



CONTROL SCIENCES

**Научно-технический
журнал**

6 номеров в год

ISSN 1819-3161

УЧРЕДИТЕЛЬ

**Институт проблем управления
им. В.А. Трапезникова РАН**

Главный редактор

Д.А. Новиков

**Заместители главного
редактора**

Л.П. Боровских, Ф.Ф. Пащенко

Редактор

Т.А. Гладкова

Выпускающий редактор

Л.В. Петракова

Издатель

ООО «СенСиДат-Контрол»

Адрес редакции
117997, ГСП-7, Москва,
ул. Профсоюзная, д. 65, к. 272.
Тел./факс (495) 334-92-00

E-mail: pu@ipu.ru
www.ipu.ru/period/pu

Оригинал-макет
и электронная версия
подготовлены
ООО «Адвансед Солюшнз»

Отпечатано с готовых диапозитивов
в типографии ГКС

Подписано в печать
5.06.2008 г.

Заказ № РВ408

Журнал зарегистрирован
в Министерстве
Российской Федерации
по делам печати,
телерадиовещания
и средств массовых
коммуникаций

Свидетельство о регистрации
ПИ №77-11963
от 06 марта 2002 г.

Журнал входит в Перечень ведущих
рецензируемых журналов и изданий,
в которых должны быть опубликованы
основные научные результаты
диссертаций на соискание ученой
степени доктора и кандидата наук

Подписные индексы:
80508 и **81708** в каталоге Роспечати
38006 в объединенном каталоге
«Пресса России»

ПРОБЛЕМЫ УПРАВЛЕНИЯ

4.2008

СОДЕРЖАНИЕ

Математические проблемы управления

- Клещев А.С.** Концепция банка математических знаний для научных исследований. Ч. 1. Метафора 2
Буркова И.В. Метод сетевого программирования в симметричной задаче коммивояжера 7
Агаев Р.П., Никифоров С.В., Андрушина Н.А. О спектре матрицы смежности орграфа кольцевой структуры и его применении 11
Рыков А.С., Матвиенко М.Ю. Недифференцируемые овражные двумерные тест-функции 16

Анализ и синтез систем управления

- Яковенко Г.Н.** Управляемые системы, робастные по начальным данным . . . 20
Соколов С.В., Кучеренко П.А. Обобщенные вероятностные критерии в задаче нелинейной параметрической идентификации 24

Управление в социально-экономических системах

- Чеботарев П.Ю., Логинов А.К., Цодикова Я.Ю.** и др. Анализ феноменов коллективизма и эгоизма в контексте общественного благосостояния 30
Бурков В.Н., Исаков М.Б., Коргин Н.А. Применение обобщенных медианных схем для построения неманипулируемых механизмов многокритериальной активной экспертизы 38
Кувшинов Б.М., Челядин А.С., Ширяев В.И. Многокритериальная оценка конкурентоспособности и ранжирование экономических субъектов, функционирующих в разнородных условиях при наличии неопределенности 48
Байкин А.А., Иванов Е.Ю., Исаева О.В. Управление выбором производителя с учетом влияния конъюнктурной информации 55

Информационные технологии в управлении

- Жевнеров В.А.** Метод эволюции параметров в нелинейных сетевых задачах на примере оптимизации распределения входящих потоков нагрузки. 65

Управление технологическими процессами

- Кирич Ю.П., Затонский А.В., Беккер В.Ф., Краев С.Л.** Идентификация технологических процессов производства губчатого титана. 71

Системный анализ и обработка данных

- Дорофеев Ю.А.** Структурно-классификационные методы анализа и прогнозирования в крупномасштабных системах управления 78

* * *

- Contents and abstracts 84

КОНЦЕПЦИЯ БАНКА МАТЕМАТИЧЕСКИХ ЗНАНИЙ ДЛЯ НАУЧНЫХ ИССЛЕДОВАНИЙ.

Ч. 1. Метафора¹

А.С. Клещев

Институт автоматики и процессов управления ДВО РАН, г. Владивосток

На основе расширяемой внутренней модели математической практики и модели аналогии между доказательствами предложена концепция системы компьютерной поддержки научной деятельности в области математики и реализующих ее механизмов. Приведена метафора системы и структура внешней модели математической практики.

ВВЕДЕНИЕ

Доказательство теорем является творческой деятельностью. Несмотря на то, что ей пытаются учить детей еще в школе, она остается трудной и для профессиональных математиков. Научная значимость автоматизации этой деятельности обсуждалась в статье [1]. Основная в области теоретической математики, она занимает значительное место и в прикладной математике, а также в науках, связанных с активным использованием математических знаний, — в физике, технических науках и др. Наконец, и в программировании доказательство правильности алгоритмов и программ составляло мечту многих выдающихся ученых. Однако ни одна из существующих систем автоматического или интерактивного доказательства теорем не используется систематически в научной деятельности.

Цель настоящей статьи заключается в разработке концепции и реализующих ее механизмов для системы компьютерной поддержки научной деятельности в области математики. В первой части приводится метафора системы и структура внешней модели математической практики.

¹ Работа выполнена при финансовой поддержке РФФИ, проект 06-070-89071-а, и ДВО РАН в рамках Программы № 15 ОЭММПУ РАН, проект 06-1-П15-055.

