pi_{\text{сист}}$, вычисленному по (10), составят:

$$\Delta\pi_{\text{сист}} = (1-\pi_{\text{ген}})(1-\pi_{\text{сети}}). \quad (11)$$

Если $\pi_{\text{сист}} = 0{,}99$, то $\pi_{\text{ген}}$ и $\pi_{\text{сети}}$ должны быть не ниже 0,995 и, следовательно, погрешность оценки $\pi_{\text{сист}}$ по формуле (11) по сравнению с (10) будет:

$$\Delta\pi_{\text{сист}} = (1-0{,}995)(1-0{,}995) = 0{,}005^2 = 25\cdot10^{-6},$$

т.е. второго порядка малости. Для $\pi_{\text{сист}} = 0{,}999$, соответственно, $\Delta\pi_{\text{сист}} = 25\cdot10^{-8}$ и т.д.

Анализ взаимосвязи и гармонизация показателей надёжности генерирующего и сетевого звеньев ЭЭС. Обеспечение системной надёжности ЭЭС

Поставленная задача формулируется как поиск экономически эффективного комплекса видов и объемов средств обеспечения необходимой надежности для задаваемых внутренних и внешних условий развития ЭЭС. Задачу можно разбить на ряд подзадач, таких как определение оптимального

состава и размещения генерирующего оборудования (синтез надёжности генерирующего звена ЭЭС); определение оптимальной конфигурации и параметров основной сети (синтез надёжности сетевого звена ЭЭС). Однако синтез надёжности указанных звеньев не представляется возможным независимо, вне рамок необходимой системной надёжности и надёжности электроснабжения потребителей. Смысл гармонизации как раз в том и состоит, чтобы синтез надёжности звеньев был согласован (скоординирован, взаимоувязан) между собой.

Выбор уровня обеспечения надёжности ЭЭС является технико-экономической проблемой. Задача оптимального обеспечения надёжности ЭЭС заключается в выборе таких параметров ЭЭС, дальнейшее улучшение которых нецелесообразно.

Определение оптимальных параметров ЭЭС должно осуществляться исходя из следующих положений.

1. Звено генерации ЭЭС должно обеспечивать нагрузку ЭЭС с учётом резерва генерирующей мощности по системе в целом, который должен быть экономически обоснован, то есть затраты на его содержание не должны быть выше ущерба у потребителей, возникшего в случае недостаточного резерва. Генерация по системе в целом не может быть ниже и даже равной абсолютному максимуму нагрузки (годовому). Норматив показателя надёжности генерирующего звена должен соответствовать этому резерву.
2. Сетевое звено ЭЭС должно обеспечивать передачу электроэнергии и мощности на уровне, необходимом для достижения принятого норматива показателя системной надёжности. Сеть должна по возможности беспрепятственно передавать генерирующую мощность потребителям во всех возможных режимах.
3. Уровни надёжности звена генерации и сетевого звена должны быть гармонизированы между собой с учётом системной надёжности.

Перед четвёртым этапом методики оказываются известными показатели системной надёжности и надёжности звеньев генерации и сети. В

процессе анализа полученных данных на этом этапе, прежде всего, используется уровень системной надёжности. Поскольку системная надёжность определяется показателями надёжности электроснабжения расчётных энергоузлов, образующих систему, в общем случае возможны следующие варианты. Системная надёжность:

а) ниже требуемого уровня во всех расчетных узлах;

б) ниже требуемого уровня только в отдельных узлах;

в) удовлетворяет требуемому уровню во всех узлах;

г) существенно выше требуемого уровня во всех узлах;

д) в одних узлах ниже требуемого уровня, в других существенно выше, в третьих – в норме.

Для выявления причин той или иной ситуации на следующем шаге анализа представляется целесообразным исследовать полученные в опыте 2 уровни надежности генерирующего звена. Здесь возможны только три варианта. Показатели надежности генерирующего звена во всех узлах:

а) ниже, требуемого уровня системной надежности;

б) удовлетворяет или не намного выше уровня системной надежности;

в) значительно выше требуемого уровня системной надежности.

Очевидно, что в случае недостаточной надежности звена генерации обеспечить требуемую системную надежность невозможно никаким другим способом в принципе, кроме как усилением генерирующего звена. В этом плане следует в первую очередь считать генерацию сетевого звена, обеспечивающего лишь транспортирование имеющейся мощности до центров питания. Поэтому первым шагом согласования («гармонизации») надежности основных технологических звеньев (генерации и основной сети) является обеспечение такого уровня надежности генерирующего звена, чтобы он (уровень) был достаточным для обеспечения системной надежности. Необходимый уровень генерирующего звена определяется из выражения (2):

$$P_{\text{ген}} = P_{\text{сист}} / P_{\text{сети}}$$

Поскольку $P_{\text{сети}} < 1$, то $P_{\text{ген}}$ должно быть больше $P_{\text{сист}}$. И если, для примера, принять, что нормативное значение $P_{\text{сист}} = 0{,}999$, то минимальное значение для $P_{\text{ген}}$ может быть принято на уровне 0,9991. При $P_{\text{ген}} = 0{,}9991$ надежность сети должна быть обеспечена согласно (2) на уровне $P_{\text{сети}} = P_{\text{сист}}/P_{\text{ген}} = 0{,}999/0{,}9991 = 0{,}9999$, что не всегда возможно или целесообразно реализовать.

Очевидно, что $P_{\text{ген}}$ можно принять и больше 0,9991 вплоть до величины, близкой к 1 (например, 0,9999), но это будет также явно дорогостоящим вариантом.

Вариант равноценной надежности звеньев находится также из (2)

$$P_{\text{ген}} = P_{\text{сети}} = P_{\text{сист}}{}^{1/2} = 0{,}999^{1/2} = 0{,}9995.$$

Аналогичные характеристики могут быть получены и для других нормативных значений $P_{\text{сист}}$ (0,996; 0,9996; 0,9998 и т. д.). Результаты данного анализа являются основой для нормирования надежности рассматриваемых технологических звеньев ЭЭС. Соответствующие данные приведены в табл.1.

Окончательный выбор рационального уровня надёжности генерирующего звена выполняется на основании технико-экономических расчётов с учётом исходной ситуации посредством традиционных приёмов проектирования ЭЭС: выбирается уровень необходимой установленной мощности по системе в целом, параметры дополнительно вводимого генерирующего оборудования и его размещение в системе по энергорайонам (узлам). С этой целью выполняются дополнительные оценки надёжности на базе проведения опыта № 2. При этом для выбора мест (узлов) размещения генерирующей мощности анализируются двойственные оценки в узлах генерирующей мощности и связей из опыта № 1. Для этого двойственные оценки ранжируются, от наибольших к наименьшим как по генерации, так и по связям. Поскольку двойственные оценки характеризуют недостаток

соответствующего ресурса в узлах, целесообразно размещать генерирующую мощность в узлах, где двойственные оценки по генерации выше всех и превосходят двойственные оценки по связям прилегающих к соответствующему узлу. Но этот выбор надо согласовать с обеспечением электростанций соответствующими первичными энергоресурсами.

Таблица 1

Диапазоны возможных значений Р для генерирующего и сетевого звеньев при заданных значениях системной надёжности

Уровень системной надежности $P_{\text{сист}}$	Минимальное значение $P_{\text{ген}}$ и соответствующее ему значение $P_{\text{сети}}$		Максимальное значение $P_{\text{ген}}$ и соответствующее ему значение $P_{\text{сети}}$		Вариант равноценной надёжности звеньев, $P_{\text{ген}} = P_{\text{сети}}$
	$P_{\text{ген мин}}$	$P_{\text{сети макс}}$	$P_{\text{ген макс}}$	$P_{\text{сети мин}}$	
0,996	0,9961	0,9999	0,9999	0,9961	0,998
0,999	0,9991	0,9999	0,9999	0,9991	0,9995
0,9991	0,9992	0,9999	0,9999	0,9992	0,99955
0,9996	0,9997	0,9999	0,9999	0,9997	0,9998
0,9997	0,9998	0,9999	0,9999	0,9998	0,99985
0,9998	0,9999	0,9999	0,9999	0,9999	0,9999

Если же в процессе проектирования на перспективу в исходном варианте имеется избыток генерирующей мощности, то используя критерий технико-экономической эффективности определяется оптимальный уровень состава генерирующего оборудования и обеспечиваемый при этом необходимый уровень надёжности генерирующего звена (опыт № 2) посредством удаления избыточной мощности.

Если же избыточный состав генерации уже установлен, то задача упрощается – принимается тот уровень надёжности генерации, который при этом получается.

Резюмируя, можно заключить, что уровень надёжности генерирующего звена должен быть выше требуемого уровня системной надёжности. В противном случае требуемая системная надёжность обеспечена быть не

может. А это значит, что установленная генерирующая мощность должна быть выше абсолютного максимума нагрузки системы в расчётном периоде на величину необходимого расчётного резерва мощности (полного резерва на период прохождения максимума нагрузки) в соответствии с устоявшейся проектной практикой.

Данный выбор надёжности генерирующего звена учитывается на следующем шаге выполнения этапа № 4 методики, на котором осуществляется гармонизация надёжности генерирующего и сетевого звеньев с учётом многих факторов: параметров и размещения генерирующего оборудования, параметров и конфигурации связей основной сети; исходной надёжности сетевого звена (результаты оценок этапа № 3 методики), характеризуемыми поузловыми показателями надёжности электроснабжения в системе; двойственных оценок сетевых элементов; недоотпусков электроэнергии от ненадёжности связей в системе; энергонадёжностных характеристик связей, а также технико-экономических характеристик генерирующего и сетевого оборудования.

В результате этих исследований определяется согласованный уровень надёжности генерации и сети при условии, что он обеспечивает требуемый (нормативный) уровень системной надёжности. Таким образом, посредством гармонизации надёжности технологических звеньев основной структуры ЭЭС выполняется синтез (оптимизация) системной надёжности.

Алгоритм гармонизации надёжности звеньев основной структуры ЭЭС показан на рис.2. Данный алгоритм тесно связан с алгоритмом декомпозиции системной (балансовой) надёжности, по сути, он является продолжением алгоритма методики декомпозиции системной надёжности. Основной акцент в данном алгоритме сделан на синтез системной надёжности исследуемой ЭЭС и здесь мы рассматриваем экономические критерии первоочередного усиления надёжности звеньев ЭЭС и подробно останавливаемся на синтезе сетевой надёжности.

Рис.2. Блок-схема методики гармонизации надёжности генерирующего и сетевого звеньев ЭЭС

Анализ сетевой надёжности ЭЭС

Для исследования надёжности связей ЭЭС и, в частности, для определения м. о. недоотпуска электроэнергии, который возникает у потребителей в результате аварийности и низкой пропускной способности связей ЭЭС предлагается использовать энергонадёжностные характеристики (ЭНХ) связей [3]. ЭНХ связи представляет собой функцию распределения мощности, передаваемой по данной связи в исследуемых условиях работы ЭЭС и при заданных надёжностных характеристиках оборудования ЭЭС. Название ЭНХ принято для того, чтобы отличать данную функцию распределения от функции состояний связи как отдельного объекта со своей производительностью, т.е. отличать от функции распределения вероятностей различных пропускных способностей связи как некоторой совокупности ЛЭП, обладающих аварийностью.

ЭНХ представляет существенную информацию об условиях работы рассматриваемой связи в рамках исследуемой ЭЭС. По этой информации

можно судить о степени использования ее пропускной способности, преимущественном направлении потоков мощности, соответствии пропускной способности имеющимся в системе возможностям взаимопомощи между узлами и т. д.

Способ определения ЭНХ связей заключается в следующем. По результатам расчетов потокораспределений в каждом из рассматриваемых расчетных состояний системы оцениваются величины и вероятности перетоков мощности. По полученным данным формируются ряды распределений фактически получающихся перетоков мощности по соответствующим связям. ЭНХ могут строиться как для всего расчетного периода (года), так и для отдельных интервалов расчетного периода (сезона, квартала, месяца). Расчеты по интервалам выявляют те периоды времени, когда интересующие нас связи слабо загружены и могут быть выведены в ремонт, реконструкцию и т.п.

В табл.2 в качестве примера приводятся значения функции ЭНХ 1-й связи, полученной по результатам расчета потокораспределений при оценке надёжности одной из ЭЭС. На рис.3 дан график ЭНХ для этой связи. На оси абсцисс отложена загрузка связи в различных режимах X в диапазоне их пропускных способностей, на оси ординат – вероятность (относительная длительность) загрузки связей p.

Таблица 2

Ряд распределения мощности, передаваемой по связи в исследуемых условиях работы ЭЭС

X, МВт	p	X, МВт	p	X, МВт	p	X, МВт	p	X, МВт	p
-900	0,0001123	-540	0,0268694	-180	0,1384509	180	0,0033820	810	0,0000274
-810	0,0000676	-450	0,1913126	-90	0,0577863	270	0,0007894	900	0,0009934
-720	0,0001709	-360	0,3045760	0	0,0250402	360	0,0002798		
-630	0,0014405	-270	0,2389479	90	0,0097187	450	0,0000340		

Рис. 3. Энергонадёжностная характеристика связи расчетной схемы исследуемой ЭЭС

Для количественной оценки сетевой надёжности при помощи ЭНХ связей была разработана методика, позволяющая находить недоотпуск электроэнергии, образующийся в результате аварийности и недостаточной пропускной способности исследуемых связей в ЭЭС.

Определение «вклада» каждой связи ΔE_j в общий недоотпуск электроэнергии производится в результате полученных данных в опыте 2 методики декомпозиции системной надёжности. Вычисляется искомый вклад по следующему выражению:

$$\Delta E_j = \Sigma(P_{ji} - \overline{P_{j0}})\,\rho_{ji}\,8760\text{, при } P_{ji} > \overline{P_{j0}}\text{,}$$

где P_{ji} – загрузка $j-$ой связи в $i-$ом расчётном режиме (принимается из ЭНХ опыта 2 методики декомпозиции системной надёжности), МВт; $\overline{P}_{j0}$ – математическое ожидание фактической пропускной способности исследуемой связи, состоящей из ЛЭП, отключаемых в аварийный или плановый ремонт, МВт; ρ_{ji} – вероятность нахождения связи в режиме передачи мощности P_{ji} (принимается из ЭНХ опыта 2 методики декомпозиции системной надёжности).

$$\overline{P}_{j0} = \sum_{k=1}^{K} \overline{P}_{jk}(1 - q_k)\text{,}$$

где $\overline{P}_{jк}$ – пропускная способность k – ой ЛЭП в связи j, МВт; $q_к$ – аварийность k – ой ЛЭП в связи; K – число ЛЭП в связи.

Графически данную процедуру можно представить следующим образом: при работе ЭЭС с искусственно завышенными пропускными способностями связей данные связи загружаются до значений, обеспечивающих беспрепятственную передачу необходимых потоков мощности в расчётных режимах ЭЭС (рис.3).

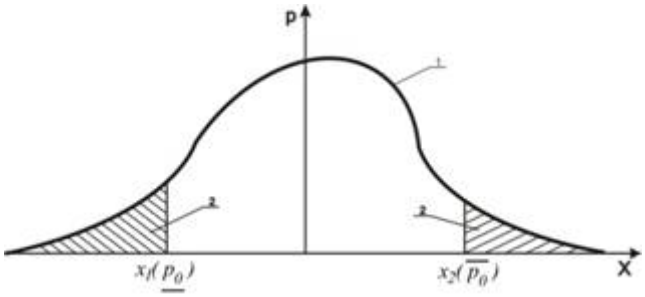


Рис. 4. Энергонадёжностная характеристика связи

На рис.4 $x_1(\underline{p}_0)$ и $x_2(\overline{p}_0)$ – фактические пропускные способности реальной связи в прямом и обратном направлениях. Области 2 являются тем недоотпуском электроэнергии, который возникает в результате малой пропускной способности конкретной связи.

Получив ΔE для всех связей, последние можно проранжировать по степени их вклада в системную надёжность и оценить приоритет первоочередного усиления соответствующих связей. После усиления связей проводится расчёт надёжности ЭЭС с усиленными связями. Если показатели системной надёжности приняли значения, удовлетворяющие принятому системному нормативу, то поставленная задача решена, если остались узлы, в которых принятый норматив не достигнут, необходимо повторить представленную выше процедуру и так далее, пока не будет получен нужный результат.

Следует отметить, что кроме описанной методики с использованием недоотпусков электроэнергии, возможно применение двойственных оценок для связей, получаемых при расчёте системной надёжности ЭЭС (опыт 1

методики декомпозиции системной надёжности). Двойственные оценки являются эффективным средством, позволяющим найти наиболее рациональные пути повышения сетевой надёжности. Их физико-техническая природа проявляется в виде вероятностей (относительных длительностей) превышения требуемых пропускных способностей связей над фактическими пропускными способностями. Получив двойственные оценки по каждой связи, последние можно проранжировать по важности первоочередного усиления связей. Как показали расчёты, двойственные оценки корреспондируют с недоотпусками электроэнергии из-за ненадёжности связей.

Заключение

1. Для оценки надёжности технологических звеньев основной структуры ЭЭС в докладе представлена методика декомпозиции системной (балансовой) надёжности. Использование данной методики позволяет определить уровни надёжности звена генерации и сетевого звена относительно системной надёжность, что, в свою очередь, обеспечивает наглядное отражение процессов, происходящих в системе с позиций надёжности и дает возможность производить синтез системной (балансовой) надёжности посредством гармонизации уровней надёжности генерирующего и сетевого звеньев.

2. Представлена методика анализа надёжности связей ЭЭС. Данная методика позволяет определить м.о. недоотпуска электроэнергии в результате ненадёжной работы (аварийности, малой пропускной способности) конкретных связей.

Список литературы

[1] Надёжность систем энергетики (Сборник рекомендуемых терминов). / Под ред. Н.И. Воропая. – М.: ИАЦ «Энергия», 2007. – 192 с.

[2] Ковалёв Г. Ф., Лебедева Л. М. Модель оценки надёжности электроэнергетических систем при долгосрочном планировании их работы // Электричество. – 2000. – №. 11. – С. 17 – 24.

[3] Крупенёв Д.С., Ковалёв Г.Ф., Лебедева Л.М. Исследование сетевой надёжности электроэнергетических систем. // Методические вопросы исследования надёжности больших систем энергетики. Вып. 60. Методы и средства исследования и обеспечения надёжности систем энергетики/ Отв.ред. Н.И. Воропай, А.И. Таджибаев (ПЭИПК). – СПб.: «Северная звезда», 2010. – С. 60 – 69.

УДК 621.311.1

ТЕОРИЯ, МЕТОДЫ И МОДЕЛИ АНАЛИЗА НАДЕЖНОСТИ ЭЭС

Ковалев Г. Ф.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: kovalev@isem.sei.irk.ru

Аннотация. Дается краткая характеристика основных положений теории надежности в применении к электроэнергетике, математические методы и вычислительные модели анализа надежности электроэнергетических систем (ЭЭС). Особое внимание уделено классификации методов и моделей, а также областей применения соответствующих методов и моделей. Характеризуются единичные свойства надежности и последовательность их реализации в практике проектирования ЭЭС. Описываются три основные подмодели модели оценки надежности исследуемого объекта.

Ключевые слова. Надёжность, объект, система, метод, модель, содержательная постановка задачи, единичные свойства, вычислительные программы, показатели, нормирование надежности.

Введение

Значительная специфичность и уникальность ЭЭС в большинстве случаев не позволяют непосредственно воспользоваться предлагаемыми в общей теории надежности математическими методами, моделями и алгоритмами вычисления показателей надежности, поскольку эта теория достаточно хорошо развита для так называемых простых систем и ограниченного класса сложных технических систем, в моделях которых не могут быть в достаточно полной мере отображены технологические особенности систем электроэнергетики, их многофункциональное и многоцелевое назначение, множественность вероятных состояний и доминирующая роль частичных отказов. Именно поэтому для оценки надежности объектов ЭЭС необходима разработка специальных математических методов и моделей, реализуемых затем в алгоритмах и вычислительных программах.

Основное назначение моделей надежности ЭЭС заключается в получении таких показателей надежности, которые могли бы напрямую или косвенно использоваться для принятия решений по обеспечению надежности объектов и ЭЭС в целом. Обеспечение надежности ЭЭС – сложная и многопла-

новая проблема, требующая своего решения в различных оценочных и оптимизационных задачах на разных уровнях территориально-временной и технологической иерархии управления. Представление же о комплексности свойства надежности как совокупности некоторых единичных свойств, актуальных для того или иного объекта, еще больше усложняет проблему анализа и синтеза надежности ЭЭС. В такой ситуации создание какой-то единой модели для решения всех задач надежности на всех временных этапах управления (при прогнозировании, проектировании, планирования развития, при долгосрочном, краткосрочном и текущем управлении функционированием), на всех территориальных уровнях (от оборудования до установки, объекта и системы различной степени объединения: регионального, единого для страны и выше), для всех технологических звеньев (звена первичных энергоресурсов, генерирующего звена, звеньев транспорта, преобразования и распределения электроэнергии)* и для всех единичных свойств (безопасности, долговечности, безотказности, ремонтопригодности, устойчивоспособности, живучести, управляемости, ресурсообеспеченности и сохраняемости) не представляется возможным не только сейчас, но и в обозримом будущем.