1. СИСТЕМЫ ИНТЕРАКТИВНОГО ПОСТРОЕНИЯ ДОКАЗАТЕЛЬСТВ

Идея формализации математики имеет долгую традицию. Возможность чисто синтаксического построения математики и механической проверки правильности доказательств была показана в первой половине XX в. Во второй половине были разработаны компьютерные системы, с помощью которых были получены некоторые нетривиальные математические результаты. В течение последних десятилетий были созданы компьютерные системы и среды для проверки правильности и помощи в построении доказательств, такие как Coq [2], Isabelle [3] и HOL [4]. Эти средства обеспечивают высокую степень уверенности в правильности полученных с их помощью доказательств. Они нашли применение для верификации элементной базы и программного обеспечения компьютеров, но построение корпуса механически верифицированной математики пока остается недостижимой мечтой [5]. Другим приложением является обучение математике [6].

Большинство из этих систем однопользовательские, а некоторые — свободно распространяемые. Их установка на персональном компьютере довольно сложна, причем пользователь в результате получает «пустую» систему (без базы знаний). Исключение составляет система ActiveMath [6], где



Интернет используется как средство общения между преподавателями и студентами.

Работа с такой системой начинается с формирования некоторого проекта, связанного, например, с поиском доказательства некоторой теоремы или с формальной верификацией программы. Проект представляет собой одну или несколько теорий, образующих сеть (граф без циклов), в которой каждый потомок расширяет одну или несколько родительских теорий. Формированием теорий занимается «специалист по спецификациям». Лишь после того, как сформированы все необходимые теории, «специалист по доказательствам» может приступить к интерактивному построению доказательств теорем [7]. Если при построении очередного доказательства выясняется, что в сформированных теориях не хватает необходимых для этого знаний (аксиом, определений или теорем), пользователь вынужден приостанавливать построение доказательства и переходить в режим формирования теорий.

При формировании теорий «специалист по спецификациям» вынужден использовать язык с фиксированным синтаксисом, являющийся либо языком исчисления предикатов, либо его расширением, допускающим некоторые классы арифметических, теоретико-множественных и иных формул. «Специалист по доказательствам» вынужден строить полные доказательства в рамках фиксированного исчисления предикатов первого или более высокого порядка. Доказательство формируется как последовательность команд. На очередном шаге «специалист по доказательствам» анализирует состояние доказательства и формирует очередную команду. Система выполняет ее и представляет ему новое состояние. Состояние доказательства характеризуется несколькими открытыми (еще не доказанными) целями. Кроме того, оно может иметь и другие элементы, такие как локальные описания и предположения. Построение доказательства заканчивается, когда достигнуто состояние, в котором нет открытых целей [8, 9].

Опыт применения таких систем показал, что построение полных доказательств в рамках фиксированного исчисления слишком трудоемко. Математик обычно имеет долговременный план для сложных доказательств, который считается непредставимым на уровне исчисления (отдельных шагов вывода). Более того, в хорошо известной ему области он имеет множество стратегий, техник и приемов доказательства. Идея моделирования этих способностей на компьютере состоит в следующем. План есть представление доказательства на более высоком уровне абстракции, чем уровень логического исчисления. Он состоит из последовательности макрооператоров, таких как применение гомоморфизмов, применение лемм, некоторые упрощения, дифференцирование или интегрирование функций, вызов системы автоматического до-

казательства или компьютерной алгебры, специализированной процедуры и др. Каждый из этих макрооператоров может быть раскрыт в последовательность шагов вывода и называется тактикой. Методы доказательств получаются из тактик добавлением к ним предусловий и постусловий. Стратегии доказательств объединяют вместе взаимосвязанные методы и поисковые эвристики. В таких расширенных исчислениях построение доказательств заменяется построением планов доказательств (автоматическим или интерактивным). План есть последовательность методов. Доказательство является результатом выполнения плана. Для облегчения процесса построения доказательств или их планов некоторые системы содержат механизм предложения команд: пользователю выдается список команд, применимых в текущем состоянии доказательства или его плана, из которого он выбирает необходимую команду и уточняет ее аргументы [10, 11].

В процессе взаимодействия с системами интерактивного построения доказательств пользователь вынужден оперировать такими понятиями, как: проект, теория, иерархия теорий, потомок теории, предок теории; логический контекст, логический базис, внелогический контекст; аксиома, теорема; формула, терм, константа, тип, сорт; доказательство, цель, главная цель, шаг доказательства, команда построения доказательства, состояние доказательства, открытая цель, попытка доказательства. Интерфейс таких систем основан на двух альтернативных принципах — текстовом вводе информации, с последующим контролем ее синтаксической правильности, и совокупности выборов информации из списков. Обычно обе эти возможности присутствуют в интерфейсе в том или ином сочетании [9].

Рассмотрим, в какой мере системы, обладающие подобными свойствами, пригодны для использования в математических исследованиях.

Сложная установка однопользовательской системы на персональном компьютере не может быть приемлемой, поскольку ученый-математик не всегда обладает достаточной для этого компьютерной грамотностью. Интернет-системы в этом отношении явно предпочтительнее. Однако прямое взаимодействие пользователей между собой во время сеансов их работы вряд ли требуется.

«Пустые» системы вообще неприменимы для научных исследований. Средства поддержки научной деятельности могут использоваться лишь тогда, когда содержат значительную часть уже накопленных математических знаний. Принцип формирования отдельных «проектов», т. е. сетей баз знаний, специально организованных для целей доказательства отдельной теоремы или совокупности содержательно связанных теорем, также неприемлем. Исследователям нужна общая база математических знаний, к которой они смогут быстро при-

выкнуть, причем организованная таким образом, чтобы они могли достаточно легко ориентироваться в ней при поиске необходимой информации.