На различных уровнях и этапах отношение к точности и полноте решения задач надежности различно, то есть и в этом плане имеющиеся модели могут существенно различаться. Более того, для решения одной и той же задачи возможны модели разной степени полноты и точности представления исходных данных и результатов решения, соответственно. Таким образом, отказ от единой универсальной модели и переход к ее декомпозиции на множество моделей, решающих частные задачи надежности являются объективно обусловленными. Но при этом возникает особая проблема согласования решений частных задач надежности, которая на практике также должна решаться или решается тем или иным путем.

Как уже отмечалось, задача оценки надежности ЭЭС на разных терри-

* Разделение задач надежности по технологическим звеньям становится актуальной в последнее время в связи с реформированием электроэнергетической отрасли на принципах либерального рынка с заменой вертикально интегрированных систем на частично горизонтально интегрированные звенья отрасли.

ториальных, временных и технологических уровнях управления методически ставится и решается по разному в зависимости от цели исследования, ограничений на время решения, принятой расчетной схемы (глубины эквивалентирования), достоверности и форм представления исходной информации, номенклатуры вычисляемых показателей надежности, требований к точности получаемых результатов и используемого математического аппарата. Сообразно этому имеет место разработка многих различных моделей надежности в электроэнергетике.

В отечественной и мировой практике существует обширное множество методик и способов выполнения расчетов надежности ЭЭС и ее объектов [1–4]. Наряду со значительным количеством различающих их "нюансов", они обладают и определенными элементами общности, одинаковостью принципиальных подходов. Многообразие используемых математических методов и таких моделей позволяет считать полезным анализ их особенностей и потенциальных возможностей с целью выявления областей преимущественного применения. В этом смысле их можно классифицировать следующим образом.

Характеристика основных методов, используемых в моделях для оценки надежности объектов и систем электроэнергетики.

Методы, на которых базируются модели для оценки надежности ЭЭС и их объектов, известны в общей теории надежности, но их применение в данном случае отличается некоторыми особенностями.

Для разработанных способов и моделей исследования надежности характерно применение большого числа различных методов.

При всем разнообразии применяемых методов, – вопрос: какой из методов лучше, – не правомерен. Практика создания и использования моделей показывает, что предпочтение тому или иному методу перед другими должно определяться содержанием решаемой задачи управления ЭЭС. При этом в каждом случае преследуется цель получения достаточно точной, быстродействующей и удобной вычислительной программы.

По методологическим соображениям здесь делается попытка имею-

щееся многообразие методов достаточно наглядно классифицировать по некоторым принципиальным признакам.

Используемые в электроэнергетике методы можно классифицировать прежде всего с точки зрения информационной обеспеченности процесса определения надежности и применяемого математического аппарата. В этом смысле предлагается различать следующие две ***группы*** методов: экспериментальной оценки надежности (группа А) и расчетов и прогнозирования надежности (группа Б), см. табл. 1, левый столбец.

В *группе А* методы экспериментальной оценки надежности основаны на исследовании результатов специальных испытаний, проводимых на самих объектах или их физических моделях. Под специальными испытаниями ("испытаниями на надежность") понимается процесс определения или проверки показателей надежности опытным путем. Основной целью таких испытаний является создание информационной базы для соответствующих расчетных методов анализа и синтеза надежности или производственный контроль обеспечиваемых уровней надежности.

Опытное определение показателей надежности возможно на всех стадиях управления (но, к сожалению, не для всех объектов): проектировании, изготовлении и эксплуатации, - и в принципе связано с полными данными об объекте как физической реальности, хотя вопрос о представительности этих данных иногда остается открытым.

Особое место в этой группе занимают методы изучения в процессе проведения испытаний физико-химических и прочих причин отказов, требующих проведения многоцелевых экспериментов соответствующей природы.

В *группе Б* методы расчетов надежности имеют целью определение численных характеристик надежности объекта исследования при известных структуре, условиях работы и показателях надежности составляющих его элементов.

Методы в зависимости от используемых методологических принципов

и математического аппарата можно подразделить также на следующие ***классы*** (табл. 1, второй слева столбец):

I. Испытания на надежность, которые подразделяются на длительные и ускоренные.

Основным принципом длительных испытаний является воспроизведение реальных условий эксплуатации объекта. Ускоренные испытания характеризуются форсированными в сравнении с условиями эксплуатации режимами нагружения объекта с целью получения необходимой информации о надежности в возможно короткое время.

II. Методы, не требующие детального (поэлементного) моделирования объекта: ретроспективные, основанные на обобщении прошлого опыта; экстраполяционные, базирующиеся на анализе и прогнозировании сложившихся тенденций; экспертные, основанные на знаниях, опыте и интуиции специалистов.

Эти методы используются при прогнозировании для оценки численных значений показателей надежности объекта в условиях неполной определенности как количественных характеристик надежности элементов, составляющих объект, так и условий его функционирования.

III. Методы, основанные на поэлементном моделировании объекта исследования. Эти методы подразделяются на так называемые детерминистические и вероятностные подходы. Среди последних, с одной стороны, различают "аналитические", основанные на использовании функциональных соотношений в виде математических зависимостей, аналитических выражений вероятностных процессов, полном или усеченном переборе возможных состояний объекта, и "статистические", использующие аппарат метода Монте-Карло или псевдостатистические методы типа $ЛП_{\tau}$-программирование и прочее. С другой стороны, эти методы могут базироваться на представлении стохастических явлений случайными событиями или случайными процессами. К этому классу относятся также и методы физического моделирования, требующие чаще всего поэлементного представления сложного объекта.

Ниже приводятся особенности и области преимущественного применения основных методов (на уровне их ***видов***, табл. 1, третий столбец слева), используемых в моделях для оценки надежности объектов и систем электроэнергетики.

Уникальность и другие особенности ЭЭС не позволяют широко использовать экспериментальные методы оценки надежности (группа А, класс 1 в табл. 1).

Таблица 1

Классификация методов определения надежности ЭЭС и их объектов

<table>
<tr><td>Группы</td><td>Классы</td><td colspan="3">Виды</td><td>Области преимущественного применения</td></tr>
<tr><td rowspan="2">А. Методы экспериментальной оценки надежности.</td><td rowspan="2">I. Испытания на надежность реального объекта.</td><td colspan="3">1. Длительные испытания</td><td rowspan="2">Определение и проверка надежности оборудования ЭЭС</td></tr>
<tr><td colspan="3">2. Ускоренные испытания</td></tr>
<tr><td rowspan="5">Б. Методы расчетного определения надежности.</td><td rowspan="3">II. Методы, не требующие поэлементного моделирования объекта.</td><td colspan="3">3. Ретроспективные методы</td><td>Краткосрочное планирование работы ЭЭС</td></tr>
<tr><td colspan="3">4. Методы экстраполяции</td><td rowspan="2">Перспективное и долгосрочное планирование развития ЭЭС и их объектов</td></tr>
<tr><td colspan="3">5. Экспертные методы</td></tr>
<tr><td rowspan="6">III. Методы, основанные на поэлементном моделировании объекта исследования.</td><td rowspan="4">6. Детерминистические подходы</td><td colspan="2">6.1. Методы физического моделирования</td><td rowspan="6">На всех территориальных уровнях и временных этапах управления развитием и эксплуатацией ЭЭС в качестве основного или вспомогательного средства</td></tr>
<tr><td colspan="2">6.2. Критерий n-i</td></tr>
<tr><td colspan="2">6.3. Метод наихудшего случая</td></tr>
<tr><td colspan="2">6.4. Нормирование</td></tr>
<tr><td rowspan="4">7. Вероятностные подходы</td><td rowspan="2">7.1. Аналитические методы, базирующиеся на представлении реальных стохастических явлений</td><td>а) случайными событиями</td></tr>
<tr><td>б) случайными процессами</td></tr>
<tr><td rowspan="2">7.2. Статистические методы, базирующиеся на представлении реальных стохастических явлений</td><td>а) случайными событиями</td></tr>
<tr><td>б) случайными процессами</td></tr>
</table>

Методами математической статистики легко показать, что для получения достоверных данных о надежности энергообъектов и электрооборудования с помощью специальных испытаний на надежность требуются либо очень длительные испытания небольшого числа элементов, либо весьма большое число единиц оборудования (объектов) при непродолжительных испытаниях. И то, и другое для большинства электроэнергетических объектов, к сожалению, неприемлемо по экономическим соображениям, а также в силу уникальности и единичности отдельных видов оборудования.

Поэтому в электроэнергетике приняты только стендовые испытания одной-двух головных установок не столько для оценки надежности, сколько для выявления и устранения конструктивных дефектов. Это, естественно, повышает надежность, но не дает оценок показателей надежности.

Из ***испытаний на надежность*** наиболее применим в электроэнергетике метод длительных испытаний в его крайней форме – в форме анализа опыта эксплуатации (статистических данных) реальных объектов (систем) и их оборудования. По этой причине в ЭЭС формируются информационные системы и службы надежности, ведется большая и трудоемкая работа по сбору, обработке, хранению и использованию данных по надежности электроэнергетических установок с целью получения достаточно полных наборов реальных обобщенных характеристик их надежности.

Применение ускоренных испытаний ограничивается в электроэнергетике уровнем несложного оборудования и отдельных узлов и элементов более сложного оборудования.

Использование в моделях оценки надежности ЭЭС и их объектов методов группы Б в общем случае предпочтительнее в силу большей общности, строгости и меньшей зависимости от субъективных факторов. Эти методы используются на всех территориально-временных уровнях в качестве основного или вспомогательного средства.

Методы, не требующие поэлементного моделирования объекта исследо-

вания, подразделяются на следующие виды (табл. 1):

Ретроспективные методы – расчетные методы оценки надежности, основанные на анализе прошлого опыта функционирования объекта и обоснованном использовании этого опыта при прогнозируемых условиях его развития. Чаще всего при этом используется метод планирования экспериментов и регрессионный анализ.

Применение данных методов возможно при условии определенной "инерционности" в развитии объекта, то есть достаточной стабильности во времени его свойств и структуры, что в общем-то характерно для ЭЭС и их объектов. Эти методы применяются тогда, когда по каким-либо причинам не представляется возможным или целесообразным использование других методов. Преимущественной областью применения этих методов является анализ функционирования ЭЭС или отдельных их элементов при перспективном и долгосрочном планировании их развития.

Методы экстраполяции – методы прогнозирования надежности, основанные на анализе тенденций изменения показателей надежности работы объекта от изменения отдельных его параметров или условий работы. Эти методы, широко использующиеся для прогнозирования надежности оборудования ЭЭС, иногда применяются также и для оценки надежности объектов, подсистем и самих систем электроэнергетики в целом. Данные методы предполагают сбор представительной статистики за достаточно длительный промежуток времени или по большому числу объектов, что представляет определенную трудность. Другой проблемой при этом является то, что ЭЭС как объекты исследования изменяются во времени сложным образом, что накладывает серьезные ограничения на точность получаемых результатов, особенно при оценке надежности объектов ЭЭС на отдаленную перспективу.

Экспертные методы – методы оценки надежности, основывающиеся на накопленном в данной области знании, опыте и интуиции специалистов. Использование экспертных методов целесообразно в тех случаях, когда процесс

определения надежности в силу ограниченной информации о закономерностях явлений, связанных с объектом, неформализуем либо данные о развитии объекта существенно неопределенны.

Модели III класса подразделяются на два больших вида: детерминистические и вероятностные.

Детерминистическими называют такие методы и критерии анализа и синтеза надежности объекта (системы), в которых не моделируются вероятностные характеристики отказов элементов объекта, а только анализируется способность объекта выдерживать любые возмущения из априори заданного класса, то есть функционирование объекта после такого возмущения должно удовлетворять заданным условиям и параметрам. Для ЭЭС и их объектов – это допустимые уровни напряжений, частоты, загрузки оборудования, бездефицитность мощности и энергии.

Детерминистические подходы – это множество различных методов и критериев, основными из которых можно назвать следующие.

Физические модели для экспериментальной оценки надежности в принципе применимы, особенно на системном уровне, однако эта возможность пока что используется для исследования отдельных состояний или режимов системы в детерминированной форме, например, при исследовании живучести и вряд ли будет широко использоваться для вероятностного анализа из-за сложности проведения такого рода экспериментов.

Надежность объекта при отказе любых i элементов (надежность по *критерию (правилу)* $n–i$) есть свойство объекта выполнять основные функции при неисправности $i = 1, 2, 3 \ldots$ элементов из n даже при неблагоприятных, но реально возможных условиях. При этом загрузка всех элементов должна оставаться в допустимых пределах, режимные параметры не должны выходить за стандартные диапазоны, однако ухудшение эффективности работы объекта, как правило, допускается (рост расхода топлива, потери энергии и т.п.).

Метод «наихудшего случая» предполагает, что объект должен выпол-

нять свои функции, когда состав и параметры его элементов с учетом окружающих условий имеют предельные значения.

Нормирование предполагает учет фактора надежности посредством применения заданных правил и показателей надежности для структуры и параметров объекта.

Вероятностными называют такие методы и критерии, с помощью которых пытаются оценить, как часто и какое время объект (система) будет находиться в частичном или полностью неработоспособном состоянии из-за отказов элементов. Для такой оценки обязательным является моделирование вероятностных процессов в системе и вероятностных характеристик отказов элементов. Считается, что вероятностные подходы дают более разносторонние, глубокие и точные характеристики надежности объекта по сравнению с детерминистическими, однако они являются более сложными и трудоемкими.

В **аналитических методах** расчета надежности используются, как правило, основные положения теории вероятностей, комбинаторики, алгебры логики, теории массового обслуживания и т.п.

Аналитические методы при наличии математического описания функциональных связей между отдельными факторами принципиально позволяют решить любую задачу по оценке надежности в электроэнергетике с необходимой точностью. В практическом плане абсолютному преимуществу этих методов перед другими препятствуют: отсутствие либо громоздкость в некоторых случаях описания функциональных связей; "проклятие размерности", делающее невозможным в ряде случаев выполнение расчетов даже на современных ЭВМ за приемлемое время; трудности, связанные с вычислением некоторых показателей.

Статистические методы – методы расчета надежности, основанные на использовании статистического моделирования. При статистическом моделировании основные процессы функционирования объекта, включая стохастические, представляются многократно испытываемой вероятностной моделью.

Данные методы являются по существу методами "математического испытания на надежность".

Статистические методы используются в тех случаях, когда имеются препятствия для применения аналитических методов. В этом смысле они дополняют друг друга. Статистические методы отличаются относительной простотой учета факторов, определяющих надежность ЭЭС, и не накладывают ограничений на вид законов распределений рассматриваемых событий. Их главным недостатком является то, что обеспечение необходимой точности получаемых результатов иногда достигается ценой значительного увеличения числа реализаций (повторений опыта), что снижает достоинства этих методов. С целью устранения данного недостатка ведутся работы по поиску новых приемов и способов "разыгрывания" состояний. Остроту проблемы существенно снижает также рост производительности вычислительной техники и новые технологии выполнения расчетов, в частности, распараллеливание счета и др.

Методы (аналитические и статистические), базирующиеся на представлении вероятностных явлений в виде *случайных событий* используются в моделях надежности более широко по сравнению с методами, основывающимися на понятии случайного процесса. Это связано с тем, что многие явлений в ЭЭС с достаточной для решаемых задач точностью могут быть описаны на уровне случайных событий, а с другой стороны – по причине их большей простоты и разработанности соответствующего математического аппарата.

Наряду с методами, базирующимися на представлении стохастических явлений случайными событиями, в моделях надежности используются и методы, базирующиеся на представлении этих явлений *случайными процессами* (как аналитические, например, с использованием комбинаторного подхода, аппарата марковских и полумарковских цепей, так и опирающиеся на статистическое моделирование). Это вызвано тем, что методы, основанные на анализе случайных процессов в сочетании со статистическим моделированием на быстродействующих ЭВМ, позволяют получить более полную, а иногда и более точную

информацию о надежности исследуемого объекта, а постоянное развитие и усложнение ЭЭС этого требуют.

Классификацию методов можно было бы продолжить на более низком иерархическом уровне, однако практическая ценность этой классификации будет невысока в связи с тем, что на этом уровне должно было бы быть охарактеризовано постоянно расширяющееся множество конкретно применяемых методик, различающихся существенными или несущественными формализованными приемами представления условий функционирования, структуры систем и характеристик элементов, а также принимаемыми упрощениями и допущениями.

Так, к примеру, в области аналитических методов можно встретить следующие модификации:

- *логико-вероятностные методы* (функции алгебры логики, таблично-логический, метод "дерева отказов"), при которых структура объекта и особенности его функционирования описываются средствами алгебры логики, а расчет надежности производится с помощью теории вероятностей. Таблично-логический метод – частный случай метода случайных событий, основанный на составлении таблиц или матриц связей между режимами объекта, отказами его элементов и отказами объекта в целом;
- *метод структурных схем ("блок-схем")*, использующий условное графическое изображение элементов системы и связей между ними (структурную схему), при этом расчет показателей надежности производится с помощью теории графов или путем последовательного эквивалентирования;
- *метод фазового пространства*, исследующий случайный процесс функционирования сложного объекта на фазовом пространстве – множестве состояний, различающихся между собой составом исправных и неисправных элементов;
- *метод марковских процессов* (*метод пространства состояний*) – частный случай метода фазового пространства, в котором процесс функциониро-

вания объекта описывается с помощью системы дифференциальных уравнений переходов элементов от состояния к состоянию;

- *сетевой метод (метод минимальных путей и сечений)*;
- *топологические методы*

и т.д.

В области статистических методов различают:

- *функционально-статистический метод*, в котором процесс функционирования системы описывается вероятностной моделью, многократно испытываемой с помощью ЭВМ;
- *логико-статистический метод*, в котором структура системы и особенности ее функционирования описываются средствами алгебры логики, а расчет надежности производится с помощью статистического моделирования

и т.п.

Методы расчета надежности различаются также в зависимости от учета или неучета восстанавливаемости оборудования, учета только моментов появления *отказов* или учета *отказовых состояний* объекта, способов оптимизации дефицитных режимов и т.д.

В заключение необходимо обратить внимание на то, что приведенная в табл. 1 принципиальная классификация методов в определенном смысле является формально-абстрактной, так как на практике в одной и той же методике (модели) можно встретить сочетание различных методов, каждый из которых используется для решения частных подзадач. Как показывает опыт, именно такие комбинированные модели оказываются наиболее приемлемыми с точки зрения необходимой точности и быстроты счета. В этом смысле комбинированные модели – наиболее перспективные модели.

Характеристика моделей оценки (синтеза) надежности объектов (систем) электроэнергетики.

Все разнообразие моделей объектов ЭЭС для оценки (синтеза) их надеж-

ности в методологических целях может быть классифицировано по характеру их применения в соответствии с табл. 2.

Таблица 2

Классификация моделей надежности ЭЭС и их объектов

Характер использования модели	Степень полноты учета влияющих факторов	Точность представления учитываемых факторов	Быстродействие соответствующих программ на ЭВМ	Цель использования модели
Самостоятельная модель надежности	Практически полный учет	Наиболее точное представление	Низкое или среднее	Исследование надежностных свойств ЭЭС и их объектов, а также свойств соответствующих моделей
	Достаточно полный учет	Достаточно точное	Среднее	Сравнительный анализ надежности вариантов синтеза объектов управления --- Оптимизация надежности ЭЭС и их объектов
Модель надежности как часть оптимизационной модели для решения конкретной задачи управления	Достаточно полный учет	Приближенное	Достаточно высокое	Оценка надежности в процессе оптимизационного расчета
	Неполный учет	Приближенное, даже грубое	Высокое	Приближенный учет надежности в оптимизационной модели

При решении задач управления ЭЭС и ее объектами с учетом надежности вычисление показателей надежности осуществляется либо по отдельной модели надежности, либо в процессе решения основной задачи (алгоритмически совместно).

Наличие самостоятельных моделей оценки (синтеза) надежности позволяет при решении отдельных задач по развитию и эксплуатации ЭЭС ограничиваться грубым предварительным учетом надежности (вплоть до применения опосредованных норм и правил учета фактора надежности) с последующим углубленным анализом с точки зрения обеспечиваемой надежности. При этом уточнение значений показателей надежности проводится только для выявленных лучших вариантов. В ряде случаев данный способ в целом может требовать

меньших усилий и меньших затрат времени и обеспечивать высокую точность результатов.

Возможен и другой путь применения моделей надежности. При достаточных полноте и точности модели, а также ее удовлетворительном быстродействии, она может быть подблоком для оценки надежности в рамках соответствующего оптимизационного программно-вычислительного комплекса. В этом случае дополнительная оценка надежности получаемого решения отпадает. Периодичность работы такой модели в рамках вычислительного комплекса: на каждой итерации или в заключение формирования какого-либо варианта или этапа расчета, – будет определяться содержанием конкретных моделей, но, видимо, периодическая работа блока по соображениям быстродействия более целесообразна.

Большая роль отводится самостоятельным моделям в области **исследований** надежности как свойства объектов и систем электроэнергетики. Здесь они используются, кроме сравнительного анализа вариантов, еще в двух аспектах, отраженных в табл. 2. Указанные исследования дают возможность более

обоснованно формулировать новые методы и модели оценки надежности ЭЭС и их элементов.

В ближайшем прошлом описанные модели и методы применялись только для исследования надежности в смысле безотказности и ремонтопригодности (восстанавливаемости), но все чаще они начинают находить применение и при исследовании других единичных свойств надежности.

Относительно *единичных свойств надежности* при создании соответствующих моделей следует иметь в виду то обстоятельство, что единичные свойства, актуальные для ЭЭС, имеют разную степень значимости. Поэтому при их исследовании очень важно соблюдать определенную последовательность.

Последовательность рассмотрения единичных свойств надежности не так актуальна для функционирующих систем. По необходимости каждое свойство может быть оценено (синтезировано) отдельно от остальных. Но при проекти-

ровании последовательность обеспечения единичных свойств является принципиальной, так как, к примеру, бессмысленно заниматься обеспечением высокой живучести, не обеспечив предварительно необходимый уровень безопасности, долговечности, безотказности, ремонтопригодности и устойчивоспособности.

Известно, что:

– единичные свойства взаимосвязаны и взаимозависимы;

– относительная стоимость обеспечения необходимого уровня различных единичных свойств различна;

– в правильно ранжированной последовательности обеспечение каждого последующего единичного свойства обходится дешевле, если обеспечены все предыдущие свойства.

С учетом этих обстоятельств, рациональная последовательность единичных свойств надежности ЭЭС должна быть такой:

1. Безопасность.
2. Долговечность.
3. Безотказность.
4. Ремонтопригодность.
5. Устойчивоспособность.
6. Живучесть.
7. Управляемость.
8. Ресурсообеспеченность.
9. Сохраняемость.

Временной аспект надежности отображается такими понятиями, как балансовая надежность, режимная надежность, текущая (коммутационная) надежность и т.п. Для них также применимы рассмотренные выше методы и модели, но с учетом специфики рассмотрения проблемы.

Относительно моделей надежности есть еще одна проблема, которая както должна быть решена. Поскольку надежность, оцененная по той или иной модели, является «расчетной надежностью», которая неизбежно не может точ-

но совпадать с фактической надежностью объекта, а лишь в лучшем случае приближенно соответствовать ей из-за вводимых в модель допущений и упрощений, то представляется, что из множества моделей должны быть выбраны такие, которым можно было бы присвоить статус «*нормативных моделей*». Применение таких моделей для соответствующих расчетов обеспечивало бы сопоставительное сравнение надежности объектов на нормативной («узаконенной») основе. В мировой практике такие прецеденты имеются. Это, например, широко известные программы MEFISTO (Бельгия), PERU (Франция), модели «А» и «В» в США и др.

Вопросам моделирования надежности ЭЭС у нас в стране уделяли внимание Волков Г.А., Гольденберг Ф.Д., Гук Ю.Б., Лялик Г.Н., Малкин П.А., Обоскалов В.П., Розанов М.Н., Руденко Ю.Н., Фокин Ю.А., Чукреев Ю.Я. и другие (см. их публикации, например [5-10]). За рубежом - Биллинтон Р., Конкордия, Доду И., Аллан Р., Сальвадери Л., Эндрени Дж. и другие (см. их публикации, например [2-4]).

Модель оценки надежности объекта в общем случае состоит из трех подмоделей (этапов):

1. Оценка вероятностей рассматриваемых расчетных состояний объекта.

2. Блок оптимизации режимов работы объекта в расчетных состояниях, определенных на первом этапе.

3. Определение показателей надежности объекта по данным, полученным на первых двух этапах.

Новая экономическая среда (формируемая не только в России), как показывает практика, требует пересмотра классических подходов к указанным подмоделям. Методы, представленные в табл. 1, относятся, в основном, к этапу 1 и для современных условий подходят, практически, без изменений.

Блок оптимизации режимов для новых условий подлежит существенной переделке по сравнению с моделями надежности для прежней вертикально интегрированной ЭЭС. Несомненно, что новые организационно-коммерческие

структуры потребуют в блоке 2 (в дополнение к техническим требованиям) учета способности выполнять контрактные обязательства при различных взаимосвязях разных технологических звеньев системы (производства, передачи, распределения).

Блок 2 является центральным блоком модели оценки надежности. К нему предъявляются повышенные требования в части полноты учета основных факторов, точности и быстродействия.

Очевидно также, что политика, ориентированная на рынок и прибыль потребует определения новых, нетрадиционных показателей надежности. Все это в совокупности станет причиной значительного пересмотра, преобразований и модернизации имеющихся моделей надежности как у нас, так и за рубежом. И такая работа сейчас ведется.

В качестве примера здесь приводится описание блока оптимизации расчетного состояния ЭЭС, в котором учитываются характеристики функционирования оптового рынка электроэнергии.

В этом плане задача оптимизации режима может быть сформулирована так:

для известных значений располагаемой генераторной мощности, удельных затрат на производство электроэнергии, требуемых уровней покрытия и коэффициентов важности нагрузок в узлах, заданных пропускных способностей связей и коэффициентов потерь мощности в них, а также значений тарифов на электроэнергию для внутреннего рынка каждого энергоузла и внешних оптовых рынков найти загрузку генерирующего оборудования и значения покрываемой нагрузки в узлах по заданному критерию оптимальности с учетом ограничений на диапазоны возможных изменений генераторной мощности и нагрузки в узлах, перетоков по связям, а также выполнения балансов в узлах с учетом потерь мощности в сетях.

Математическая формулировка задачи выглядит следующим образом [12]:

$$\max \sum_{m=1}^{M} \{C_m(\overline{P}_m^{\text{н}} - y_m) + C_m^{\text{э}} \sum_{n \in N_m} x_n - f_m y_m - d_m(\overline{P}_m^{\text{г}} - z_m) + g_m z_m\} \quad (1)$$

при ограничениях

$$z_m - y_m + \sum_{n=1}^{N}(a_{mn}x_n + b_{mn}k_{\text{пот}\,n}x_n^2) = \ell_m, \quad m = 1,...,M;$$

$$0 \le y_m \le \overline{P}_m^{\text{н}}, \quad m=1,...,M; \quad (2)$$

$$0 \le z_m \le \overline{P}_m^{\text{г}}, \quad m = 1,...,M;$$

$$\underline{P}_n \le x_n \le \overline{P}_n, \quad n=1,...,N. \quad (3)$$

В приведенных выражениях M – число узлов, N – число связей расчетной схемы; N_m – множество связей, по которым m-й узел продает (покупает) мощность оптового рынка более высокого уровня по сравнению с локальным (внутренним) рынком; C_m – цена отпускаемой потребителям электроэнергии в узле m; $C_m^{\text{э}}$ – цена электроэнергии, предлагаемая m-ым узлом (частной компанией) на оптовый рынок; y_m – дефицит мощности в m-ом узле; z_m – избыток генераторной мощности в m-ом узле; x_n – поток мощности по n-ой связи; f_m – удельный ущерб (компенсационные затраты) за дефицит мощности в m-ом узле; d_m – удельные затраты на выработку и передачу электроэнергии в m-ом узле; g_m – стоимость топлива, необходимого для выработки единицы электроэнергии в m-ом узле; a_{mn} – элементы m-ой строки и n-го столбца матрицы связей; b_{mn} – элементы матрицы связи размерности $m \times n$, причем

$$b_{mn} = \begin{cases} 1, & \text{если } a_{mn} \cdot x_n > 0, \\ 0, & \text{если } a_{mn} \cdot x_n \le 0; \end{cases}$$

$$\ell_m = \overline{P}_m^{\text{г}} - \overline{P}_m^{\text{н}},$$

где $\overline{P}_m^{\text{г}}$ – располагаемая генераторная мощность в m-ом узле; $\overline{P}_m^{\text{н}}$ – требуемая мощность нагрузки в m-ом узле;

$k_{\text{пот}\,n}$ – коэффициент потерь мощности в n-ой связи; $\underline{P}_n$, $\overline{P}_n$ – пропускные

способности *n*-ой связи в обратном и прямом направлениях.

По результатам расчета режима с помощью данной модели можно получить:

– стоимость отпущенной потребителям электроэнергии;

– ущерб от недоотпуска электроэнергии;

– затраты на выработку электроэнергии;

– народнохозяйственный (экономический) эффект по системе в целом и по узлам;

– коммерческий эффект (прибыль) по системе в целом и по узлам (энергокомпаниям).

Вид функционала (1) может меняться в случае других условий функционирования конкурентного рынка и другого состава его субъектов.

Третий этап работы модели не составляет особых сложностей и представляет из себя обычный решатель, использующий данные, накопленные по результатам работы первых двух этапов.

Список литературы

[1] Ковалев Г.Ф., Малкин П.А. Программное обеспечение для оптимизации и оценки надежности при проектировании ЭЭС // Методические вопросы исследования надежности больших систем энергетики. Вып. 22. – Иркутск: СЭИ СО АН СССР, 1981. – С. 7–16.

[2] Methods and Techniques for Reliability Assessment of Interconnected Systems / Prepared by CIGRE T.F. 38.03.11. July 1997. – 60 pp.

[3] Биллинтон Р, Аллан Р. Оценка надежности электроэнергетических систем: пер. с англ. – М.: Энергоатомиздат, 1988. – 288 с.

[4] Эндрени Дж. Моделирование при расчетах надежности в электроэнергетических системах: пер. с англ. / Под ред. Ю.Н. Руденко. – М.: Энергоатомиздат, 1983. – 180 с.

[5] Васильев А.П., Гук Ю.Б., Карпов В.В. Надежность электроэнергетических установок и систем. Теория и практика. – СПб.: "ГУ Ленгосэнергонадзор", 2000. – 413 с.

[6] Обоскалов В.П. Надежность обеспечения баланса мощности электроэнергетических систем. – Екатеринбург: ГОУ ВПО УГТУ-УПИ, 2002. – 210 с.

[7] Надежность электроэнергетических систем. Справочник, т. 2 / под ред. М.Н. Розанова. – М.: Энергоатомиздат, 2000. – 568 с.

[8] Ковалев Г.Ф., Лебедева Л.М. Модель оценки надежности электроэнергетических систем при долгосрочном планировании их работы // Электричество, 2000. – № 11. с. 17–24.

[9] Розанов М.Н. Надежность электроэнергетических систем. – М.: Энергоатомиздат, 1984. – 280 с.

[10] Фокин Ю.А. Вероятностно-статистические методы в расчетах систем электроснабжения. – М.: Энергоатомиздат, 1985. – 210 с.

[11] Справочник по проектированию электроэнергетических систем / В.В. Ершевич, А.Н. Зейлигер, Г.А. Илларионов и др. Под ред. С.С. Рокотяна и И.М. Шапиро. – М.: Энергоатомиздат, 1985. – 352 с.

[12] Ковалев Г. Ф., Лебедева Л. М. Комплекс моделей оптимизации режимов расчетных состояний при оценке надежности электроэнергетических систем. Препринт № 7 ИСЭМ СО РАН. – Иркутск: ИСЭМ СО РАН, 2000. – 73с.

УДК 622.69.019.3

ПРОГРАММНО-ВЫЧИСЛИТЕЛЬНЫЙ КОМПЛЕКС ОЦЕНКИ НАДЁЖНОСТИ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ «ЯНТАРЬ». АЛГОРИТМИЧЕСКИЕ ОСОБЕННОСТИ

Лебедева Л. М.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: lebedeva@isem.sei.irk.ru

Аннотация. Приведено описание программно-вычислительного комплекса «Янтарь» оценки надёжности электроэнергетических систем и отмечены некоторые его особенности.

Ключевые слова: Электроэнергетическая система (ЭЭС), надёжность, программно-вычислительный комплекс, вероятность, оптимизация.

Введение

В модели оценки надёжности ЭЭС (рис. 1), реализованной в виде программно-вычислительного комплекса «Янтарь» выделяют три блока:

- вероятностный блок, где формируются случайные состояния режимов функционирования ЭЭС с использованием датчиков случайных чисел. Случайность позволяет «проигрывать» разные возможные ситуации при функционировании ЭЭС (случайные колебания нагрузки, отказы генерирующего и сетевого оборудования и др. [1]);
- блок оптимизации режимов расчётных состояний (оценка дефицита мощности по подсистемам (узлам) и системе (ЭЭС) в целом), сформированных в первом блоке в результате статистических испытаний) [2];
- блок вычисления показателей надёжности (в том числе вероятности бездефицитной работы системы и узлов, математического ожидания дефицита мощности и др.), в котором обрабатывается информация, накопленная в результате многократных оценок дефицитности или бездефицитности случайных состояний [1].

Рис. 1. Укрупнённая схема ПВК «Янтарь»

Из названных – блок оптимизации режимов (потокораспределения) расчётных состояний ЭЭС имеет основное значение. Характеристики этого блока определяют эффективность работы вычислительной модели надёжности ЭЭС в целом. Этим блоком, прежде всего, определяется точность расчётов надёжности и быстрота счёта, а также возможность учёта различных факторов, влияющих на надёжность и оптимальность функционирования ЭЭС.

Практика разработки моделей надёжности ЭЭС в ИСЭМ СО РАН также показывает, что технические и экономические изменения в работе современных ЭЭС в связи с изменениями внешних условий их функционирования отражаются на формировании новых подходов и требований к блоку оценки режима расчётного состояния. Для получения более полноценного результата оценки надёжности, отражающего и технические аспекты и экономические, в том числе и рыночные отношения, самым доступным способом является построение и использование нескольких моделей (блоков) оптимизации режимов расчётных состояний.

Вероятностный блок.

1. Вычисление комплексных характеристик факторов, определяющих надёжность узлов и системы в целом.

К комплексным показателям относятся:

– абсолютные максимумы нагрузок по узлам и системе в целом;
– требуемые выработки электроэнергии по узлам и системе в целом;
– число часов использования максимумов нагрузки.
– среднее значение нагрузки;
– коэффициенты заполнения графиков нагрузки;
– м.о. относительного пребывания оборудования в аварийном простое;
– и т.д.

2. Вычисление условных часовых нагрузок узлов, включающих в себя кроме мощностей электропотребления также мощности генерирующего оборудования, находящегося в плановых ремонтах.

3. Вычисление функции (ряда) распределения состояний генерирующей мощности каждого узла. В алгоритме «Янтаря» реализована возможность назначить шаг по мощности и ограничение на количество определяемых членов ряда распределения, которые обеспечивают необходимую точность расчётов. Отсутствие необходимости предварительного эквивалентирования состава агрегатов снижает трудоёмкость подготовки исходных данных.

4. Перемножение рядов распределений генерирующих мощностей. В «Янтаре» этот алгоритм реализован универсально, можно сказать «на все случаи жизни».

5. Вычисление функции (ряда) распределения состояний ЛЭП реализован так, позволяет вычислять ряды распределения состояний связей, состоящих практически, из любого числа параллельно работающих ЛЭП, и имеется возможность ограничения числа расчётных состояний наиболее вероятными.

6. Перемножение рядов распределений состояний ЛЭП и алгоритм

аналогично п. 4 имеет достаточно универсальный характер.

7. Композиция состояний системы. Под композицией состояний системы понимается определение и ранжирование возможных её состояний и соответствующих им вероятностей по полученным в программе рядам распределения состояний генерирующих мощностей и графикам нагрузки с учётом их случайных колебаний в узлах и состояниям ЛЭП по связям.

Рассмотрим сначала 3-й блок, а затем несколько подробнее 2-й.

Сбор результатов расчётов по интервалам и за год, и их статистическая обработка не представляет ничего сложного. Здесь в полном объёме используются статистические формулы для вычисления математических ожиданий, дисперсий и т.д. В этом блоке реализовано:

- вычисление энергонадёжностных характеристик ЛЭП (вероятности загрузки линий в пределах заданных ограничений с заданным шагом);
- вычисление показателей надёжности для каждого узла;
- вычисление показателей надёжности системы;
- обработка результатов расчётов.

Оптимизация режимов расчётных состояний. Этот блок может быть представлен решением разных задач и соответственно разными моделями и методами в зависимости от поставленной задачи. Для «Янтаря» – это в основном задача минимизации дефицитов мощности.

Изначально при разработке модели оценки надёжности ЭЭС блок оптимизации режимов был представлен моделью, обеспечивающей возможность минимизации дефицитов мощности (в дефицитных режимах) с учётом только 1-го закона Кирхгофа и распределением суммарного дефицита по узлам пропорционально нагрузкам в них. Это – линейная постановка задачи и как уже было сказано, рассматривает только дефицитные по системе состояния [3, 4].

Затем была разработана вторая модель, учитывающая потери мощности в сетях. Эта модель обеспечивает однозначное распределение суммарного дефицита мощности по узлам. Модель – нелинейная, неявно учитывает 2-ой закон Кирхгофа, но опять анализируются и рассчитываются только дефицитные

по системе состояния.

При рыночной экономике наряду с минимизацией дефицита должен использоваться и критерий минимизации затрат на производство и передачу электроэнергии. Так появилась третья модель, обеспечивающая распределение суммарного дефицита мощности по узлам по 2-ой модели в дефицитных режимах и минимизацию затрат на производство и транспорт электроэнергии в бездефицитных состояниях.

И, наконец, ещё одна модель, аналогичная 3-ей модели, но учитывающая особенности работы ЭЭС в условиях оптового рынка, хотя сам принцип существующего рынка неважен, важным является задание цены на продажную и покупную электроэнергию для каждого узла по каждой примыкающей к нему связи.

Формирование нескольких различных блоков оптимизации режимов расчётных состояний реализует комплексный подход к оценке надёжности и преследует цель – оценить влияние различных стратегий оптимизации режимов на результирующую надёжность ЭЭС.

Итак, задача, решаемая в программе Янтарь, имеет оценочный характер. Как оценочную вычислительную модель расчёта показателей надёжности ЭЭС рекомендуется использовать в тех случаях, когда временные или ресурсные ограничения не позволяют в полной мере обеспечить удовлетворение экономически обоснованных нормативов резервирования, а выбор средств обеспечения надёжности настолько сужен, что возможно инженерное формирование лишь нескольких альтернативных вариантов. В таких случаях с помощью оценочной модели осуществляется технический анализ вариантов, в результате которого рассчитываются и сравниваются с нормативами получающиеся показатели надёжности. Предпочтение, как правило, отдаётся варианту с более высоким уровнем надёжности и меньшими затратами.

Задача оценки надёжности с четвёртым вариантом блока оптимизации режимов расчётных состояний помимо технико-экономических характеристик, через стоимости электроэнергии учитывает особенности действия

рынков электроэнергии регионального и федерального уровней, поскольку эти особенности сказываются на результатах оптимизации как бездефицитных, так и дефицитных режимов. В этом плане задача оптимизации режима может быть сформулирована следующим образом:

для известных значений располагаемой генераторной мощности, удельных затрат на производство электроэнергии, требуемых уровней покрытия и коэффициентов важности нагрузок в узлах, заданных пропускных способностей связей и коэффициентов потерь мощности в них, а также значений тарифов на электроэнергию для внутреннего рынка каждого узла и внешних оптовых рынков найти загрузку генерирующего оборудования и значения покрываемой нагрузки в узлах по заданному критерию оптимальности с учетом ограничений на диапазоны возможных изменений генераторной мощности и нагрузки в узлах, перетоков по связям, а также выполнения балансов в узлах с учётом потерь мощности в сетях.

Вычислительный комплекс «ЯНТАРЬ» предназначен для оценки надёжности в смысле безотказности и ремонтопригодности (восстанавливаемости) больших сложных ЭЭС, представляемых любой (радиальной, кольцевой) многоузловой расчётной схемой с ограниченными пропускными способностями связей между узлами. Задача решается в условиях управления развитием и долгосрочного планирования функционирования на уровнях Единой, объединенных и районных ЭЭС.

Электроэнергетическая система представляется как совокупность узлов совместной нагрузки и генерации, узлов только нагрузки и узлов только генерации (может быть с учётом собственных нужд) и связей между ними. В качестве элементов системы принимается основное оборудование ЭЭС (генерирующие агрегаты и линии электропередачи (ЛЭП)). Каждый узел в такой схеме является концентрированным, который в нынешней интерпретации может представлять «зону свободного перетока». Связи между узлами расчётной схемы представляют собой совокупности всех ЛЭП между соответствующими регионами, замещаемыми данными узлами.

С помощью «Янтаря» возможен интерактивный режим оптимизации величины и размещения резервов генераторной мощности, пропускных способностей связей и запасов энергоресурсов с учётом фактора надёжности.

Исходные данные

- расчётная схема ЭЭС (эквивалентные узлы и связи между ними);
- характерные суточные графики нагрузок в каждом из узлов, отнесенные к единому (например, московскому) времени или задаваемым отклонением часового пояса;
- продолжительности соответствующих расчётных интервалов (например, месяца года) и периодов в интервале (например, рабочих и выходных дней);
- с.к.о. нагрузок от регулярных почасовых значений;
- графики максимумов нагрузки по расчётным интервалам и узлам (например, месячные);
- состав и параметры генерирующих агрегатов для каждого из узлов и каждого расчётного интервала;
- состав и параметры линий электропередачи для каждой связи в каждом расчётном интервале;
- вероятности аварийных простоев генерирующих агрегатов и линий электропередачи;
- нормативы на проведение текущих ремонтов элементов ЭЭС в течение года;
- нормативы на проведение капитальных и средних ремонтов элементов ЭЭС за год;
- ограничения на использование первичных энергоресурсов (топлива на ТЭС, воды на ГЭС);
- удельные приведенные затраты в оборудование ЭЭС по узлам и связям;
- компенсационные затраты при недоотпуске электроэнергии потребителям по узлам.