Разделение ролей «специалиста по спецификации» теорий и «специалиста по доказательству» теорем, а также строго последовательный порядок их работы противоречат математической практике, поскольку в работе исследователя эти два вида деятельности тесно переплетены и постоянно чередуются.

Хотя при написании формул математик вынужден соблюдать целый ряд правил синтаксиса, однако переход при их вводе с математического диалекта на его формальный аналог, который требует знания значительного числа дополнительных правил, также является неприемлемым. Столь же неприемлем фиксированный синтаксис такого языка, поскольку математический диалект постоянно развивается по мере развития математики. Полные доказательства не строятся в математической практике, несмотря на призывы ревнителей их достоверности. Поэтому необходимость построения полных доказательств отпугивает от таких систем большинство математиков. Наконец, неприемлема и фиксированная формальная система (исчисление), в рамках которой строятся доказательства, поскольку до сих пор неизвестны попытки практически приемлемой кодификации правил математического рассуждения. Построение доказательства в виде последовательности команд можно считать приемлемым, если набор возможных команд удовлетворителен для пользователей.

Возможность расширять исчисление, несомненно, прогрессивная идея. Однако такие расширения должны выполняться либо самими математиками в ходе построения доказательств, либо системой на основе анализа этих доказательств. Использование процедурного метаязыка (или макроязыка) для описания повторно используемых тактик, методов и стратегий исключает возможность участия большинства математиков в расширении исчисления. Выделение же работы по расширению исчисления в самостоятельную деятельность требует организации оперативного взаимодействия между ее исполнителями (программистами) и математиками, у которых возникает потребность в таких расширениях при построении доказательств. Самостоятельной проблемой выступает обеспечение правильности тактик, методов и стратегий, представленных в виде программ, относительно их спецификации (предусловий и постусловий) и содержательного описания (без которого их использование математиками оказывается невозможным). Это же, но в еще большей степени, относится и к повторно используемым планам доказательств. Важен механизм предложения команд при условии, что математик понимает, что ему предлагается.

Желательно, чтобы при взаимодействии с подобными системами пользователю приходилось оперировать как можно меньшим числом нематематических понятий, из которых большую часть желательно выдать за понятия, в терминах которых в математической литературе обычно описывается процесс построения доказательств. В процессе построения доказательства текстовый ввод информации с последующим синтаксическим контролем неприемлем по двум причинам: он требует знания математиками синтаксиса такого языка (а они не хотят учить новые формальные языки); в случае совершения синтаксических ошибок процесс их обнаружения и устранения может оказаться для математиков слишком сложным (из-за неточной диагностики и наведенных ошибок). Выбор информации из списков не обладает этими недостатками, но не всегда привычен.

Таким образом, на пути применения систем интерактивного построения доказательств для научных применений лежит еще много нерешенных проблем.

2. МЕТАФОРА БАНКА МАТЕМАТИЧЕСКИХ ЗНАНИЙ

В качестве метафоры для компьютерной поддержки математических исследований (назовем такую систему «Банк математических знаний» (БМЗ)) возьмем организацию математической практики. Чем ближе такой банк будет к этой метафоре, тем более востребованным он может оказаться.

Цель математических исследований состоит в расширении математических знаний, в том числе в формулировании новых аксиоматических систем, определений и математических утверждений (теорем, лемм и т. п.), а также их доказательстве. Доступные для выполнения этой деятельности и расширяемые в результате нее математические знания, помимо персональных знаний ее участников, практически недоступны для других, распределены между учебниками и справочниками (где их легче всего найти), монографиями (где их найти несколько труднее), а также научными статьями (где их найти труднее всего в силу обилия таких статей). В этой деятельности принимает участие большое число специалистов, которых можно условно разделить на потребителей знаний (преподавателей, студентов и прикладных специалистов), генераторов знаний или исследователей (тех, кто занимается получением новых математических знаний) и интеграторов знаний (тех, кто занимается систематизацией математических знаний). Источниками для потребителей знаний служат учебники, справочники, реже монографии, еще реже научные статьи. Результатом деятельности исследователей являются научные статьи, а результатом деятельности интеграторов — монографии, учебники и справочники. Если исследователь в статье использует (и вводит) лишь такое множес-



тво понятий, которое достаточно, чтобы представить формулировки и доказательства новых утверждений, то интегратор вынужден формировать более широкую систему понятий, в которой могут быть представлены все включенные в монографию, справочник или учебник утверждения и доказательства. Поэтому интегратор, используя свои и (или) чужие статьи, формирует более или менее значительный фрагмент математических знаний. Если в результате деятельности исследователей знания растут экстенсивно, то интеграторы превращают те или иные их части в системы (способствуют интенсивному накоплению знаний).

Банк математических знаний следует рассматривать как систему для накопления достоверных математических знаний. Реальное накопление знаний возможно лишь в том случае, если в этой работе может принять участие любой заинтересованный исследователь. Не менее важна и деятельность интеграторов, в задачу которых входит приведение разрозненных знаний, соответствующих интересам отдельных исследователей, в системы. Наконец, знания, накапливаемые в БМЗ, должны быть широко доступны, чтобы все заинтересованные потребители знаний могли не только ими пользоваться, но и страховать возможные ошибки исследователей и интеграторов. Интернет является той средой, которая может обеспечить широкое и удобное участие всех заинтересованных лиц в работе БМЗ. Принципы организации подобных систем рассматривались в статье [12].

Пользователей БМЗ можно разделить на четыре группы — гостей (потребителей знаний), исследователей (генераторов знаний), интеграторов знаний и администраторов. Гость может лишь просмотреть общую базу знаний БМЗ, а также сообщить администраторам о замеченных в ней недостатках. Математик может сделать заявку на участие в работе БМЗ в качестве исследователя, в ответ получить от администратора полномочие на развитие знаний БМЗ и использовать это полномочие, формируя свою персональную базу знаний (виртуальный аналог своего рабочего стола). Интеграторы знаний рассматривают результаты деятельности каждого исследователя по развитию знаний БМЗ, находящиеся в его персональной базе знаний. Они решают вопрос о включении результата исследователя в общую базу знаний БМЗ. Системные процессы БМЗ поддерживают деятельность исследователей и интеграторов знаний, освобождая их от части рутинной работы.

3. СТРУКТУРА ВНЕШНЕЙ МОДЕЛИ МАТЕМАТИЧЕСКОЙ ПРАКТИКИ

В работе [13] были введены понятия внутренней и внешней моделей математической практики и предложена внутренняя модель для БМЗ. Далее

вводится структура внешней модели для БМЗ и устанавливается соответствие между компонентами математической практики и ее внешней модели.

Математическому диалекту может быть поставлена в соответствие грамматика модели математического диалекта, представленная в виде расширенных форм Бекуса—Наура. Исследователи и интеграторы знаний должны иметь возможность включать в текущее состояние этой грамматики новые правила либо модифицировать существующие для расширения языка. Грамматика не охватывает язык пропозициональной логики, поскольку он является фиксированным (и может быть встроен в БМЗ жестко). Язык описания нелогических правил рассуждения (метаязык внутренней модели) является надстройкой над моделью математического диалекта, расширяется благодаря ей и определяется лишь набором типов синтаксических переменных. Каждый тип характеризуется нетерминальным символом грамматики модели математического диалекта, терминальные порождения которого образуют множество возможных значений синтаксических переменных этого типа. Исследователи и интеграторы знаний должны иметь возможность вводить новые типы синтаксических переменных.

Математическим знаниям соответствует общая база знаний, представляющая собой сеть разделов. Всею сетью разделов должен управлять (вводить новые разделы и назначать для них интеграторов знаний) глобальный администратор БМЗ. Информацией внутри каждого раздела должен управлять интегратор знаний этого раздела. Каждый раздел может иметь несколько предков и потомков. Раздел содержит онтологические знания (системы аксиом, определения терминов и новых формальных способов записи), теоремы и модели их интуитивных доказательств. В разделе-потомке справедливы онтологические знания всех его предков. Информации, содержащейся на рабочем столе исследователя и относящейся к его научной работе, соответствует его персональная база знаний. В ней он может вводить свои разделы, онтологические знания, теоремы и доказательства.

Онтологические знания и теоремы представляются средствами модели математического диалекта. Модель каждой теоремы, кроме формулировки, может иметь название (для облегчения ее поиска).

Интуитивным доказательствам соответствуют модели интуитивных доказательств. В теории доказательств обычно различают полное, формальное и интуитивное доказательства. Доказательство является полным, если оно правильно (т. е. правильные правила рассуждения и методы доказательства на каждом шаге правильно применяются к посылкам для вывода следствий), показывает, как доказываемое утверждение следует из посылок, и при этом не ссылается ни на какие недока-

занные утверждения, кроме онтологических знаний. Доказательство является формальным, если оно выполнено в рамках некоторой формальной системы. Последнее означает, что все математические утверждения, использованные в доказательстве, представлены на некотором формальном языке, а все использованные в нем правила рассуждения являются правилами вывода некоторой формальной системы, представленными на некотором формальном метаязыке. Понятно, что эти два определения являются конструктивными — полное или формальное доказательство можно разобрать на части (математические утверждения, правила рассуждения или вывода, методы доказательства), для каждой части можно установить ее правильность, а для всего доказательства — правильность и полноту его сборки из этих частей. Доказательство является интуитивным (ИД), если оно представлено в математической литературе. Очевидно, что это определение не является конструктивным; из него следует, что публикация доказательства в математической литературе, делает его интуитивным. Однако смысл этого определения в другом — все доказательства, имеющиеся в математической литературе, считаются интуитивными (если доказательство принимается математическим сообществом, то оно является интуитивным).

У системы, не поддерживающей построения ИД, нет шансов быть востребованной в науке. Это означает, что усилия, которые математик должен затратить на построение любого доказательства с ее помощью, не должны заметно превышать усилий на построение соответствующего ИД обычным способом. Одновременно, такая система должна гарантировать правильность ИД, построенных с ее помощью.

Отсюда возникают два вопроса: как устроено ИД, и как можно проверить (обеспечить) его правильность? Ответ на второй вопрос очевиден — ИД является правильным, если его можно расширить до полного и формального доказательства. На первый вопрос нельзя дать исчерпывающий ответ (для этого требуется провести анализ всех существующих в математической литературе доказательств), но можно попытаться к нему приблизиться. Модели ИД, способам их интерактивного формирования и обеспечения правильности посвящена вторая часть этой статьи.

Модели способов рассуждения могут быть представлены в специальной базе, разделенной на логический (пропозициональные тавтологии) и нелогический (метаматематические аксиомы) разделы. Модель способа рассуждения может иметь название (для облегчения его поиска).