Из-за отсутствия реальной информации часто для получения исходных

данных приходится обращаться к [5].

Результаты расчётов

– Показатели надёжности по узлам и системе в целом за каждый расчётный интервал и год в целом:

- вероятность бездефицитной работы;
- средний недоотпуск электроэнергии;
- коэффициент обеспеченности потребителей электроэнергией.

– По известным методикам [1] в модели рассчитываются мощности и состав агрегатов, находящихся в текущем и капитальном (среднем) ремонтах в каждом расчётном интервале.

– Значения расчётных резервов различных видов по узлам и системе в целом:

- резерв на проведение капитальных (средних) ремонтов;
- резерв на проведение текущих ремонтов;
- резерв на модернизацию;
- оперативный резерв;
- полная величина резервов всех видов.

– Значения дополнительных пропускных способностей связей для обеспечения взаиморезервирования ЭЭС и функции распределения загрузки этих связей в течение года.

– Значения требуемых первичных энергоресурсов.

– Двойственные (объективно обусловленные) оценки дефицитности основных ресурсов (генераторной мощности и энергоресурсов в узлах и пропускных способностей связей) для обеспечения надёжности электроснабжения потребителей.

Метод решения

Задача решается методом имитационного моделирования работы ЭЭС в течение расчётного периода времени, за который принимается год. Состояние нагрузок и оборудования системы разыгрываются методом Монте-Карло.

Предварительно рассчитываются ряды распределения мощностей агрегатов в узлах и ЛЭП в связях, находящихся в аварийном простое. В качестве производящей функции используется биноминальное выражение.

Для решения задач оптимизации режимов расчётных состояний использовался, один метод в разных его модификациях – это метод внутренних точек [4].

Программное обеспечение.

Программно-вычислительный комплекс оценки надёжности электроэнергетических систем «Янтарь» написан на языке программирования Фортран. Этот язык программирования был выбран не случайно. Он, как раньше, так и сейчас не теряет своей актуальности программирования алгоритмов решения физико-технических задач, где важным является точность выполнения арифметики, а Фортран изначально имел лучшую реализацию выполнения арифметики на вычислительной технике, в том числе при работе одновременно с большим разбросом числовой информации (от 10^{-18} до 10^{+18}). Это особенно важно при работе с вероятностными характеристиками, например при перемножении рядов распределения вероятностей или при работе оптимизационных методов, где необходимо перемножение очень больших чисел. Кое-где в алгоритме возникает необходимость использовать Double Precision.

Максимальные параметры расчётной схемы: 100 узлов, 160 связей. Количество рассматриваемых характерных суточных графиков максимально 600, количество членов строящихся рядов распределения вероятностей максимально – 60 – это достаточно много, количество линий в одной связи максимально – 12, и т.д.

Для иллюстрации (без комментариев) вид одной из схем, по которой проводились расчёты по оценке надёжности ЕЭС России приведен на рис. 2.

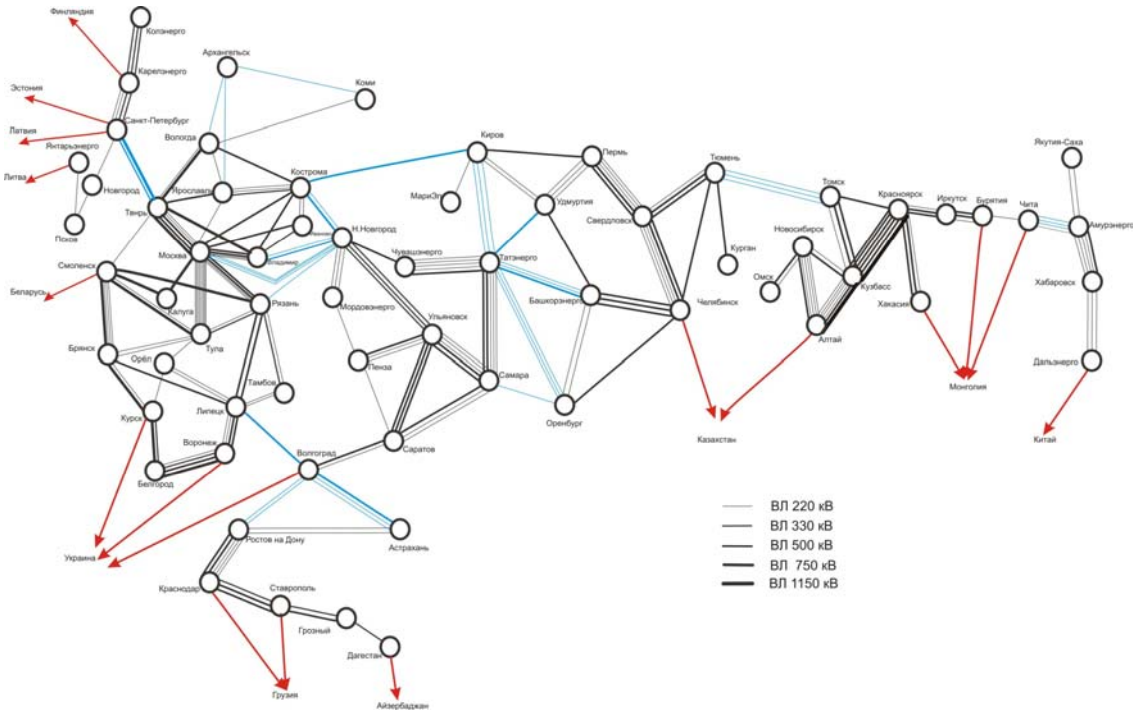


Рис. 2. Схема единой электроэнергетической системы России

Список литературы

[1] Ковалев Г.Ф., Лебедева Л.М. Модель оценки надежности ЭЭС при долгосрочном планировании их работы // Электричество, 2000. –№11. С. 17 – 24.

[2] Ковалев Г. Ф., Лебедева Л. М. Комплекс моделей оптимизации режимов расчётных состояний при оценке надёжности электроэнергетических систем. Препринт № 7 ИСЭМ СО РАН. – Иркутск: ИСЭМ СО РАН, 2000. – 73 с.

[3] Дикин И.И. Применение алгоритмов метода внутренних точек к минимизации дефицита мощности в электроэнергетической системе // Методические вопросы исследования надёжности больших систем энергетики. – Иркутск: АН СССР, СЭИ, 1978. – С. 23 – 33.

[4] Дикин И.И., Зоркальцев В.И. Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). – Новосибирск: Наука, 1980. – 144 с.

[5] Справочник по проектированию электроэнергетических систем / В.В. Ершевич, А.Н. Зейлигер, Г.А. Илларионов и др. Под ред. С.С. Рокотяна и И.М. Шапиро. – М.: Энергоатомиздат, 1985. – 352с.

УДК 621.3.019.3: 658.26: 697.34

МЕТОДЫ КОМПЛЕКСНОГО АНАЛИЗА НАДЕЖНОСТИ ТЕПЛОСНАБЖЕНИЯ

Стенников В. А.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: sva@isem.sei.irk.ru

Постников И. В.

Институт систем энергетики им. Л.А. Мелентьева СО РАН

E-mail: postnikov@isem.sei.irk.ru

Аннотация. В работе рассматривается проблема комплексного анализа надежности теплоснабжения потребителей, которая заключается в количественной оценке влияния всех этапов производства и распределения тепловой энергии на надежность теплоснабжения каждого потребителя системы. Представлен краткий обзор существующих методов анализа надежности топливо- и теплоснабжения, на основе которого обосновываются ключевые подходы к решению задачи комплексного анализа надежности теплоснабжения. Предлагается методический подход, в рамках которого формулируются математические модели и методы узловой оценки надежности теплоснабжения потребителей, а также излагаются результаты исследований по влиянию на ее уровень различных элементов систем топливо- и теплоснабжения. В основе математического моделирования лежат марковские случайные процессы, модели потокораспределения в тепловой сети, детерминированные зависимости теплофизических процессов потребления тепловой энергии и некоторые другие модели.

Ключевые слова. Теплоснабжающий комплекс, анализ надежности теплоснабжения, метод статистических испытаний, марковский случайный процесс, аварийный режим, декомпозиционный анализ надежности, узловой показатель надежности.

Введение

Теплоснабжение является важнейшей составляющей в обеспечении жизнедеятельности населения и развития практически всех отраслей экономики. Высокая социально-экономическая значимость сферы теплоснабжения предъявляет повышенные требования к надежности и экономичности теплоснабжающих систем (ТСС), которые объединяют в единой структуре источники тепловой энергии (ИТ) и тепловой сети (ТС). В течение длительного периода своего формирования и развития эти системы приобрели сложную структуру и большие размеры. Такие особенности современных ТСС, как множество источников тепловой энергии различного типа, большая протя-

женность тепловых сетей (сотни километров), объединенных в сложные многокольцевые схемы, множество потребителей тепловой энергии, многоуровневая структура организации управления теплоснабжающими системами, обуславливают сложность задач исследования, обеспечения и повышения их надежности. Эти задачи связаны главным образом с определением уровня надежности теплоснабжения потребителей и учетом нормативных требований к значениям показателей надежности при проектировании систем и их эксплуатации. Рост масштабов ТСС и соответствующее усложнение их структуры, связанные с увеличением числа потребителей и их нагрузок, выявили необходимость создания эффективных методов оценки их надежности и мер по обеспечению надежного теплоснабжения. Положение усугубляется разобщенностью структуры управления ТСС, техническим их несовершенством, значительным износом теплопроводов и оборудования, низким технологическим уровнем эксплуатации некоторых объектов системы, теплогидравлической разрегулированностью тепловых сетей и абонентских установок. Также необходимо развитие и совершенствование методической базы по нормированию надежности теплоснабжения.

Первые исследования по надежности ТСС с использованием методов теории надежности начали развиваться в конце 1960-х – начале 1970-х годов [1, 2]. Первоначально повышение надежности ТСС связывали только с улучшением качества элементов и конструкций, а соответствующие теоретические подходы не учитывали многих особенностей функционирования этих систем. Но надежная элементная база так и не была создана, а основной тенденцией в сфере теплоснабжения стало развитие крупных централизованных систем. В таких условиях одним из основных методов обеспечения надежности многие специалисты считали резервирование схем тепловых сетей. В дальнейшем было проведено большое количество исследований по надежности ТСС и ее элементов, были разработаны методики и модели решения задач анализа надежности источников тепловой энергии и тепловых сетей, обладающие спецификой, отражающей их различные свойства, разработан ал-

горитм построения ТСС с требуемым уровнем надежности, обоснована система нормативов, необходимых для ее решения [2].

Функционирование ТСС начинается с обеспечения источников топливом. В связи с этим надежность топливоснабжения неразрывно связана с надежностью и качеством отпуска тепловой энергии потребителям и учет топливной составляющей является неотъемлемым элементом при системном подходе к решению задач анализа и синтеза надежности теплоснабжения наряду с надежностью производства и передачи тепла.

Практическая реализация всех методических разработок в области надежности теплоснабжения в значительной степени зависит от наличия информации о показателях надежности элементов систем тепло- и топливоснабжения и систематизации информации о режимах топливопотребления в зависимости от различных внешних условий. За многолетний период эксплуатации ТСС накоплено некоторое количество статистических данных по отказам и восстановлениям их элементов, в результате обработки которых получены характеристики надежности этих элементов, в том числе нормативные. Однако, проектирование новых систем и реконструкция существующих производятся с учетом внедрения современного энергосберегающего оборудования, характеристики надежности которого невозможно объективно оценить вследствие недостаточного опыта его эксплуатации. В связи с этим необходимо создание систем автоматического сбора информации на объектах всего технологического процесса производства и распределения тепловой энергии.

В последние годы актуальность работ по исследованию надежности теплоснабжения возросла в связи с высокой степенью износа оборудования ТСС (например, основное оборудование многих ТЭЦ в России построено более 50 лет назад, износ тепловых сетей в среднем по стране достигает 60%), формированием новых форм управления тепловым хозяйством городов с участием множества собственников в единой системе, а также перспективными направлениями развития ТСС на основе принципов распределенной

генерации, что требует новых подходов к анализу, нормированию и обеспечению надежности реконструированных систем.

Краткий анализ существующих методов анализа надежности топливо- и теплоснабжения

Особенности процессов топливоснабжения и их моделирование. Совокупность узлов добычи топлива, его хранения и транспортировки до ИТ представляет собой систему топливоснабжения (СТС). Системы топливоснабжения различны по виду топлива, структуре и масштабам. Общая укрупненная технологическая цепочка топливоснабжения состоит из множества узлов добычи топлива (шахты, скважины), системы транспортирования топлива до предприятий переработки (обогатительные фабрики, нефте- и газоперерабатывающие заводы), системы вторичного транспортирования готового продукта (транспорт общего пользования, продуктопроводы), объекты хранения топлива (угольные базы и склады, нефтебазы, подземные хранилища газа) и системы распределения до потребителей топлива (ИТ и других предприятий).

Оценка и обеспечение надежности этих систем являются сложными задачами, решение которых затруднено распределенной структурой объектов, функционирующих в условиях влияния множества внутренних и внешних факторов (в том числе случайных), и в большей степени, чем для ИТ и ТС, недостатком информации о надежности элементов системы. Проблема надежного топливоснабжения ИТ в течение отопительного сезона связана также с неравномерностью тепловых нагрузок, имеющих случайный характер вследствие изменения метеорологических условий. Так как топливоснабжение ИТ включает в себя различные стадии добычи, транспортировки, хранения топлива, то анализ его надежности должен осуществляться комплексно – с учетом взаимодействия всех элементов, участвующих в этом процессе.

Основным геофизическим показателем, влияющим на изменение потребности в топливе на цели теплоснабжения, является колебание температуры наружного воздуха, менее значительными факторами являются измене-

ния ветровых нагрузок и количества солнечного излучения [3, 4]. Технологические факторы связаны с характером производственных процессов предприятий, производящих, транспортирующих и потребляющих топливо: технологии добычи топлива, качество топлива и способы его сжигания, взаимозаменяемость топлив между ИТ одной системы, возможность складирования топлива, пропускная способность транспортных систем и т.п. Существует также класс факторов, связанных с организационными мероприятиями, влияющими на режимы работы топливодобывающих и топливоперерабатывающих предприятий, изменениями плановых заданий, нарушениями норм расхода топлива и энергии, планов подачи транспортных средств и т.п.

Все указанные выше факторы сложным образом влияют на режимы топливопотребления и топливоснабжения, обуславливая регулярные и случайные колебания их в суточном, сезонном и многолетнем разрезах. Постоянно проявляется несоответствие во времени неравномерных режимов топливопотребления и добычи (производства) отдельных видов топлив [5, 6, 7].

Надежность топливоснабжения ИТ характеризуется соотношением потребностей и поставок топлива за определенный период времени с учетом имеющихся запасов. Проблема надежного обеспечения топливом решается в первую очередь управлением запасами топлива на различных технологических и временных уровнях с учетом взаимозаменяемости топлив между его потребителями, а также созданием резервов производственных и мощностей добычи и переработки топлива.

Существующие методы анализа надежности топливоснабжения ИТ, как правило, предназначены для определения укрупненных показателей обеспечения топливом на уровне топливно-энергетического комплекса региона или его отдельных узлов. Задача надежного обеспечения топливом ИТ решается в первую очередь при оценке надежности, так как на данном этапе рассчитываются дефициты, определяющие потребности в резервах и запасах. В условиях практической сложности построения расчетных схем СТС с учетом характеристик надежности ее элементов одним из наиболее обоснованных ме-

тодов анализа надежности топливоснабжения является метод статистических испытаний, не требующий данных об отказах и восстановлениях элементов. Метод статитстических испытаний (метод Монте-Карло) широко применяется при моделировании различных систем как для оценки их надежности, так и для решения прочих задач [8, 9].

Методика для анализа надежности топливоснабжения, основанная на модели статистических испытаний, предложена в [10]. Имитационная модель предназначена для изучения надежности топливоснабжения теплогенерирующих установок с учетом влияния на надежность обеспечения потребности в топливе следующих факторов: объемов страховых запасов топлива и емкости складов, объемов планируемых на год поставок топлива и достоверности реализации этих планов и др. При этом в модели не учитываются внутригодовые неравномерности процессов топливоснабжения, потребность в топливе, выделяемые ресурсы, поставки, потребность и дефицит даются в объеме годовых показателей. Объемы выделяемых ресурсов топлива и годовой потребности описываются как случайные величины, варьируемые вокруг некоторого планового уровня с заданными плотностями вероятности их реализаций. Алгоритм строится на основе многовариантных процедур генерации этих случайных величин, имитирующих в многолетнем разрезе реальный процесс поставок, создания и использования запасов топлива. В результате расчетов определяются частота возникновения и средняя величина дефицита и избытка топлива, вероятности возникновения дефицита и математическое ожидание дефицита топлива. Проведенные в рассматриваемой работе исследования позволяют укрупнено для всей энергосистемы оценить необходимые запасы топлива для обеспечения заданного уровня надежности топливоснабжения. Для согласования этих решений с результатами совместной оценки надежности ИТ и ТС для определения их суммарного влияния на теплоснабжение потребителей требуется развитие предложенных методов.

Методы анализа надежности ИТ. Существуют различные методы анализа надежности ИТ, которые можно разделить на аналитические, основан-

ные на использовании марковских [2, 11, 12, 13, 14, 15], полумарковских моделей [16, 17, 18], логико-вероятностных методов, и методы статистического моделирования [8, 9].

Методы первой группы рассматривают ИТ как совокупность функционально соединенных элементов, объединенных в расчетную схему для расчета надежности. Использование марковской модели для анализа надежности сложных схем ИТ связано с большой размерностью систем уравнений случайного процесса, решение которых на современных компьютерах само по себе не представляет трудностей, однако в некоторых случаях для уменьшения объема вычислений и построения более компактных моделей используют известный из теории надежности технических систем метод декомпозиции [11]. В частности, согласно методике оценки надежности ИТ, представленной в [12], исходная расчетная схема разбивается на отдельные и независимые подсистемы (блоки элементов или ветвей), для каждой из которых производится расчет показателей надежностей по некоторому заданному алгоритму. Затем по простым зависимостям, учитывающим последовательное соединение подсистем, определяются показатели надежности ИТ в целом. Такой подход применим в рамках объекта, элементы которого связаны технологически единым процессом, и значительно упрощает расчет, хотя предположение о независимости подсистем вносит некоторую погрешность.

Более приближенное к реальным условиям описание функционирования ИТ дает применение аппарата полумарковских процессов (ПМП). Основное отличие этих подходов от марковских моделей состоит в формулировке отказа, который для моделей полумарковских процессов понимается как событие, заключающееся не только в снижении количества отпускаемой ИТ тепловой энергии ниже требуемого (до этого необходимо в общих положениях пояснить о критерии отказа), но и в пребывании его в таком состоянии в течение времени, превышающего временной резерв, связанный в основном с тепловой аккумуляцией, запасами горячей воды и другими факторами. Авторы подхода [19] называют его функционально-технологическим.

Он предполагает совместное применение вероятностных моделей процессов функционирования и детерминированных моделей теплофизических процессов, что значительно усложняет расчет, но позволяет получить более точную характеристику надежности.

Альтернативными методами оценки надежности ИТ основаны на использовании имитационных расчетных алгоритмов. Ряд подобных методов предложены в [20, 21]. Метод статистических испытаний, широко используемый в исследовании надежности систем энергетики, позволяет получить достаточно объективные оценки надежности функционирования ИТ, однако он требует большого объема предварительных работ по сбору и подготовке исходной информации.

Методы анализа надежности ТС. Существующие ТС крупных централизованных систем теплоснабжения представляют собой сложные многокольцевые структуры, распределенные по территории города, включающие элементы линейной части (трубопроводы), элементы управления (НС, запорная и регулирующая арматура), потребительские установки. Методы анализа их надежности основаны на сочетании гидравлического моделирования и вероятностных методов описания их состояний во времени.

Общие принципы расчета надежности и резервирования тепловых сетей были сформулированы в 1972 г. в работе [1]. Этот подход, базирующийся на оценке узловых показателей, развивался затем в работах Сибирского энергетического института (СЭИ) – сейчас Институт систем энергетики (ИСЭМ) СО РАН [2, 22, 23]. Общие положения предложенного подхода исходят из сетевой специфики и особенностей функционирования ТС. Методика оценки надежности ТС основывается на сочетании детерминированных методов анализа нормальных и аварийных режимов в ТС с вероятностными методами вычисления показателей надежности теплоснабжения каждого потребителя (узловых ПН).

Ключевые положения данного метода исходят из специфических свойств ТС, связанными с их назначением, структурой и условиями функ-

ционирования. Отказ функционирования формулируется с точки зрения потребителя как снижение температуры воздуха в здании ниже граничного (минимально допустимого) для данного уровня надежности значения. В соответствии с несколькими уровнями надежности рассматриваются также несколько уровней теплоснабжения потребителей: расчетный и пониженный (резервный). Пониженный уровень должен поддерживаться в системе во время ликвидации аварийных ситуаций, возникших в резервируемой части сети. Он характеризуется подачей потребителям пониженного по сравнению с расчетным количества теплоты, называемого нормой аварийной подачи. Задача анализа надежности ТС состоит в определении узловых показателей, оценивающих рассматриваемые уровни надежности теплоснабжения потребителей по характеристикам надежности элементов сети.

Наряду с узловым подходом к оценке надежности ТС, существует другая концепция, основанная на использовании интегрального показателя надежности [24, 25]. Содержательная сущность интегрального показателя отражает степень выполнения задачи системой в течение отопительного периода, однако никак не интерпретируется с точки зрения качества выполнения функций ТС по отношению к потребителям. Это обстоятельство не позволяет использовать результаты анализа надежности для выделения слабых (узких) мест и наиболее тяжелых аварий, приводящих к длительному отключению потребителей. Следовательно, только узловая оценка может служить основной для принятия дальнейших решений по реконструкции ТС с целью повышения надежности теплоснабжения потребителей. Методы и модели, предлагаемые далее в настоящей работе, также используют узловой подход к оценке надежности теплоснабжения потребителей.

Методика комплексного анализа надежности теплоснабжения потребителей

1. Технологическая структура ТСК и исходные положения к анализу его надежности

В сложившейся проектной и эксплуатационной практике оценка надежности ТСС осуществляется раздельно на уровне ИТ и ТС. Такое разделение задач приводит к получению обособленных результатов, не учитывающих системные свойства ТСС и их взаимовлияние на надежность теплоснабжения, которые наиболее значимо проявляются при работе нескольких источников на единые тепловые сети. Оценки дефицитов топлива, полученные при анализе надежности СТС, также не позволяют количественно оценить влияние нарушений топливоснабжения на итоговую надежность теплоснабжения. В данной работе предлагаются подходы, позволяющие учесть связи между этими системами с точки зрения их совместного влияния на надежность обеспечения потребителей тепловой энергией в течение расчетного периода.

Технологическая связанность, непрерывность и взаимовлияние внутренних и внешних факторов процессов добычи, поставки топлива, производства, транспорта и потребления тепловой энергии обуславливают необходимость комплексного рассмотрения надежности их функционирования. Это предполагает совместное моделирование, расчет и анализ всей технологической цепочки производства и распределения тепловой энергии. Агрегирование соответствующих технологических процессов позволяет представить их в виде двух относительно самостоятельных систем: СТС, состоящей из подсистем добычи (СДТ) и транспортировки топлива (СТТ), и ТСС, включающей ИТ и ТС. В совокупности они формируют теплоснабжающий комплекс (ТСК), укрупненная структура которого представлена на рис.1.

Основная задача комплексного анализа надежности теплоснабжения потребителей ТСК (или надежности ТСК) состоит в получении количественной оценки интегрального воздействия всех элементов ТСК на надежность теплоснабжения потребителей и выделении степени влияния каждого из них. Ее решение представляет методически сложную проблему, связанную с получением обоснованных оценок надежности и поддержкой многих вспомогательных задач анализа надежности, обеспечивающих накопление и обработ-

ку информации о свойствах элементов системы, оценку воздействия внешних факторов на показатели функционирования подсистем ТСК, разработку нормативных требований и другие. Многие из этих частных вопросов составляют предмет отдельных исследований, но излагаемые далее методические положения являются основой для их постановки и решения.

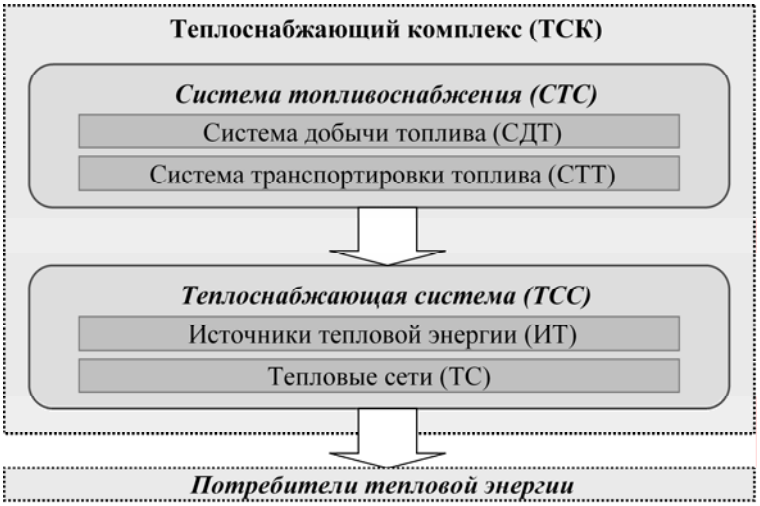


Рис. 1. Технологическая структура ТСК

Сетевая специфика теплоснабжающих систем и необходимость индивидуального учета каждого потребителя обуславливают применение следующих основных принципов расчета надежности теплоснабжения [1, 2]:

• оценка надежности узловыми показателями надежности (ПН);

• рассмотрение двух или более уровней обеспечения надежности теплоснабжения – нормальный (расчетный) и аварийные (пониженные);

• сочетание вероятностных методов расчета надежности в сочетании с детерминированными методами анализа режимов функционирования подсистем ТСК, в частности, аварийных теплогидравлических режимов в ТС [22, 23].

Транспортные схемы поставок топлива и распределения тепловой энергии в теплоснабжающих системах обладают сетевой структурой, что позволяет применять для решения задач надежности ТСК методы теории графов и «избыточных» проектных схем [23], обеспечивающие представление техно-

логической цепочки комплекса в виде единого сетевого графа, учитывающего индивидуальные особенности всех его элементов.

2. Алгоритм оценки надежности ТСК

Общий алгоритм комплексного анализа надежности ТСК укрупненно показан на рис.2. Его вычислительная схема представлена тремя уровнями – на уровне СТС, ТСС и всего ТСК в целом. Все множество рассматриваемых задач может быть поделено на два подмножества – вероятностное моделирование функционирования комплекса и анализ (расчет) режимов функционирования его подсистем.

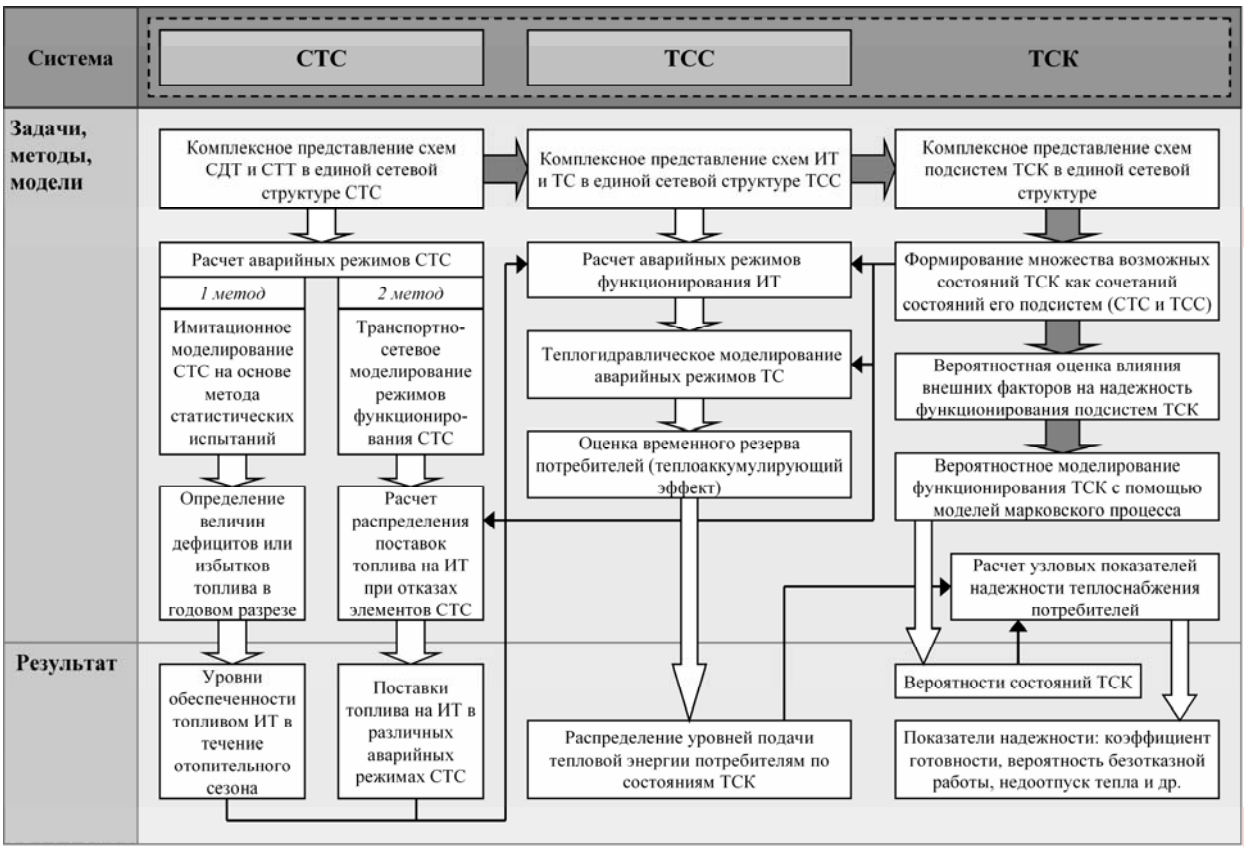


Рис. 2. Общий алгоритм оценки надежности ТСК

Топологической основой для решения первого подмножества задач является функциональная расчетная схема ТСК, сформированная путем объединения схем всех его подсистем, которые приводятся к единообразному отображению структуры и связей элементов. На основе анализа технологически возможных событий в подсистемах комплекса формируется множество состояний ТСК как сочетание состояний его подсистем. Данное множество

может также учитывать комплекс факторов внесистемного происхождения (внешних возмущений), однако, это требует предварительного исследования и статистического анализа их влияния на показатели функционирования рассматриваемых объектов. Оценка вероятностей состояний ТСК производится с учетом заданных характеристик надежности элементов (интенсивностей отказов и восстановлений) на основе марковской модели.

Методы анализа аварийных режимов функционирования подсистем ТСК учитывают их технологические особенности и различаются математическими моделями. Оценка надежности функционирования топливоснабжения может быть проведена с помощью одного из двух предложенных методов – структурного моделирования транспортных потоков топлива и метода статистических испытаний. Применение последнего метода не требует элементной схемы СТС, поэтому комплексное представление и построение марковской модели ТСК может осуществляться в рамках объединения ИТ и ТС. Уровни обеспеченности топливом ИТ (дефицит и избытки), полученные на этапе моделирования системы топливоснабжения, определяют их производительности на расчетный период (отопительный сезон). Эти показатели имеют как самостоятельное значение для анализа надежности системы топливоснабжения, так и используются при моделировании потокораспределения в ТСС с учетом возможных отказов оборудования ИТ и ТС. По их результатам оцениваются уровни аварийной подачи тепловой энергии потребителям во всех состояниях ТСК, учитывающих нарушения работы всех его подсистем.

Вероятности состояний ТСК и уровни подачи тепловой энергии потребителям в этих состояниях используются при вычислении узловых ПН, которые характеризуют надежность функционирования всего комплекса (см. рис.1) относительно каждого потребителя. При определении этих показателей также учитывается временная избыточность в системе, обусловленная теплоаккумулирующими свойствами зданий.

3. Имитационное моделирование СТС

Особенности исследования надежности функционирования СТС. Надежность функционирования системы топливоснабжения определяется соотношениями уровней потребностей и поставок топлива в течение рассматриваемого периода времени (раздел 2). Несогласованность этих уровней (с учетом резервного запаса) определяет дефицит или избыток (накопленный запас) топлива. Отмеченная ранее существенная сложность получения показателей надежности элементов СТС обуславливает использование для количественной оценки обеспеченности поставок топлива имитационного алгоритма, основанного на применении метода статистических испытаний.

Уровни поставок топлива на ИТ зависят от множества факторов как внутрисистемного, так и внешнего происхождения: аварийные ситуации в местах добычи и на линиях транспортировки топлива, обусловленные как отказами оборудования, так и воздействием окружающей среды, экономические риски, складывающиеся во взаимоотношениях поставщиков и потребителей топлива, и др.

Требуемое количество топлива, потребляемого на ИТ, также определяется множеством факторов, среди которых основным является изменение температуры наружного воздуха. Нерасчетное снижение ее среднегодовой величины приводит к увеличению тепловых нагрузок и теплопотребления, а соответственно к росту необходимых объемов топлива, поставляемого на ИТ. Компенсация дефицитов топлива, возникающего из-за нерасчетного похолодания, осуществляется из годовых и текущих запасов, рассчитываемых из условий ряда показаний предыдущих лет по среднему значению температуры наружного воздуха. Этот запас ограничивается с одной стороны, производительностью топливодобывающих предприятий и располагаемыми объемами топливных хранилищ. В связи с этим возможно такое снижение наружной температуры или увеличение продолжительности ее стояния, при которых потребность в топливе может превысить текущие объемы добычи и запасы. Оценка вероятности наступления такого события должна основываться на

статистических данных многолетних наблюдений, их анализе, обработке и аналитическом представлении для учета в расчетах.

Моделирование топливоснабжения для оценки надежности СТС. Построение математических моделей, описывающих условия поставки и потребления топлива на ИТ и учитывающих влияние на них перечисленных и других факторов, представляется достаточно сложной задачей. Вместе с тем, статистическая обработка данных об изменениях поставок топлива и потребностей, обусловленных воздействием наиболее значимых возмущений, позволяет получить распределения случайных величин уровней поставки и потребности топлива для каждого ИТ в системе. Полученные распределения используются при имитационном моделировании функционирования СТС на основе метода статистических испытаний. Данный подход позволяет учесть специфические особенности системы при множестве влияющих на нее внешних возмущений и компенсировать недостаток исходной информации о параметрах функционирования ее элементов. Положительный опыт использования данного метода применительно к СТС подтверждается рядом исследований, выполненных по данной тематике. Например, в [10] представлена имитационная модель топливоснабжения, с помощью которой производится оценка надежности функционирования СТС в динамике годовых показателей.

Предлагаемая в данной работе модель оценки надежности СТС ориентирована на определение потребности и поставки топлива по заданным характерным временным интервалам отопительного сезона (или распределения дефицитов и избытков топлива по интервалам этого периода). Это обеспечивает возможность рассмотрения состояний ТСК при различных условиях топливоснабжения.

Уровни потребности и поставки топлива за некоторый интервал времени n могут быть представлены величинами a_n и b_n соответственно. Непрерывные или дискретные распределения этих величин задаются в качестве исходных данных (в случае дискретного распределения используются суммы):

$$\int_{a_n^{\min}}^{a_n^{\max}} f(a_n)da_n = 1\ ,\ \int_{b_n^{\min}}^{b_n^{\max}} f(b_n)db_n = 1\ ; \qquad (1)$$

$$f(a_n) = 0,\ f(b_n) = 0\ \text{, если}\ a_n \notin [a_n^{\min}, a_n^{\max}],\ b_n \notin [b_n^{\min}, b_n^{\max}]; \qquad (2)$$

где $f(a_n)$ и $f(b_n)$ – плотности вероятностей случайных величин a_n и b_n, $a_n^{\min}, a_n^{\max}$ и $b_n^{\min}, b_n^{\max}$ – диапазоны их изменений.

Расчетный временной период τ_0, как правило, принимается равным отопительному сезону. Он разбивается на N временных интервалов продолжительностью τ_n, $n = \overline{1, N}$. Общее число источников тепловой энергии, включенных в ТСК соответствует множеству I. Возможность замещения топлива между источниками, как на практике, так и при моделировании осуществляется путем компенсации дефицитов топлива на одних ИТ за счет избытков на других. Источники, работающие на замещающих друг друга видах топлива, объединяются в группы I_g, где g – индекс, обозначающий принадлежность к группе.

Алгоритм оценки надежности СТС. Алгоритм решения задачи оценки уровней обеспеченности топливом ИТ (анализа надежности СТС) схематично представлен на рис. 3 и состоит из трех основных этапов.

1. Формирование исходной информации. Для каждого интервала n определяются:

– средняя потребность в топливе i-го источника $\overline{a}_n^i$, $i = \overline{1, I}$, рассчитываемая по средней температуре наружного воздуха для n-го интервала в соответствии с производительностью источника;

– средняя величина поставки топлива i-му источнику $\overline{b}_n^i$, $i = \overline{1, I}$, принимаемая на основе ретроспективных данных о работе ИТ в течение длительного времени;

– распределения случайных величин потребности и поставок топлива a_n^i и b_n^i, а также диапазоны их изменения $a_n^i \in [a_n^{i\min}, a_n^{i\max}]$, $b_n^i \in [b_n^{i\min}, b_n^{i\max}]$, удовлетворяющие условиям (1) и (2);

– для каждого i-го источника в интервале $n = 1$ задается начальный (сезонный) запас топлива (c_1^i), для более точного отражения реальных условий можно учитывать и величины текущих запасов ($c_{(\text{тек})n}^i$) как нормированный избыток к средней величине поставок топлива на ИТ;

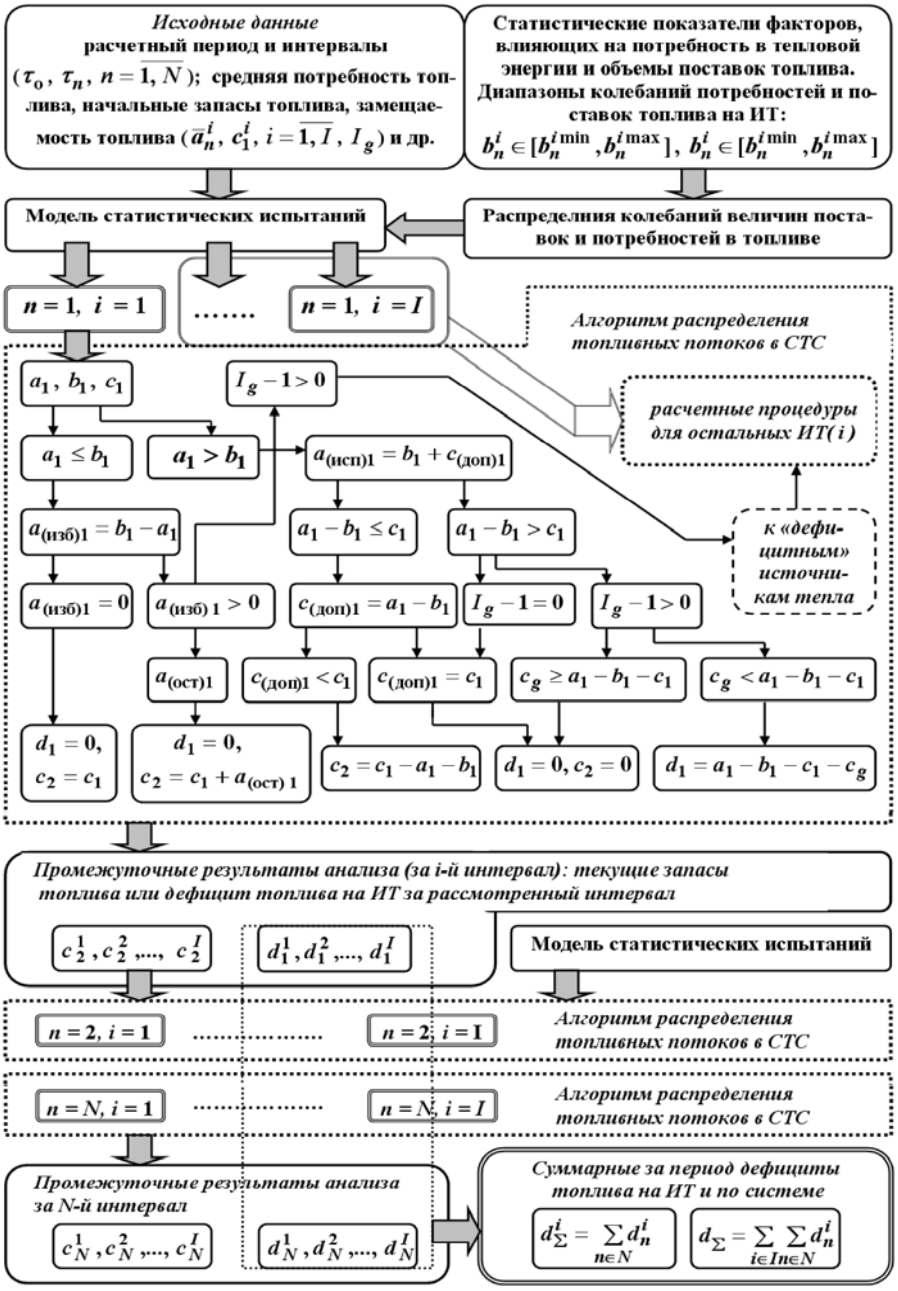


Рис. 3. Алгоритм оценки надежности СТС

2. Определение потребности и поставок топлива на ИТ в каждый интервал n расчетного периода. С помощью специальной модели статистиче-

ских испытаний (генератора случайных величин) для каждого i – го ИТ в интервале n получим уровни потребностей и поставок топлива – a_n^i и b_n^i.