Модели методов доказательства также могут быть представлены в специальной базе. Модель метода может иметь название (для облегчения его поиска).

ЗАКЛЮЧЕНИЕ

Предложена концепция Интернет-системы компьютерной поддержки научной деятельности в области математики. Основная цель такой системы — накопление математических знаний, обеспечение их достоверности, повышение их доступности и поддержка современного стиля работы математиков. Концепция допускает стихийность в работе исследователей в сочетании с централизованным управлением общими знаниями со стороны интеграторов знаний. Для этого вводится разделение базы знаний на общую и множество персональных. Работа исследователей в своих персональных базах прямо не влияет на содержимое общей базы, которая доступна всем пользователям. Интеграторы знаний формируют общую базу на основе персональных.

ЛИТЕРАТУРА

1. Гаврилова Т.Л., Клещев А.С. Анализ подходов к решению проблемы правильности математических знаний // Проблемы управления. — 2005. — № 3. — С. 13—19.
2. *The Coq Proof Assistant User's Guide*, version 6.1 / Cornes C., Courant J., Filliatre J.-C., et al. INRIA-Rocquencourt et CNRSSENS. — Lyon, 1996.
3. Paulson L.C. Isabelle: A Generic Theorem Prover. LNCS 828. — Springer, 1994.
4. Gordon M.J.C. and Melham T.F. (editors). Introduction to HOL: A theorem proving environment for higher order logic. Cambridge University Press, 1993.
5. Aspinall D. Eclipse Proof General. 2004 (<http://proofgeneral.inf.ed.ac.uk/Kit/docs/EIG04.pdf>).
6. Melis E., Meier A. and Pollet M. Adaptive Access to a Proof Planner // Third International Conference on Mathematical Knowledge Management. — Springer-Verlag, 2004. — P. 251—264.
7. Aitken S., Gray Ph. On using GOMS to analyse definition-making in interactive proving. 1997. (<http://www.dcs.gla.ac.uk/~stuart/ITP/ITP.html>).
8. Proof General in Eclipse. System and Architecture Overview / D. Aspinall, et al. 2006. (<http://www.cs.mcgill.ca/%7Emartin/etx2006/papers/30.pdf>).
9. Voelker N. Thoughts on Requirements and Design Issues of User Interfaces for Proof Assistants. — 2003 (<http://www.informatik.uni-bremen.de/~cxl/uitp03/entcs/08-Voelker.pdf>).
10. Melis E. and Siekmann J.H. Concepts in Proof Planning // *Intellectics and Computational Logic*. Kluwer. — 2000. — P. 263—276.
11. Meier A., Melis E., Pollet M. Adaptable Mixed-Initiative Proof Planning for Educational Interaction. — 2003. (<http://www.informatik.uni-bremen.de/~cxl/uitp03/entcs/06-PolletMellisMeier.pdf>).
12. Мультидисциплинарная система управления информационными ресурсами различных уровней общности / И.Л. Артемьева, Т.Л. Гаврилова, В.В. Грибова и др. // Проблемы управления. — 2006. — № 4. — С. 64—68.
13. Гаврилова Т.Л., Клещев А.С. Внутренняя модель математической практики для систем автоматизированного конструирования доказательств теорем // Проблемы управления. — 2006. — № 4. — Ч. 1. — С. 32—35; — № 5. — Ч. 2. — С. 68—73; — № 6. — Ч. 3. — С. 68—71.

☎ (4232) 31-04-24; e-mail: kleshev@iacp.dvo.ru

Статья представлена к публикации членом редколлегии О.П. Кузнецовым. □



МЕТОД СЕТЕВОГО ПРОГРАММИРОВАНИЯ В СИММЕТРИЧНОЙ ЗАДАЧЕ КОММИВОЯЖЕРА

И.В. Буркова

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Сформулирована двойственная задача, состоящая в разбиении ограничений на две группы с соответствующим делением длин дуг на две части и решении двух полученных оценочных задач; сумма целевых функций оптимальных решений оценочных задач дает нижнюю оценку для исходной задачи. Решение оценочной задачи сведено к построению i -деревьев кратчайшей длины. Предложен новый способ получения нижних оценок для оценочных задач, в основе которого лежит построение дерева кратчайших путей. Показано, что построение i -деревьев и дерева кратчайших путей для исходной матрицы расстояний не дает оптимального решения двойственной задачи.

ВВЕДЕНИЕ

Метод сетевого программирования [1] является достаточно универсальным инструментом решения или получения оценок в задачах дискретной оптимизации. Он применялся для точного решения задач о ранце, максимальном потоке, для получения оценок в задачах о камнях, о разрыве контуров, о покрытии двудольного графа и др.

Суть метода для линейных целевых функций заключается в разбиении исходной задачи на несколько подзадач с соответствующим делением коэффициентов целевой функции. Сумма значений целевых функций в оптимальных решениях этих подзадач дает оценку (нижнюю — в случае задачи на минимум) для исходной задачи. Если для каждой подзадачи имеется метод получения нижних оценок целевой функции, то сумма оценок всех подзадач также дает нижнюю оценку для исходной задачи. Задача поиска такого разбиения коэффициентов целевой функции, при котором сумма оценок подзадач максимальна, называется двойственной задачей. В статье рассматривается применение этого подхода к симметричной задаче коммивояжера.