3. Распределение топливных потоков в СТС. В пределах заданного временного интервала ($n = \overline{1, N}$) и для каждого i – го источника реализуется последовательность расчетных процедур, которая заключается в управлении топливными резервами в системе с учетом возможного замещения топлива между ИТ.

Для каждого принятого временного интервала n осуществляется следующая последовательность логических процедур и расчетных вычислений.

Проверяется соотношение величин потребности и поставок топлива: возможны два случая – либо $a_n \le b_n$ (а), либо $a_n > b_n$ (б).

В первом случае (а) количество использованного топлива равно его потребности $a_{(\text{исп})n} = a_n$, а избыток топлива, образующийся на ИТ, равен $a_{(\text{изб})n} = b_n - a_n$ и направляется на «дефицитные» по топливу источники $i \in I_g$. При неполном его использовании образуется остаток $a_{(\text{ост})n} = a_{(\text{изб})n} - \sum_{i \in I_g}(a_n^i - b_n^i - c_n^i)$, который переходит в запас топлива на следующий временной интервал (c_{n+1}). Во втором случае (б) используемое количество топлива складывается из поставленного объема и дополнительного топлива из запасов $a_{(\text{исп})n} = b_n + c_{(\text{доп})n}$. Если $a_n - b_n \le c_n$, то $c_{(\text{доп})n} = a_n - b_n$. При полном использовании запасов $c_{n+1} = 0$. В случае, когда $a_n - b_n > c_n$, то возможны два варианта использования замещающего топлива в системе, которое определяется по выражению $c_g = \sum_{i \in I_g}(a_n^i - b_n^i - c_n^i)$. При условии $c_g \ge a_n - b_n - c_n$ используемые резервы замещающего топлива покрывают дефицит полностью ($c_{(\text{доп})n} \le c_g$); если $c_g < a_n - b_n - c_n$, то $c_{(\text{доп})n} = c_g$ и на ИТ образуется топливный дефицит $d_n = a_n - a_{(\text{исп})n} = a_n - b_n - c_n - c_g$. В слу-

чае, когда группа источников I_g включает только рассматриваемый ИТ, то $d_n = a_n - b_n - c_n$.

Приведенные вычисления производятся последовательно для каждого i-го источника. В результате получаем величины дефицита и запасов топлива на последующий интервал $n+1$, для которого производится рассмотренный выше расчет по алгоритму 3 этапа на основе полученных данных имитационного моделирования (2 этап). Последний шаг соответствует интервалу $n = N$.

В результате формируется N пар величин (дефицита и запасов) для каждого ИТ. Средняя суммарная оценка недоотпуска тепловой энергии может производиться на основе значений годового дефицита (математического ожидания дефицита) для каждого ИТ и по системе в целом:

$$d_{\Sigma}^{i} = \sum_{n \in N} d_n^i , \tag{3}$$

$$d_{\Sigma} = \sum_{i \in I} \sum_{n \in N} d_n^i . \tag{4}$$

В случае нерасчетного снабжения топливом ИТ, полученные значения дефицитов топлива используются для определения недоотпуска тепла потребителям от нарушений работы СТС (см. рис.3).

4. Вероятностное моделирование функционирования ТСС

Функционирование ТСС с точки зрения вероятностного описания характеризуется последовательностью событий отказов и восстановлений, происходящих с определенной частотой на всех технологических этапах производства и распределения тепловой энергии. Для описания этой последовательности применяется модель марковского случайного процесса, которая является одним из наиболее обоснованных и универсальных аппаратов для оценки надежности восстанавливаемых систем [12].

Множество состояний ТСС формируется на основе представления возможных сочетаний состояний ИТ и ТС в единой структуре событий [26]. В зависимости от целей и ожидаемых результатов расчетов применяются ста-

ционарная или нестационарная модель функционирования ТСС. Предполагается, что надежность элементов системы не изменяется в течение расчетного периода (отопительного сезона), так как учет старения и износа оборудования имеет определяющее значение в более длительном периоде времени и учитывается на уровне формирования модели данных. Стационарная марковская модель функционирования ТСС представляется следующей системой линейных алгебраических уравнений (уравнения Колмогорова для стационарного режима):

$$p_s \sum_{z \in E_1} \nu_{sz} = \sum_{z \in E_2} p_z \nu_{zs} \, , \; s \in E \, , \tag{5}$$

$$\sum_{s \in E} p_s = 1 \; ; \tag{6}$$

где s и z – номера состояний ТСС; p_s и p_z – вероятности состояний ТСС; E – все возможные состояния системы; E_1 – подмножество состояний системы, в которые она может непосредственно перейти из состояния s; E_2 – подмножество состояний системы, из которых она может непосредственно перейти в состояние s; ν_{sz} – интенсивности переходов системы из состояния s в z, где $z \in E_1$; ν_{zs} – интенсивности переходов в состояние s из состояний z, где $z \in E_2$.

Условие стационарности характеристик надежности элементов и вероятностей состояний ТСК является достаточно адекватным реальным свойствам рассматриваемых подсистем. Существенные изменения последствий аварийных ситуаций и их вероятностей, обусловленные неравномерностью климатических условий, учитываются при расчете показателей надежности с использованием графика тепловых нагрузок в течение расчетного периода. В то же время существуют задачи анализа и синтеза надежности, требующие учета изменения за определенный период времени вероятностей состояний комплекса. Для их решения необходимо использовать марковскую модель, описывающую функционирование ТСК с переменными вероятностями состояний на основе уравнений Колмогорова для нестационарного режима:

$$p_s(t)\sum_{z\in E_1}\nu_{sz} = p'_s(t) + \sum_{z\in E_2} p_z(t)\nu_{zs}\ ,\ s\in E\,, \tag{7}$$

$$p_0(0) = p_0\,; \tag{8}$$

где $p_s(t)$ и $p_z(t)$ – функции вероятностей состояний от времени, p_0 – вероятность состояния $s = 0$ в момент времени $t = 0$, задаваемая как начальное условие для решения системы дифференциальных уравнений (7).

Соответствие результатов стационарного и нестационарного моделирования достигается в момент времени выхода системы в режим функционирования с постоянными характеристиками надежности элементов и других факторов, влияющих на надежность подсистем. С учетом этого, в расчетах используется в основном стационарная модель.

5. Моделирование послеаварийных режимов ТСК

Моделирование послеаварийных режимов ТСК заключается в определении уровня подачи тепловой энергии потребителям в каждом из возможных ее состояний. Этот процесс осуществляется последовательно для СТС и затем для ТСС с соответствующей взаимоувязкой результатов (см. рис.2).

Рассчитанные на этапе имитационного моделирования СТС уровни обеспеченности топливом ИТ позволяют определить их производительности при различных состояниях системы. Определение уровня подачи теплоносителя каждому j-му потребителю в различных состояниях ТСС s осуществляется с помощью модели расчета гидравлических режимов ТС (модели потокораспределения), реализующей методы теории гидравлических цепей (ТГЦ).

Узловая модель потокораспределения (гидравлических режимов) в ТС, представленная в матричной форме, имеет вид [22, 23]:

$$Ax = g\,, \tag{9}$$

$$\overline{A}^{\mathrm{T}} p = h - H\,, \tag{10}$$

$$SXx = h\,, \tag{11}$$

где A – матрица соединений для линейно независимых узлов, x – вектор расходов теплоносителя на участках сети, т/ч; g – вектор расходов теплоносителя в узлах расчетной схемы, т/ч; $\overline{A}^{\mathrm{T}}$ – транспонированная матрица $\overline{A}$ ($\overline{A}$ – полная матрица соединений узлов и ветвей схемы сети); p – полный вектор давлений в узлах сети, м в.ст.; h, H – векторы потерь напора и действующих напоров на ветвях, м в.ст.; S, X – диагональные матрицы гидравлических сопротивлений ветвей, составленные из величин s_i – гидравлических сопротивлений ветвей, (м·ч^2/т^2), и абсолютных значений расходов на них $|x_i|$, (т/ч).

В результате расчета всей системы (30)-(32) для каждого из состояний ТСК (s) определяются величины g_{sj} – расходы теплоносителя в каждом узле потребления j, с помощью которых рассчитываются уровни подачи тепловой энергии q_{sj}, Гкал/ч, по следующей зависимости [27]:

$$q_{sj} = g_{sj} c (t_{\text{под}} - t_{\text{обр}}), \qquad (12)$$

где c – теплоемкость теплоносителя, Гкал/(т·ºС); $t_{\text{под}}$ и $t_{\text{обр}}$ – температуры сетевой воды в подающем и обратном трубопроводах соответственно.

6. Определение узловых показателей надежности

Результаты вероятностного моделирования функционирования ТСС (вероятности состояний системы) и рассчитанные на предыдущем этапе уровни подачи тепла потребителям используются при осуществлении итоговой количественной оценки надежности ТСК с помощью узловых ПН теплоснабжения каждого потребителя.

В соответствии с методическими положениями, изложенными выше, надежность расчетного (первого) уровня определяется коэффициентом готовности (K_j). Он соответствует периоду в отопительном сезоне, в течение которого у данного потребителя обеспечивается расчетное значение температуры внутреннего воздуха. Надежность пониженного (второго) уровня оценивается показателем вероятности безотказной работы (R_j). Он представляет

собой вероятность того, что в течение отопительного периода внутренняя температура воздуха в зданиях не опустится ниже ее некоторого граничного значения $t_{j\min}$.

Коэффициент готовности и вероятность безотказной работы определяются для каждого потребителя j с учетом их временного резерва, обусловленного теплоаккумулирующей способностью зданий, по следующим формулам:

$$K_j = 1 - \sum_{s \in E} p_s \left[\frac{1}{1-\omega} \left(1 - \frac{1}{q_{\text{op}j}} \left(\varphi_j t_{sj} - \varphi_j \left(\frac{C_1 - C_2 \exp B}{C_3 (1 - \exp B)} \right) - q_{\text{гв}j} \right) \right) \right]^{\alpha}, \quad (13)$$

$$R_j = \exp \left[- \sum_{z \in E} \sum_{s \in E_1} \nu_{zs} \dot{p}_0 \tau_{\text{o}} \left(\frac{1}{1-\omega} \left(1 - \frac{1}{q_{\text{op}j}} \left(\varphi_j t_{sj} - \varphi_j \left(\frac{C_1 - C_2 \exp B}{C_3 (1 - \exp B)} \right) - q_{\text{гв}j} \right) \right) \right)^{\alpha} \right], \quad (14)$$

$$\omega = q_{\text{он}j} / q_{\text{op}j},\ \delta = q_{\text{oc}j} / q_{\text{op}j},\ \alpha = (1-\delta)/(\delta - \omega), \quad (15)$$

$$C_1 = t_{0j}(1 - \overline{q}_{sj}),\ C_2 = t_{j\min} - t_{0j}\overline{q}_{sj},\ C_3 = 1 - \overline{q}_{sj},\ B = \beta^{-1} (\nu_{szj}^{\max})^{-1}, \quad (16)$$

где ω – коэффициент неравномерности графика тепловой нагрузки; $q_{\text{op}j}$, $q_{\text{он}j}$, $q_{\text{oc}j}$ – нагрузки отопления: расчетная, соответствующая началу отопительного периода и средняя за отопительный период; $q_{\text{гв}j}$ – расчетная тепловая нагрузка горячего водоснабжения; $\overline{q}_{sj} = q_{sj} / q_{0j}$ – относительное снижение подачи тепловой энергии в состоянии s (здесь q_{0j} – суммарная расчетная тепловая нагрузка потребителя); φ_j – постоянный коэффициент, зависящий от теплофизических свойств здания потребителя j [27]; $\dot{p}_0$ – вероятность полностью работоспособного состояния системы при условии поглощающих отказовых состояний; τ_0 – продолжительность отопительного периода; t_{0j} – расчетная внутренняя температура воздуха; $\nu_{szj}^{\max}$ – максимальная из возможных для данного состояния s интенсивность перехода в работоспособное для данного потребителя состояние z.

7. Декомпозиционный анализ надежности подсистем ТСК

Анализ надежности ТСК, осуществляемый с целью дальнейшего принятия эффективного решения по его резервированию, обуславливает необходимость получения информации о степени влияния подсистем комплекса или их отдельных элементов на надежность теплоснабжения потребителей. Распределение показателей надежности по подсистемам и их элементам может быть осуществлено с помощью методического подхода, основанного на использовании принципа декомпозиционного (поэлементного) анализа состояний функционирования ТСК. Его идея заключается в выделении в расчетной схеме ТСК анализируемой на надежность подсистемы, а оценка надежности осуществляется при условии, что никакие отказы элементов других подсистем не снижают уровень подачи тепловой энергии. Вероятностная оценка системы производится при исходных параметрах надежности элементов. Рассчитанные таким образом ПН учитывают последствия отказов только рассматриваемой подсистемы с учетом состава событий, объединяющей элементы всех подсистем.

Оценка степени влияния какого-либо элемента схемы ТСК на надежность теплоснабжения потребителя осуществляется при условии, что отказ данного элемента считается невозможным (элемент с абсолютной надежностью), при этом его производительность (пропускная способность) остается на расчетном уровне. Это условие принимается на этапе вероятностного моделирования функционирования ТСК путем обнуления интенсивностей переходов, связанных с отказом рассматриваемого элемента. Отношение величины ПН, полученной при оценке надежности всей системы, к его значению, полученному относительно рассматриваемого элемента, показывает, в какой мере этот элемент оказывает влияние на надежность теплоснабжения потребителя. Чем меньше данное отношение, тем значительнее рассматриваемый элемент снижает надежность теплоснабжения.

Расчет укрупненной схемы ТСК

1. Подготовка исходных данных. Предложенные методические подходы достаточно универсальны и могут применяться для расчетов надежности различных схем ТСК. На рис.4-а представлена исходная схема теплоснабжающего комплекса, которая включает СТС, два источника тепловой энергии (ИТ1, ИТ2), группу потребителей (узлы 1-75) и кольцевую ТС, состоящую из 142 участков. СТС на схеме представлена условно, транспортные связи по топливоснабжению ИТ1 и ИТ2 на рис.4-а показаны пунктиром.

Эквивалентирование представленной на рис.4-а схемы позволяет ее упростить до вида, показанного на рис.4-б. Полученная таким образом схема ТСС состоит всего из 18 участков тепловой сети и 7 обобщенных потребителей, которые объединяют группы потребителей, как показано на рис.4-а. Принципиальная элементная схема ИТ показана на рис.4-в. Она состоит из обобщенных основных элементов – котла 1, турбины 2, сетевого подогревателя 3 и насоса 4, последние два элемента продублированы (5 и 6) и представляют технологический резерв на источнике.

В соответствии с предлагаемым методическим подходом к исследованию надежности теплоснабжения [26] преобразованная расчетная схема ТСС, приведенная на рис.4-г, представляется в виде единой структуры, связывающей элементы схем ТС и ИТ. Элементы 1-18 являются участками сети, элементы 19-30 соответствуют источникам.

Исходные данные, представленные в табл.1, включают технические параметры ТС (диаметры и длины трубопроводов), нагрузки потребителей, мощности ИТ и надежностные характеристики схемы, включая интенсивности отказов и восстановлений элементов ТС и ИТ. Интенсивности восстановления элементов рассчитаны в соответствии с приведенными в [2] зависимостями интенсивности восстановления трубопроводов от их диаметра, интенсивности отказов участков сети приняты на уровне 0,00002 1/(км·ч). Аналогичные характеристики для элементов ИТ приняты по данным анализа эксплуатационной надежности энергоблоков ТЭЦ.

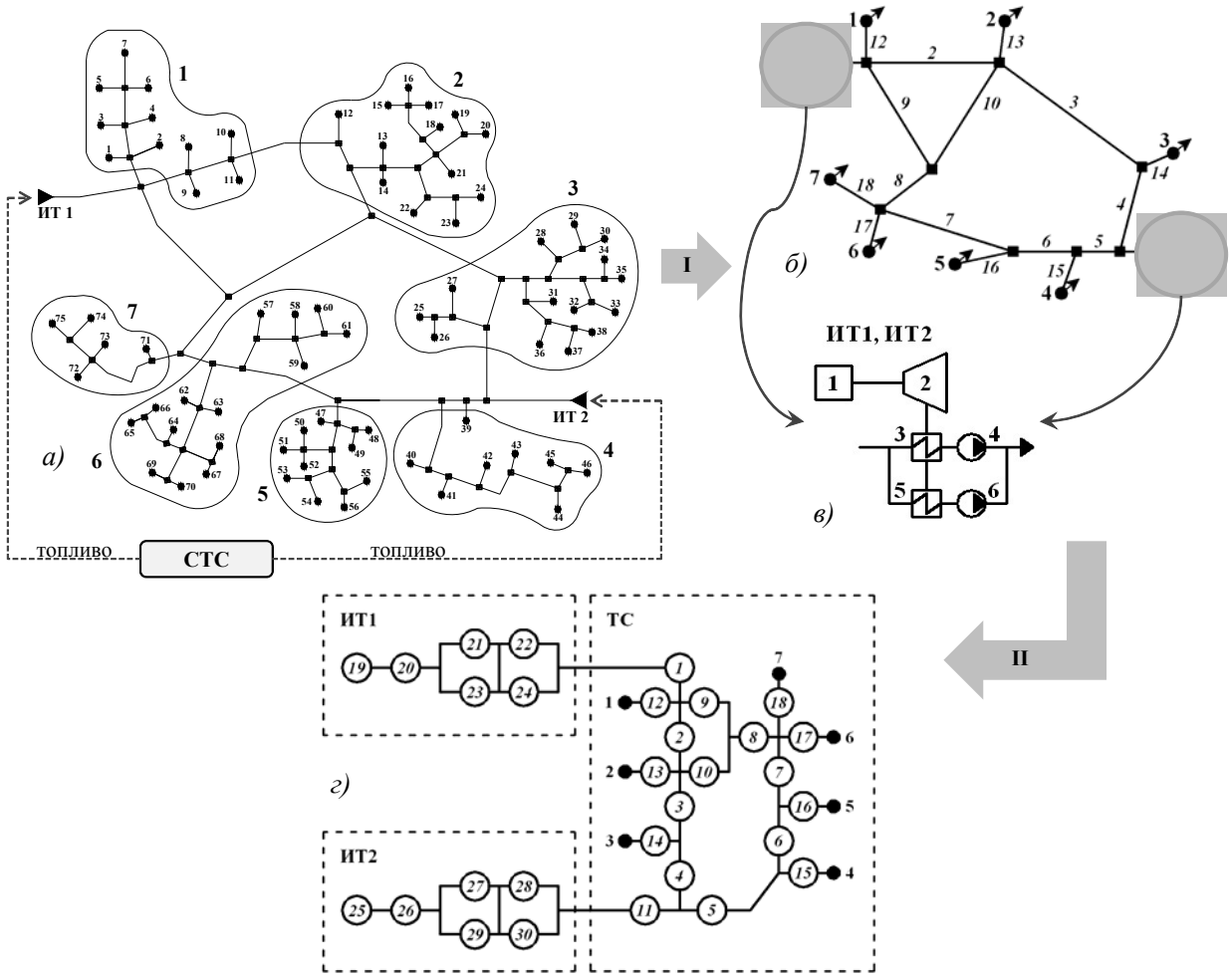


Рис. 4. Этапы построения расчетной схемы ТСК:
I – эквивалентирование исходной схемы; II – формирование расчетной схемы ТСС для анализа надежности ТСК; а) исходная схема ТСС; б) эквивалентированная схема ТСС; в) упрощенная схема ИТ; г) расчетная схема ТСС

Информация, необходимая для расчета надежности функционирования СТС, представлена в табл.2. Она содержит средние за расчетный интервал отопительного сезона величины потребностей и поставок топлива для каждого ИТ, принятые по показателям работы ТЭЦ Иркутской области. Отопительный сезон разделен на восемь расчетных интервалов, равных одному месяцу. В табл.2 они указаны в числовом формате, начиная с 10-го месяца (октябрь).

Таблица 1

Исходные данные для анализа надежности ТСК

Параметры ТС и надежностные характеристики элементов ТС					Нагрузки потребителей и мощность ИТ			
ТС					Потребитель	Нагрузка, Гкал/ч	ИТ	Мощность, Гкал/ч
Участок ТС (рис. 1-а, в)	Диаметр, м	Длина, м	Интенсивность восстановления, 1/ч	Интенсивность отказов, 1/ч				
1	0,65	2449	0,0236	0,00004898	1	281	ИТ1	1257
2	0,6	5864	0,0246	0,00011728	2	344	ИТ2	1467
3	0,55	7811	0,0257	0,00015622	3	268	-	-
4	0,5	3867	0,0269	0,00007734	4	503	-	-
5	0,6	1909	0,0246	0,00003818	5	352	-	-
6	0,55	2799	0,0257	0,00005598	6	402	-	-
7	0,5	6147	0,0269	0,00012295	7	197	-	-
8	0,55	2887	0,0257	0,00005775	сумма	2346	сумма	2724
9	0,55	5540	0,0257	0,00011080	Характеристики надежности элементов ИТ1, ИТ2			
10	0,5	5580	0,0269	0,00011159				
11	0,7	3128	0,0227	0,00006256	Элемент ИТ1, ИТ2 (рис.1-в)		Интенсивность восстановления, 1/ч	Интенсивность отказов, 1/ч
12	0,35	1844	0,0312	0,00003688				
13	0,4	1860	0,0296	0,00003720				
14	0,35	1508	0,0312	0,00003016				
15	0,5	1962	0,0269	0,00003925	19, 25		0,0100	0,0001065
16	0,4	2620	0,0296	0,00005239	20, 26		0,0089	0,0000330
17	0,45	1932	0,0282	0,00003864	21, 23, 27, 29		0,0074	0,0000558
18	0,35	2575	0,0312	0,00005150	22, 24, 28, 30		0,0115	0,0000535

Таблица 2

Потребности и поставки топлива на ИТ, тыс. т у.т.

Источник		Месяц							
		10	11	12	01	02	03	04	05
ИТ1	потребности	31,0	37,7	47,5	45,8	38,8	38,1	27,9	24,4
	поставки	30,4	37,1	46,6	45,0	38,1	37,4	27,4	24,0
ИТ2	потребности	32,6	44,3	51,5	47,2	41,8	40,6	31,3	25,6
	поставки	32,0	43,4	50,5	46,3	40,9	39,8	30,7	25,1

Исходные климатические параметры (температура наружного воздуха, продолжительность отопительного сезона и др.) приняты для условий г. Иркутска, теплоаккумулирующая способность зданий соответствует величине $\beta = 60$ ч.

В соответствии с вышеописанной методикой ниже излагается последовательность выполнения расчетов.

2. Вероятностное моделирование функционирования ТСС. Вероятностное описание функционирования ТСС основано на следующих положениях. Каждый элемент может пребывать в двух состояниях – работоспособном и

отказовом. Поток событий в пределах одной подсистемы (ТС, ИТ1 и ИТ2) является простейшим. Данное условие предполагает одновременный отказ нескольких элементов из разных подсистем ТСС, здесь же для наглядности ограничимся рассмотрением состояний совместного отказа не более двух элементов. В соответствии с этим множество состояний формируется возможными состояниями ТС, ИТ1, ИТ2 и их сочетаниями: ТС+ИТ1, ТС+ИТ2 и ИТ1+ИТ2. Граф состояний, отражающий заданную структуру событий системы, в сокращенном виде изображен на рис.5. Номер элемента графа соответствует номеру отказавшего элемента, обозначенного на рис.4-г. Состояния одновременного отказа двух элементов обозначены их номерами, записанными через знак «+».

Марковский случайный процесс, удовлетворяющий заданным условиям, описывается системой из 283 уравнений вида (5)-(6), в результате решения которой получены значения вероятностей состояний системы. Ввиду большого массива информации эти результаты здесь не приводятся.

3. Расчет послеаварийных режимов. Определение уровней подачи тепловой энергии потребителям в различных состояниях ТСС осуществляется на основе многовариантных расчетов потокораспределения в ТС с помощью модели (9)-(11) и зависимости (12), при этом также учитываются отказы оборудования ИТ и недоотпуск тепла потребителям вследствие дефицитов топлива, рассчитанных в п. 1.

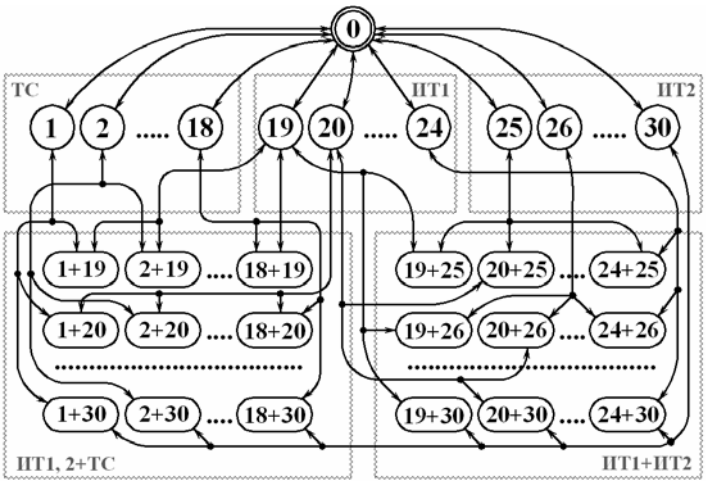


Рис. 5. Граф состояний ТСС

4. Расчет показателей надежности. Показатели надежности рассчитываются по зависимостям (13)-(16). Коэффициент готовности рассчитывается на расчетный уровень подачи тепла потребителям при $t_{j\min}^{(1)} = 20°C$, а вероятность безотказной работы – на пониженный, соответствующий $t_{j\min}^{(2)} = 16°C$.

Результаты комплексной оценки надежности ТСК, учитывающей воздействие всех его подсистем на надежность теплоснабжения потребителей, приведены в табл.3 (раздел ТСК).

Требования к надежности теплоснабжения потребителей, приведенные в [28], устанавливают следующие нормативные значения ПН: вероятность безотказной работы для ИТ – 0,97, для ТС – 0,9 и в целом для ТСС – 0,86, коэффициент готовности для ТСС – на уровне 0,97. Сопоставление этих значений и расчетных величин ПН, представленных в табл.3 для ТСК, показывает, что полученный коэффициент готовности не превышает 90% (0,8707) от его требуемого значения, а вероятность безотказной работы – 92% (0,7945). Это свидетельствует о неудовлетворительной надежности данной системы.

5. Декомпозиционный анализ надежности подсистем ТСК. Для адекватного сопоставления результатов оценки надежности с нормативными требованиями, дифференцированными по отдельным подсистемам, был проведен декомпозиционный анализ согласно принципам, изложенным выше. Результаты полученной таким образом «распределенной» оценки надежности представлены в табл.3. в разделах СТС, ТСС, ИТ и ТС. Наглядное представление о соотношении значений ПН, рассчитанных для различных подсистем, в сравнении с общими их уровнями для всего ТСК дают диаграммы, приведенные на рис. 6 – 8.

Интегральные оценки ПН имеют более низкие значения величин, полученных по результатам раздельных расчетов. Например, если коэффициент готовности в целом для ТСК не превышает 0,87 (рис.6), что составляет менее 90% от нормативного значения для потребителей, то при выделении доли

ТСС тот же показатель для потребителей 3, 4 практически достигает своей нормативной величины, равной 0,97. Нормативные уровни ПН на рисунках изображены пунктирными линиями. Аналогичное соотношение результатов расчетов справедливо и для показателя вероятности безотказной работы, максимальное значение которого для ТСК в целом соответствует потребителю 2 и равно 0,7945 (рис.7), что составляет 92% от норматива. В то же время его величина, рассчитанная только для ТСС, выше интегральной оценки и удовлетворяет нормативному значению для всех потребителей.

Таблица 3

Узловые показатели надежности теплоснабжения потребителей (*K* – коэффициент готовности, *R* – вероятность безотказной работы)

Показатель	Потребитель						
	1	2	3	4	5	6	7
ТСК							
K	**0,8612**	**0,8629**	**0,8707**	**0,8518**	**0,8565**	**0,8439**	**0,8660**
R	**0,7745**	**0,7945**	**0,7852**	**0,7688**	**0,7674**	**0,7678**	**0,7878**
СТС							
K	0,8890	0,8980	0,8983	0,8782	0,8890	0,8791	0,8980
R	0,8890	0,8980	0,8983	0,8782	0,8890	0,8791	0,8980
ТСС							
K	0,9687	0,9609	0,9693	0,9699	0,9634	0,9599	0,9644
R	0,8712	0,8848	0,8742	0,8754	0,8632	0,8734	0,8773
ИТ							
K	0,9897	0,9824	0,9904	0,9906	0,9841	0,9811	0,9850
R	0,9801	0,9795	0,9800	0,9804	0,9794	0,9797	0,9804
ТС							
K	0,9788	0,9782	0,9787	0,9791	0,9790	0,9784	0,9791
R	0,8889	0,9033	0,8920	0,8929	0,8814	0,8915	0,8948

Декомпозиция ТСС на составляющие ее подсистемы ИТ и ТС также представляет более детальный анализ надежности каждого из этих звеньев системы и обеспечивает получение более подготовленной базы для принятия решений по развитию и функционированию комплекса. На рис.8 показаны значения вероятности безотказной работы, рассчитанные для ТСС в целом и раздельно для ИТ и ТС. При выполнении нормативных требований в целом для ТСС, показатель надежности, рассчитанный только относительно тепловой сети, не соответствует нормативу кроме одного потребителя 2, в то же

время вероятность безотказной работы источников тепла превышает нормативное значение для всех потребителей.

Сравнительный анализ значений ПН, представленных в табл.3 и на рис.6-8, позволяет укрупненно сформулировать приоритетные направления повышения надежности ТСК. Так, значения коэффициента готовности принимают наименьшие значения для СТС (рис.6), а для ТС не выполняются нормативные требования по вероятности безотказной работы (рис.8).

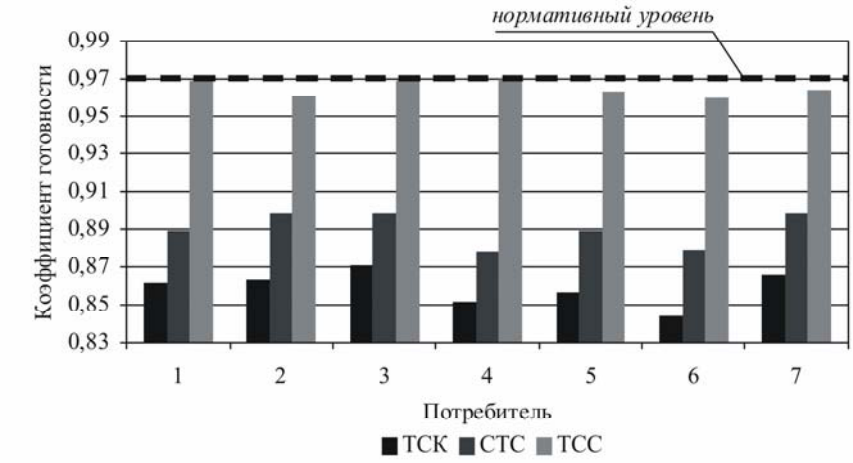


Рис. 6. Распределение коэффициента готовности по подсистемам ТСК

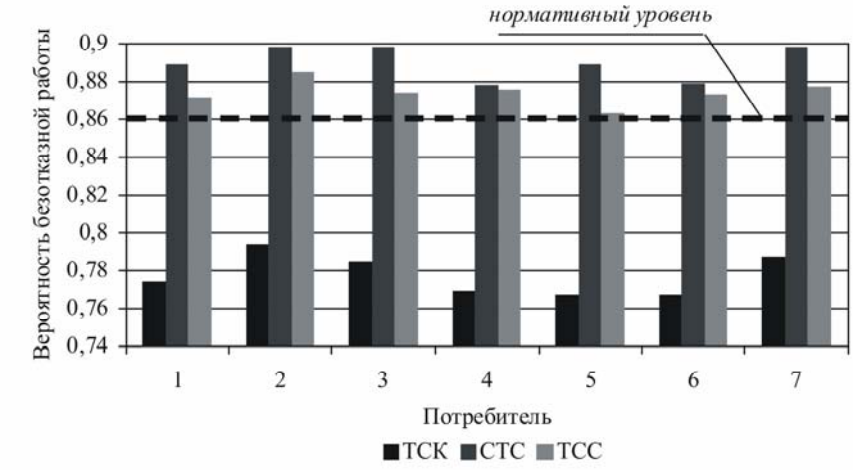


Рис. 7. Распределение вероятности безотказной работы по подсистемам ТСК

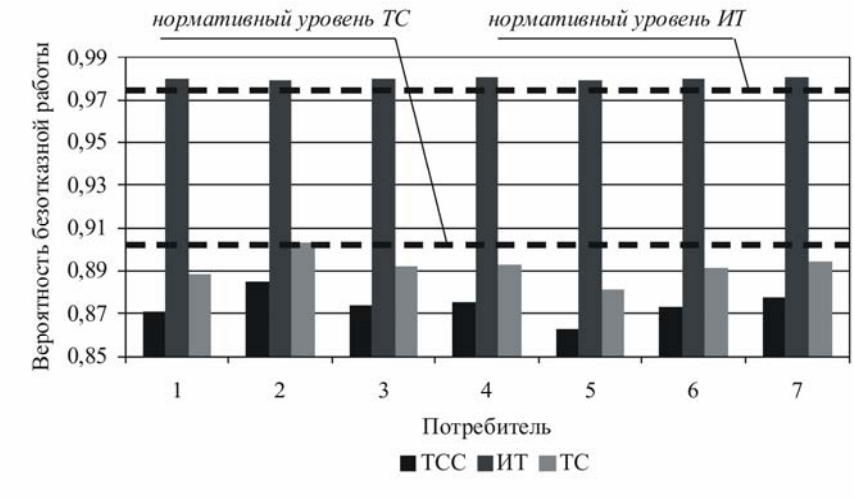


Рис. 8. Распределение вероятности
безотказной работы по подсистемам ТСС

Из этого следует, что увеличение запасов топлива, формирование структурного резерва в СТС в виде дополнительных источников топлива, более надежной системы его транспортировки позволит значительно повысить надежность расчетного теплоснабжения потребителей рассматриваемой системы. Реализация комплекса мероприятий по функциональному и структурному резервированию ТС обеспечит требуемый уровень пониженного теплоснабжения потребителей в аварийных ситуациях. Более детальный состав мероприятий по повышению надежности подсистем ТСК обоснованно определяется при решении в последующем задачи оптимального синтеза его надежности. Исходной основой решения данной задачи являются результаты исследований, полученных в процессе комплексного анализа надежности теплоснабжения потребителей.

Выводы

Комплексный подход (объектный и содержательный) к анализу надежности ТСК обеспечивает новый уровень исследований, увеличивает возможности резервирования подсистем, максимально используя взаимосвязи и взаимозаменяемость их элементов. Этот подход приобретает наибольшую актуальность при организации совместной работы ИТ на единые тепловые сети, хотя и при их раздельном функционировании позволяет получить наи-

более рациональные решения по надежности. Предлагаемые модели и методы количественной оценки надежности теплоснабжения потребителей значительно расширяют спектр решаемых задач и обладают следующими особенностями:

– представление всех этапов производства и распределения тепловой энергии в единой технологической структуре позволяет получить системную оценку надежности теплоснабжения потребителей;

– оценка состояний ТСК относительно конкретного потребителя характеризуется простой логикой событий (отказ или работоспособность комплекса), что обеспечивает возможность использования марковских моделей для вероятностного описания функционирования ТСК;

– применение методического подхода, основанного на принципах элементной декомпозиции расчетной схемы, для исследования влияния на уровень обеспечения теплоснабжения потребителей каждой подсистемы ТСК (группы элементов, отдельных элементов) позволяет на этапе анализа надежности сформулировать рациональные направления повышения надежности как для ТСК в целом, так и для составляющих его подсистем;

– предлагаемые модели и методы являются достаточно универсальными и применимы к ТСК любой структуры, сложности и масштаба.

Практические расчеты по оценке надежности ТСК показывают, что обеспечение требуемой надежности теплоснабжения потребителей возможно за счет различных подсистем (СТС, ИТ, ТС). Их степени влияния и возможные доли в повышении надежности во многом зависят от особенностей и условий функционирования комплекса в целом.

Список литературы

[1] Хасилев В.Я., Такайшвили М.К. Об основах методики расчета и резервирования тепловых сетей // Теплоэнергетика. – 1972. – №4 – С. 14 – 19.

[2] Надежность систем теплоснабжения / Справ., Отв. ред. Е. В. Сеннова. – Новосибирск: Наука, 2000. – 360 с.

[3] Надежность топливоснабжения электростанций: методы и модели исследований / Отв. ред. А.С. Некрасов, А.Ш. Резниковский. – М.: Наука, 1990. – 200 с.

[4] Некрасов А.С., Великанов М.А. Многолетнее регулирование расходов топлива на отопление и вентиляцию // Достижения и перспективы. Сер. Энергетика. – М.: МЦНТИ КСА при президиуме АН СССР, 1986. – вып. 46. – С. 85–98.

[5] Мазур Ю.Я. Проблемы маневренности в развитии энергетики. – М.: Наука, 1986.

[6] Зоркальцев В.И., Иванова Е.Н. Интенсивность и синхронность колебаний потребности в топливе на отопление по экономическим районам страны. – Известия АН СССР, энергетика и транспорт, 1990. – №6. – С. 14 – 22.

[7] Зоркальцев В.И., Иванова Е.Н. Анализ интенсивности и синхронности колебаний потребности в топливе на отопление. – Сыктывкар: Коми науч. Центр УрО АН СССР, 1989. – 24 с.

[8] Claudio M. Rocco S. A rule induction approach to improve Monte Carlo system reliability assessment / Reliability Engineering & System Safety, 2003, 82. – P. 85 – 92.

[9] Naess A., Leira B.J., Batsevych O. System reliability analysis by enhanced Monte Carlo simulation / Structural Safety, 2009, 31. – P. 349 – 355.

[10] Зоркальцев В.И., Колобов Ю.И. Имитационная модель для изучения надежности топливоснабжения теплогенерирующих установок// Труды Коми филиала Академии наук СССР. – 1984 – С. 33 – 39.

[11] Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д. Математические методы в теории надежности. – М.: Наука, 1965. – 524 с.

[12] Надежность технических систем: Справочник / Под ред. И.А. Ушакова. – М.: Радио и связь, 1985. – 608 с.

[13] Половко А.М. Основы теории надежности. – М.: Наука, 1964. – 446 с.

[14] Haghifam Mahmood Reza, Manbachi Moein. Reliability and availability modelling of combined heat and power (CHP) systems / International Journal of Electrical Power & Energy Systems, 2011, 33. – P. 385 – 393.

[15] Lisnianski A., Elmakias D., Laredo D., Hanoch Ben Haim. A multi-state Markov model for a short-term reliability analysis of a power generating unit / Reliability Engineering & System Safety, 2012, 98. – P. 1 – 6.

[16] Королюк В.С., Турбин А.Ф. Полумарковские процессы и их приложения. – Киев: Наук. Думка. – 1976. – 138 с.

[17] Rudenko Yo.N., Ushakov I.A., Cherkesov G.N. Theoretical and methodical aspects of reliability analysis of energy systems// Energy Reviews Scientific and Engeneering problems of Energy System Reliability. – N.Y.: Harwood Acad. Publishers. – 1984. – Vol. 3. – P. 82 – 94.

[18] Woo S. Jung, Nam Z. Cho. Semi-Markov reliability analysis of three test/repair policies for standby safety systems in a nuclear power plant / Reliability Engineering & System Safety, 1991, 31. – P. 1 – 30.

[19] Смирнов А.В. Функционально-технологический подход к надежности источников теплоты в системах теплоснабжения // Тепло- и энергосбережение, теплометрия. –Киев: Ин-т проблем энергосбережения АН УССР, 1990. – С. 50 – 57.

[20] Бусленко Н.П. Моделирование сложных систем. – М.: Наука, 1978. – 400с.

[21] Бусленко Н.П., Калашников В.В., Коваленко И.Н. Лекции по теории сложных систем. – М.: Сов. радио, 1983. – 440 с.

[22] Меренков А.П., Хасилев В.Я. Теория гидравлических цепей. М.: Наука, 1985. 278 с.

[23] Сеннова Е.В., Сидлер В.Г. Математическое моделирование и оптимизация развивающихся теплоснабжающих систем. - Новосибирск, 1985 г. – 222 с.

[24] Ионин А.А. Критерии для оценки и расчета надежности тепловых сетей //Водосн. и сан.техника. 1978, №12. С. 9 – 10.

[25] Ионин А.А. Надежность систем тепловых сетей. – М.: Стройиздат, 1989. – 265 с.

[26] Стенников В.А., Постников И.В. Комплексная оценка надежности теплоснабжения потребителей. – Трубопроводные системы энергетики. Развитие теории и методов математического моделирования и оптимизации – Новосибирск: Наука, 2008. – 312 с., раздел 2.6 – С. 139 – 157.

[27] Соколов В.Я. Теплофикация и тепловые сети. – М.: Издательство МЭИ, 1999. – 472 с.

[28] СНиП 41-02-2003 «Тепловые сети». М.: Государственный комитет Российской Федерации по строительству и жилищно-коммунальному комплексу, 2000.

УДК 519.8

ОПТИМАЛЬНЫЕ НОРМИРОВАННЫЕ ВЕКТОРЫ КОНЕЧНОГО ПРОДУКТА И ДОБАВЛЕННОЙ СТОИМОСТИ В ПРОДУКТИВНОЙ МОДЕЛИ ЛЕОНТЬЕВА

Стецюк П. И.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: stetsyukp@gmail.com

Аннотация. Рассматривается задача нелинейного программирования с билинейной целевой функцией, линейными ограничениями для прямой и двойственной моделей Леонтьева, двумя квадратичными равенствами, нормирующими векторы конечного продукта и добавленной стоимости. Если матрица Леонтьева (матрица коэффициентов прямых затрат) продуктивна, то оптимальное решение задачи выражается через собственные векторы, соответствующие максимальным собственным числам некоторых неотрицательных симметричных матриц. Если матрица Леонтьева продуктивна и неразложима, то задача имеет единственное решение, которое интерпретируется как оптимальные нормированные структуры конечного продукта и добавленной стоимости для экономики по Леонтьеву. Приведены расчеты для 15-отраслевого баланса Украины за 2003–2009 годы.