1. ЗАДАЧА, ДВОЙСТВЕННАЯ ЗАДАЧЕ КОММИВОЯЖЕРА

Задан $n + 1$ вершинный граф с длинами дуг l_{ij} , которые интерпретируются как расстояние между городами i и j . Без ограничения общности можно считать все $l_{ij} \geq 0$. Требуется найти кратчайший гамильтонов контур (маршрут коммивояжера). Обозначим $x_{ij} = 1$, если дуга (i, j) входит в маршрут, $x_{ij} = 0$, в противном случае. Тогда задача сводится к минимизации критерия

$$L(x) = \sum_{i,j} l_{ij} x_{ij}$$

при ограничениях

$$\sum_i x_{ij} = 1, \quad \forall j, \quad (1)$$

$$\sum_j x_{ij} = 1, \quad \forall i, \quad (2)$$

$$u_i - u_j + nx_{ij} \leq n - 1, \quad (3)$$

где $u_i \geq 0$; $i, j = 1, 2, \dots, n$; $i \neq j$. Условие (4) гарантирует получение гамильтонова контура [2]. Рас-

сма триваемая задача относится к NP -трудным [3]. Поэтому представляет интерес получение нижних оценок целевой функции.

Рассмотрим симметричную задачу коммивояжера, т. е. $l_{ij} = l_{ji}, \forall i, j$.

Разобьем ограничения задачи на две группы. В одну группу включим ограничения (1) и (3), а во вторую — (2) и (3). Соответственно разделим на две части коэффициенты l_{ij} , т. е. представим l_{ij} в виде

$$l_{ij} = p_{ij} + q_{ij}, \quad p_{ij}, q_{ij} \geq 0, \quad \forall i, j. \quad (5)$$

Рассмотрим две оценочные задачи.

Задача 1. Определить значения $\{x_{ij}\}$, удовлетворяющие ограничениям (1) и (3) и минимизирующие критерий $P(x) = \sum_{ij} p_{ij} x_{ij}$.

Задача 2. Определить значения $\{x_{ij}\}$, удовлетворяющие ограничениям (2) и (3) и минимизирующие критерий $Q(x) = \sum_{ij} q_{ij} x_{ij}$.

Обозначим $L_1(p)$ ($L_2(q)$) значение целевой функции в оптимальном решении первой (второй) задачи, $p = \{p_{ij}\}$, $q = \{q_{ij}\}$.

В соответствии с основной теоремой [1], сумма

$$L_1(p) + L_2(q) = L(p, q) \quad (5)$$

дает оценку снизу для исходной задачи коммивояжера. Таким образом, двойственная задача заключается в определении значений p и q , удовлетворяющих соотношению (4) и максимизирующим сумму (5).

2. РЕШЕНИЕ ОЦЕНОЧНЫХ ЗАДАЧ ДЛЯ СИММЕТРИЧНОЙ ЗАДАЧИ КОММИВОЯЖЕРА

Метод решения оценочных задач для симметричных матриц основан на понятии кратчайшего i -дерева [2].

Выбираем некоторую вершину i . Для подграфа без вершины i построим кратчайшее i -дерево T_i , т. е. дерево кратчайшей длины $l(T_i)$. Для этого пос-

	1	2	3	4	5	6
1	—	2	2	3	3	3
2	2	—	1	1	1	3
3	2	1	—	3	3	3
4	3	1	3	—	3	3
5	3	1	3	3	—	1
6	3	3	3	3	1	—

Рис. 1. Матрица расстояний

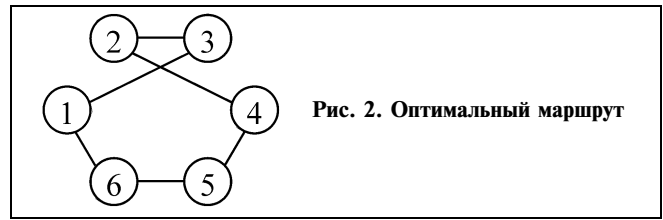


Рис. 2. Оптимальный маршрут

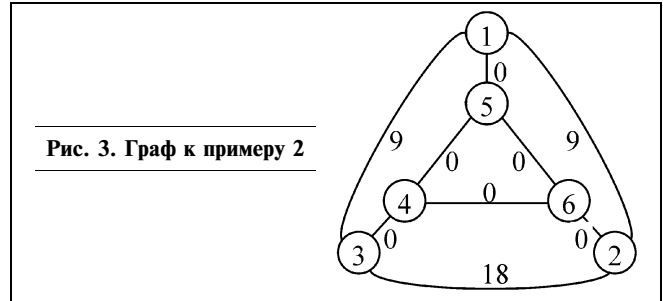


Рис. 3. Граф к примеру 2

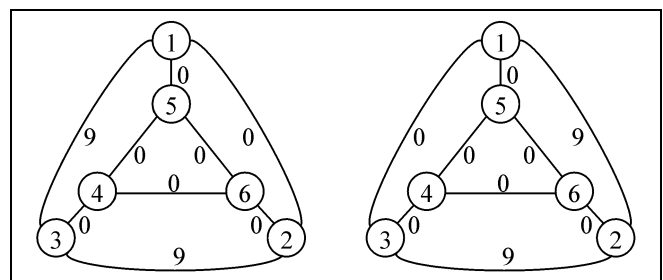


Рис. 4. Разбиение графа (см. рис. 3) на подграфы

троим кратчайшее дерево на вершинах графа без вершины i и добавим к нему два кратчайших ребра, инцидентных вершине i [2].

Оценка снизу для каждой задачи определяется i -деревом, для которого длина $l(T_i)$ максимальна.

Пример 1. Рассмотрим симметричную задачу из работы [2, § 3.5]. Матрица расстояний приведена на рис. 1.

В работе [2] оценка снизу определялась на основе 1-дерева, для которого $l(T_1) = 8$. Однако для 2-дерева $l(T_2) = 11$, что является достижимой оценкой и, более того, определяющей оптимальный маршрут, представленный на рис. 2.

В данном случае оптимальное решение двойственной задачи $P = L = \{l_{ij}\}$, $Q = 0$. Однако это не всегда так, что показывает следующий пример. ♦

Пример 2. Рассмотрим граф, рис. 3. Оценка снизу определяется любым i -деревом, и $l(T_i) = 9$, $i = \overline{1, 6}$. Разобьем граф на 2 подграфа (рис. 4). Имеем для первого графа $l(T_3) = 9$, а для второго $l(T_2) = 9$. Оценка снизу равна $l(T_3) + l(T_2) = 18$, т. е. в два раза больше. Более того, она достижима. Оптимальный маршрут — (1, 2, 6, 5, 4, 3, 1). ♦

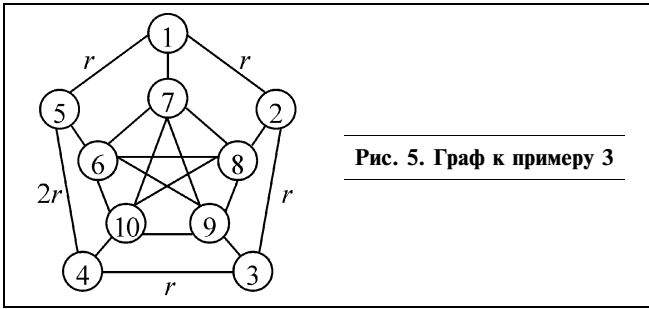


Рис. 5. Граф к примеру 3

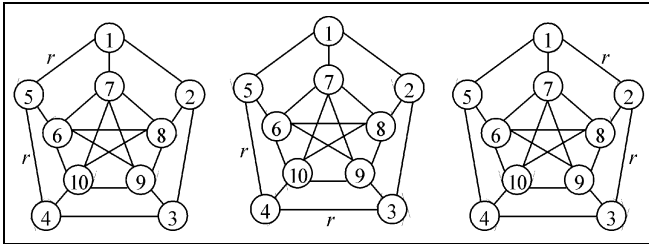


Рис. 6. Разбиение графа (см. рис. 5) на подграфы

Возникает вопрос, можно ли улучшить оценку, разбивая граф на более, чем 2 графа? Рассмотрим k матриц $P^s, s = 1, \dots, k$ таких, что

$$\sum_s p_{ij}^s = l_{ij}, \quad p_{ij} \geq 0, \quad \forall i, j. \quad (6)$$

Обозначим $\Phi(P^s)$ — значение целевой функции в решении s -й оценочной задачи. Очевидно, что сумма

$$\Phi(P) = \sum_s \Phi(P^s) \quad (7)$$

определяет оценку снизу для исходной задачи. Двойственная задача заключается в определении матриц $P^s, s = \overline{1, k}$, максимизирующих сумму (7) при условии (6). ♦

Пример 3. Рассмотрим граф, рис. 5 (длины всех дуг, кроме указанных на рисунке, равны нулю, $r > 0$).

Разобьем граф на $k = 3$ графа, рис. 6.

Имеем: $\Phi(P^1) = \Phi(P^2) = \Phi(P^3) = r$, оценка снизу $\Phi(P) = 3r$, и эта оценка достижима.

Оценка снизу, определяемая i -деревом исходной задачи равна r , а оценка, полученная в результате разбиения на 2 графа, равна $2r$. ♦

3. ПОЛУЧЕНИЕ ОЦЕНОК НА ОСНОВЕ ДЕРЕВА КРАТЧАЙШИХ ПУТЕЙ

Рассмотрим новый способ получения нижних оценок для оценочных задач, в основе которого лежит построение дерева кратчайших путей.

Определение 1. i -деревом кратчайших путей R_i называется дерево кратчайших путей из вершины i в остальные вершины графа. ♦

Сначала рассмотрим задачу коммивояжера с длинами дуг

$$l_{ij} = |\lambda_j - \lambda_i|, \quad \forall i, j, \quad \lambda_i \geq 0 \quad \forall i, \quad \lambda_1 = \min_i \lambda_i = 0.$$

Определение 2. Последовательность (i_1, i_2, \dots, i_n) называется *однопиковой*, если $\lambda_{i_1} \leq \lambda_{i_2} \leq \dots \leq \lambda_{i_{k-1}} \geq \lambda_{i_k} \geq \lambda_{i_{k+1}} \geq \dots \geq \lambda_{i_n}$, где $\lambda_{i_1} = 0, \lambda_{i_{k-1}} = \max_i \lambda_i = \lambda_{\max}$. ♦

Теорема. Любая однопиковая последовательность определяет оптимальный маршрут, длина которого равна $2\lambda_{\max}$.