Ключевые слова. Матрица Леонтьева, статические модели Леонтьева, экстремальная квадратичная задача, собственные числа и собственные векторы, сингулярное число.

Введение

Леонтьевские модели "затраты-выпуск" и равновесных цен являются ключом к многим экономическим явлениям и политэкономическим величинам. Так, например, с их помощью профессор Хайнц Д. Курц (университет Граца, Австрия) исследует проблему добавленной стоимости, использование основного капитала и проблему технических изменений [1], профессор Генрих Бортис (университет Фрибурга, Швейцария) исследует связь между структурой труда и заработной платой [2]. В работе [1] подчеркивается, что хотя исследуемые в ней проблемы очень сложные, но перспективы обнадеживают и нет опасений, что аналитикам леонтьевских моделей скоро придется искать новые области исследований, потому что старые уже исчерпаны. Данная работа призвана в некоторой степени подтвердить этот тезис. В ней исследуются оптимальные соотношения между основными векторными величинами леонтьевской экономики, такими

как валовой и конечный продукт, цены, нормы добавленной стоимости. Величины, характеризующие оптимальные соотношения, тесно связаны с максимальным сингулярным[1] числом неотрицательных матриц.

Матрицы Леонтьева *A* и *B*

Рассматривается экономика с n чистыми отраслями, т.е. каждая отрасль производит один вид продукта и разные отрасли выпускают разные продукты. Пусть $A \geq 0$ – неотрицательная $n \times n$–матрица

$$A = \{a_{ij}\},\ i,\ j = \overline{1,n},$$

где коэффициент $a_{ij} \geq 0$ обозначает величину затрат продукта отрасли i на изготовление единицы продукта отрасли j. Величины a_{ij} могут быть заданы в натуральном или в стоимостном выражении. Матрица A называется матрицей Леонтьева (матрицей коэффициентов прямых затрат, матрицей технологических коэффициентов). Для экономики страны (региона) матрица A несет информацию о сложившейся структуре межотраслевых связей, о существующей технологии общественного производства и т.д.

Матрицу A называют неразложимой, если одновременной перестановкой строк и столбцов ее нельзя привести к виду

$$A = \begin{Bmatrix} A_1 & A_2 \\ 0 & A_3 \end{Bmatrix},$$

где A_1 и A_3 – квадратные подматрицы размеров $k \times k$ и $(n-k) \times (n-k)$, соответсвенно. Неразложимость матрицы A означает, что каждая отрасль использует (хотя бы косвенно) продукты всех других отраслей.

Пусть λ_A – число Фробениуса, оно равно $\lambda_{max}(A)$ – максимальному собственному числу матрицы A. Число Фробениуса λ_A для неотрицательной матрицы $A \geq 0$ всегда положительно и не меньше, чем абсолютное значение любого другого собственного числа матрицы A (теорема Фробениуса–Перрона).

[1]Сингулярным числом матрицы A есть арифметическое значение квадратного корня соответствующего

Теорема 1 [3]. *Неотрицательная матрица A продуктивна тогда и только тогда, когда $\lambda_A < 1$.*

Если матрица A является продуктивной, то неотрицательной будет матрица

$$B = (I - A)^{-1},$$

где I – единичная $n \times n$-матрица. Матрица B называется матрицей коэффициентов полных затрат (обратной матрицей Леонтьева). Для продуктивной матрицы A и соответствующей ей матрицы B числа Фробениуса связаны соотношениями

$$\lambda_B = \frac{1}{1 - \lambda_A} \qquad \text{и} \qquad \lambda_A = 1 - \frac{1}{\lambda_B}.$$

Прямая и двойственная модели Леонтьева

Прямой моделью Леонтьева является модель "затраты-выпуск", описываемая равенством

$$y = (I - A)x, \tag{1}$$

где $x = (x_1, \ldots, x_n)^T$ – вектор валового продукта и $y = (y_1, \ldots, y_n)^T$ – вектор конечного продукта, I – единичная $n \times n$ – матрица. Здесь и везде далее T – символ траспонирования. Если матрица A является продуктивной, то

$$x = By, \quad \text{где} \quad x \geq 0 \qquad \text{для любого } y \geq 0. \tag{2}$$

Двойственной моделью Леонтьева является модель равновесных цен, описываемая равенством

$$w = (I - A^T)p, \tag{3}$$

где $p = (p_1, p_2, \ldots, p_n)^T$ – вектор цен (p_i – цена единицы продукта i–ой отрасли), $w = (w_1, w_2, \ldots, w_n)^T$ – вектор норм добавленной стоимости. Если матрица A является продуктивной, то

$$p = B^T w, \quad \text{где} \quad p \geq 0 \qquad \text{для любого } w \geq 0. \tag{4}$$

собственного числа матрицы $\mathrm{A}^T\mathrm{A}$ или, что то же самое, матрицы $\mathrm{A}\mathrm{A}^T$.

Прямую и двойственную модели Леонтьева связывает соотношение

$$p^T y = w^T x, \tag{5}$$

которое означает, что национальный продукт совпадает с национальным доходом. Соотношение (5) следует из справедливости следующей цепочки равенств

$$p^T y = p^T (I - A)x = \left((I - A^T)p\right)^T x = w^T x.$$

Квадратичная экстремальная задача

Пусть вектор $y \geq 0$ $(\|y\| = 1)$ задает нормированную структуру конечного выпуска в прямой модели Леонтьева (1), а вектор $c \geq 0$ $(\|c\| = 1)$ задает нормированную структуру добавленной стоимости в двойственной модели Леонтьева (3). Здесь $\|\cdot\|$ – евклидова норма вектора.

Рассмотрим следующую задачу нелинейного программирования [4]

$$f^* = (p^*)^T y^* = \max_{y \in R^n, p \in R^n} p^T y \quad \equiv \quad \max_{x \in R^n, w \in R^n} w^T x = (w^*)^T x^* \tag{6}$$

при ограничениях

$$y = (I - A)x, \qquad x \geq 0, \qquad y \geq 0, \tag{7}$$

$$w = (I - A^T)p, \qquad p \geq 0, \qquad w \geq 0, \tag{8}$$

$$\|y\|^2 = 1, \qquad \|w\|^2 = 1, \tag{9}$$

где матрица A – заданная матрица Леонтьева (неотрицательная $n \times n$-матрица), а неизвестными являются компоненты n–мерных векторов x, y, p, w.

Условие эквивалентности "≡" в целевой функции (6) означает использование либо целевой функции $p^T y$, либо $w^T x$. Это следует из справедливости соотношения (5), которое связывает прямую и двойственную модели Леонтьева, заданные ограничениями (7) и (8) соответственно. Следовательно, в задаче (6) – (9) требуется найти такие нормированные векторы конечного выпуска (вектор y^*) и добавленной стоимости (вектор

w^*), чтобы с точностью до некоторого постоянного множителя максимума достигал национальный продукт (он же равен национальному доходу). Оптимальным значениям векторов y^* и w^* будут соответствовать такие значения векторов x^* и p^*, которые с точностью до постоянных множителей будут определять оптимальный валовой продукт в прямой модели Леонтьева и оптимальные цены в двойственной модели Леонтьева.

В общем случае нахождение векторов y^* и w^*, x^* и p^* требует использования численных методов оптимизации для решения квадратичной экстремальной задачи (6) – (9). Здесь целевая функция (6) является билинейной функцией либо от переменных p и y, либо от переменных w и x, а ограничения (9) содержат два квадратичных равенства. Однако оказывается, что в важных для экономических приложений случаях задача (6)–(9) может быть решена аналитически в терминах собственных чисел и собственных векторов симметричных матриц.

Алгоритм для продуктивной матрицы A

Если матрица A продуктивна и матрица $B = (I - A)^{-1}$, то задача (6) – (9) может быть решена в два этапа:

Этап 1. Находим векторы y^* и w^* путем решения квадратичной задачи

$$f^* = (w^*)^T B y^* = \max_{y \geq 0,\, w \geq 0} w^T B y \qquad (10)$$

при ограничениях

$$\sum_{i=1}^{n} y_i^2 = 1, \qquad \sum_{i=1}^{n} w_i^2 = 1.$$

Этап 2. Вычисляем $x^* = By^*$ и $p^* = B^T w^*$.

Теорема 2 [5]. *Если матрица A – продуктивна, то решение задачи* (6) – (9) *имеет вид*

$$y^* = \xi, \quad x^* = By^*, \quad w^* = \eta, \quad p^* = B^T w^*,$$

где ξ *и* η *– неотрицательные нормированные собственные векторы матриц* $B^T B$ *и* BB^T, *отвечающие их максимальным собственным числам* $\lambda_{max}(B^T B)$ *и* $\lambda_{max}(BB^T)$. *Оптимальное значение целевой функции (6) равно*

$$f^* = \sqrt{\lambda_{max}(B^T B)} = \sqrt{\lambda_{max}(BB^T)} = \sigma_B, \qquad (12)$$

где σ_B *– максимальное сингулярное число матрицы* B.

Для продуктивной матрицы A векторы y^* и w^* не обязательно определяются единственным образом. Для их единственности достаточно чтобы продуктивная матрица A была еще и неразложимой.

Теорема 3 [6]. *Если матрица Леонтьева* A *– продуктивна и неразложима, ей соответствует матрица* $B = (I - A)^{-1}$, *то задача* (6) – (9) *имеет единственное решение, все компоненты которого положительны. Это решение имеет вид*

$$y^* = \xi, \quad x^* = By^*, \quad w^* = \eta, \quad p^* = B^T w^*,$$

где ξ *и* η *– положительные нормированные собственные векторы матриц* $B^T B$ *и* BB^T, *соответствующие их максимальным собственным числам* $\lambda_{max}(B^T B)$ *и* $\lambda_{max}(BB^T)$.

Дадим эквивалентную формулировку теоремы 3, которая проясняет содержательный смысл собственных векторов и максимального сингулярного числа.

Теорема 4. *Если матрица* A *– продуктивна и неразложима,* σ_B *– максимальное сингулярное число матрицы* $B = (I - A)^{-1}$, *тогда задача* (6) – (9) *имеет единственное решение*

$$f^* = \sigma_B, \quad y^* = \xi, \quad x^* = \sigma_B \bullet \eta, \quad w^* = \eta, \quad p^* = \sigma_B \bullet \xi,$$

где ξ *and* η *– положительные нормированные собственные векторы матриц* $B^T B$ *и* BB^T, *соответствующие их максимальным собственным числам.*

Для продуктивной и неразложимой матрицы Леонтьева A теорема 4 определяет оптимальную нормированную структуру конечного продукта (вектор y^*) и оптимальную нормированную структуру добавленной стоимости (вектор w^*). Первой соответсвуют компоненты собственного вектора ξ, а второй – компоненты собственного вектора η.

Векторы y^* и w^* для Украины (15 отраслей)

Межотраслевой баланс экономики Украины с 2000 года ведется по 38 отраслям (кодам видов экономической деятельности). Он публикуется в ежегодных статистических сборниках Госкомстата Украины, которые можно найти на сайте http://www.ukrstat.gov.ua. Широко используется агрегированный 15-отраслевой баланс, который построен в результате объединения нескольких отраслей из 38-отраслевого баланса в одну агрегированную: так, например, добыча угля и торфа, добыча углеводородов и добыча неэнергетических материалов группируются в сектор добывающей промышленности. Названия всех 15 агрегированных отраслей приведены на рис. 1 для фрагмента матрицы Леонтьева за 2009 год.

В табл. 1 даны оптимальные нормированные структуры конечного продукта (векторы y^*) и оптимальные нормированные структуры добавленной стоимости (векторы w^*) для 15-отраслевых матриц Леонтьева [7]. Эти матрицы построены на основе таблиц "затраты-выпуск" в ценах потребителей за 2003 – 2009 годы [8]. Использовался следующий способ их построения. Для каждой отрасли j из таблиц были взяты ее валовый выпуск V_j и объем продукции $\tilde{a}_{ij}$ отрасли i, израсходованный отраслью j в процессе производства. Коэффициенты матрицы Леонтьева a_{ij} получены в результате деления этих чисел: $a_{ij} = \tilde{a}_{ij}/V_j$.

матрица Леонтьева

№	Название отрасли / № отрасли	1	2	3	4	5
1	Сельское хозяйство, охотничье и лесное хозяйство	0,25644	0,07763	0,00214	0,03313	0,00017
2	Рыбное хозяйство	0,00017	0,07457	0,00001	0,00049	0,00001
3	Добывающая промышленность	0,01008	0,00367	0,06446	0,11893	0,33387
4	Перерабатывающая промышленность	0,18065	0,18032	0,15941	0,29734	0,11449
5	Производство и распределение электроэнергии, газа и воды	0,01163	0,02934	0,08042	0,02750	0,07150
6	Строительство	0,00019	0,00000	0,00107	0,00025	0,00176
7	Торговля, ремонт автомобилей, бытовых изделий и предметов личного пользования	0,12371	0,21394	0,06930	0,20672	0,00137
8	Деятельность гостиниц и ресторанов	0,00025	0,00122	0,00194	0,00116	0,00254
9	Деятельность транспорта и связи	0,04232	0,08924	0,12953	0,04982	0,01114
10	Финансовая деятельность	0,00225	0,00428	0,00749	0,00809	0,01582
11	Операции с недвижимым имуществом, аренда, инжиниринг и предоставление услуг	0,00860	0,01284	0,01234	0,01477	0,01549
12	Государственное управление	0,00032	0,00122	0,00207	0,00242	0,00726
13	Образование	0,00006	0,00000	0,00042	0,00011	0,00052
14	Здравоохранение и предоставление соц. помощи	0,00032	0,00244	0,00134	0,00045	0,00093
15	Предоставление коммунальных и индивидуальных услуг, деятельность в сфере культуры и спорта	0,00018	0,00061	0,00151	0,00083	0,00222

Рис. 1. Фрагмент 15-отраслевой матрицы Леонтьева A за 2009 год

Таблица 1.

Оптимальные векторы y* и w* (Украина, 15 отраслей)

	2003		2004		2005		2006		2007		2008		2009	
λ_B	2.418		2.408		2.476		2.409		2.338		2.305		2.323	
σ_B	2,914		2,937		3,107		2,980		2,865		2,884		2,866	
	y*	w*	y*	w*	y*	w*	y*	w*	y*	w*	y*	w*	y*	w*
1	0,26	0,24	0,26	0,21	0,27	0,20	0,28	0,20	0,28	0,20	0,28	0,18	0,28	0,19
2	0,28	0,10	0,26	0,09	0,30	0,10	0,29	0,10	0,27	0,10	0,28	0,11	0,26	0,10
3	0,33	0,35	0,32	0,33	0,30	0,29	0,28	0,28	0,25	0,25	0,24	0,27	0,28	0,30
4	0,55	0,76	0,57	0,78	0,55	0,81	0,55	0,80	0,56	0,78	0,56	0,79	0,54	0,76
5	0,26	0,20	0,24	0,17	0,22	0,16	0,22	0,16	0,23	0,16	0,23	0,16	0,25	0,18
6	0,30	0,11	0,33	0,12	0,33	0,12	0,35	0,12	0,35	0,14	0,38	0,14	0,37	0,14
7	0,19	0,27	0,17	0,26	0,23	0,29	0,24	0,30	0,23	0,32	0,22	0,30	0,23	0,32
8	0,27	0,10	0,26	0,10	0,28	0,09	0,24	0,09	0,23	0,09	0,22	0,09	0,25	0,10
9	0,21	0,22	0,21	0,24	0,23	0,22	0,24	0,23	0,26	0,24	0,26	0,25	0,25	0,26
10	0,12	0,11	0,11	0,14	0,09	0,09	0,07	0,07	0,08	0,09	0,07	0,08	0,06	0,07
11	0,17	0,15	0,17	0,16	0,16	0,14	0,19	0,16	0,21	0,20	0,19	0,18	0,21	0,22
12	0,15	0,06	0,16	0,07	0,11	0,05	0,10	0,04	0,11	0,05	0,12	0,05	0,09	0,04
13	0,10	0,04	0,09	0,03	0,10	0,03	0,10	0,04	0,09	0,03	0,10	0,04	0,11	0,04
14	0,18	0,07	0,18	0,07	0,17	0,06	0,17	0,06	0,17	0,06	0,17	0,06	0,17	0,06
15	0,16	0,07	0,16	0,07	0,15	0,06	0,15	0,07	0,15	0,08	0,13	0,07	0,14	0,08

Из табл. 1 видим достаточно высокую устойчивость компонент для ряда отраслей в парах векторов (y^*, w^*) для разных лет. Кроме того, максимальные сингулярные числа σ_B на 20-25% больше, чем числа Фробениуса λ_B. Это означает, что по критерию максимизации национального дохода оптимальные нормированные структуры конечного выпуска и добавленной стоимости лучше, чем их нормированные аналоги, найденные с помощью векторов Фробениуса.

Похожая ситуация имеет место для 22-отраслевых матриц Леонтьева (табл. 1) из системы таблиц "затраты-выпуск" России за 2001 – 2003 годы. Оптимальные векторы y^* и w^* для этих матриц даны в работе [9].

Заключение

Модель (6) – (9), объединяющая прямую и двойственную статические модели Леонтьева, позволяет использовать для анализа экономической системы сингулярные числа и собственные векторы некоторых симметричных матриц. С их помощью можно исследовать связи между затратами на производство продукции и ценами при распределении продукции в экономической системе. Это пополняет арсенал средств для анализа качественных свойств леонтьевских моделей, который можно осуществить с помощью чисел и векторов Фробениуса. Усовершенствование модели (6) – (9) позволит оценить долю косвенных затрат в процессе производства как системы в целом, так и ее отдельных отраслей. Это дает возможность обнаружить такие коэффициенты в матрице Леонтьева, изменения которых необходимо отслеживать в первую очередь.

Работа выполнена при частичной финансовой поддержке SNSF (Швейцария), проект №127962 "Analysis of Institutional and Technological Changes in Market and Transition Economies on the Background of the Present Financial Crisis". Автор благодарен доктору Жан-Франсуа Эмменеггеру и профессору Генриху Бортису (университет Фрибурга, Швейцария) за внимание к этой работе, активное ее обсуждение и полезные замечания.

Список литературы

[1] Kurz H.D. Who is going to kiss sleeping beauty? on the 'classical' analytical origins and perspectives of input-output analysis // Review of Political Economy, 2011. – Vol. 23(1). – 25 – 47.

[2] Bortis H. Keynes and the Classics: Notes on the Monetary Theory of Production, in: Modern Theories of Money. The Nature and Role of Money in Capitalist Economies, Rochon L.-P. and Sergio Rossi (eds), 2003. – Edward Elgar: UK, USA. – Pp. 411 – 474.

[3] Ашманов С.А. Введение в математическую экономику. – М.: Наука, 1984. – 296 с.

[4] Стецюк П.И., Кошлай Л.Б., Пилиповский А.В. О задаче оптимального соотношения между спросом и добавленной стоимостью в моделях Леонтьева // Теорія оптимальних рішень. – 2010. – № 9. – С. 136 – 143.

[5] Стецюк П.И., Кошлай Л.Б. Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева // Кибернетика и системный анализ. – 2010. – №5. – С. 51 – 59.

[6] Стецюк П.И., Кошлай Л.Б. Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева // Спектральные и эволюционные задачи. – 2011. – Т. 2. – №2. – С. 164 – 169.

[7] Стецюк П.И., Бондаренко А.В. О спектральных свойствах модели Леонтьева // Теорія оптимальних рішень. – 2011. – №10. – С. 84 – 90.

[8] Статистическая информация [Электронный ресурс]: Таблица «Затраты-выпуск» (в ценах потребителей) / Госкомстат Украины. – http://www.ukrstat.gov.ua. – Режим доступа: свободный.

[9] Стецюк П. И. О спектральных свойствах матриц Леонтьева // Сборник трудов Всероссийской научно-практической конференции «Статистика, моделирование, оптимизация» (28 ноября – 2 декабря 2011 г., г. Челябинск, Южно-Уральский государственный университет).

Научное издание

Кнопов Павел Соломонович
Зоркальцев Валерий Иванович
Иваньо Ярослав Михайлович и др.

Стохастическое программирование и его приложения

Технические редакторы
Е.И. Литвинцева, С.М. Пержабинский

ISBN 978-5-93908-110-8

Издано на электрон. опт. диске (CD-ROM): цв.; 12 см. – Систем. требования: ПК с процессором Pentium3; Microsoft Windows XP или выше; 2-скоростной дисковод CD-ROM; 256 цв. SVGA дисплей; Acrobat reader 7.0 или выше. – Загл. с экрана. – Иркутск. – ИСЭМ СО РАН. – 2012. – 493с.

Тираж 500 экз.