Доказательство. Пусть вершины пронумерованы по возрастанию λ_i ($\lambda_1 = 0$). Длина пути от вершины 1 до вершины n не может быть меньше, чем λ_n , и равна λ_n при прохождении вершин по возрастанию номеров. Длина обратного пути из вершины λ_n в вершину λ_1 также не может быть меньше, чем λ_n , и равна λ_n при прохождении вершин по убыванию номеров. Поэтому любой однопиковый маршрут оптимален, и его длина равна $2\lambda_n$. Теорема доказана. ♦

Для получения оценок задач 1 и 2 определим $\lambda_i, i = \overline{1, n}, \min_i \lambda_i = 0$, так, чтобы $|\lambda_i - \lambda_j| \leq l_{ij}$ и значение $\lambda_m = \max_i \lambda_i$ было максимально.

Для построения i -дерева полагаем $\lambda_i = 0$ и строим дерево кратчайших путей из вершины i в остальные вершины. Обозначим λ_j длину кратчайшего пути из вершины i в вершину $j, \lambda(R_i) = \max_{j \neq i} \lambda_j$.

Согласно теореме, величина $2\lambda(R_i)$ является нижней оценкой для соответствующей задачи коммивояжера. Максимальная оценка, очевидно, определяется i -деревом кратчайших путей, для которого $\lambda(R_i) = \max_j \lambda(R_j)$.

Пример 4. Рассмотрим граф, рис. 7.

Имеем $\lambda(R_1) = 5, \lambda(R_2) = 7, \lambda(R_3) = 5, \lambda(R_4) = 7, \max_i \lambda(R_i) = 7$. Следовательно, оценка снизу равна 14,

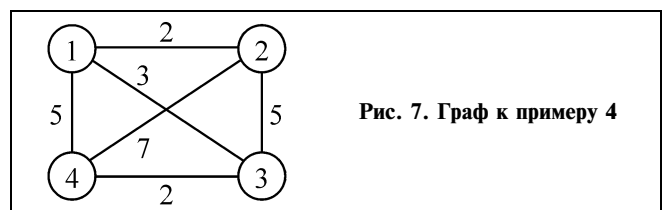


Рис. 7. Граф к примеру 4

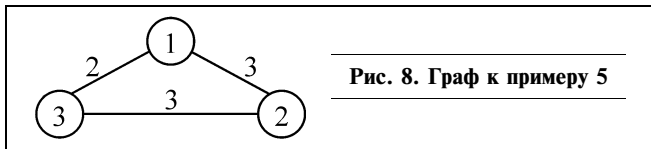


Рис. 8. Граф к примеру 5

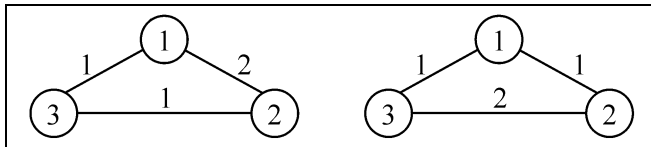


Рис. 9. Разбиение графа (см. рис. 8) на подграфы

она достижима. Оптимальный маршрут (1, 2, 3, 4, 1) или (1, 3, 4, 2, 1). Заметим, что оценка, получаемая на основе построения кратчайших i -деревьев хуже. Действительно, имеем $l(T_1) = 12$, $l(T_2) = 12$, $l(T_3) = 12$, $l(T_4) = 12$, и оценка равна $12 < 14$. ♦

Однако нетрудно привести примеры, когда оценка путем построения кратчайшего i -дерева лучше, чем оценка, получаемая путем построения i -дерева кратчайших путей.

Описанный метод получения нижних оценок естественно применить для оценки решений оценочных задач 1 и 2.

Пример 5. Рассмотрим граф, рис. 8.

Разобьем его на 2 графа, рис. 9.

Для первого графа $\lambda(R_2) = 2$, а для второго $\lambda(R_3) = 2$, и оценка равна 8. Эта оценка достижима.

Заметим, что для исходного графа $\lambda(R_1) = \lambda(R_2) = \lambda(R_3) = 3$, и оценка снизу равна 6.

ЗАКЛЮЧЕНИЕ

Рассмотренный подход к получению нижних оценок для симметричной задачи коммивояжера на основе двойственных задач дает новый метод получения оценок. Фактически мы имеем три типа двойственных задач. В задачах первого типа оценочные задачи решаются на основе построения кратчайших i -деревьев. В задачах второго типа оценочные задачи решаются на основе построения i -деревьев кратчайших путей. В задачах третьего типа одна из оценочных задач решается на основе построения кратчайших i -деревьев, а вторая — на основе построения i -деревьев кратчайших путей. Разработка методов решения двойственной задачи требует дальнейших исследований.

ЛИТЕРАТУРА

1. Бурков В.Н., Буркова И.В., Попок М.В. Метод сетевого программирования // Проблемы управления. — 2005. — № 3. — С. 23—29.
2. Сигал И.Х., Иванова А.П. Введение в прикладное дискретное программирование. — М.: Физматлит, 2007. — 304 с.
3. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. — М.: Мир. — 1982. — 416 с.

☎ (495) 334-79-00; e-mail: irbur27@mail.ru

Статья представлена к публикации членом редколлегии А.Д. Цвиркуном. □

Конференция «Технические и программные средства систем управления, контроля и измерения» (УКИ-08)

Институт проблем управления им. В.А. Трапезникова РАН
10—12 ноября 2008 г., г. Москва

Цели конференции:

- Обмен информацией о результатах современных исследований и разработок технических и программных средств систем управления, контроля и измерения (УКИ)
- Обсуждение, обобщение и распространение научных достижений в области создания и эксплуатации технических и программных средств систем УКИ
- Возобновление вынужденно прерванных контактов и связей между специалистами, работающими над проблемами теории и практики создания и применения технических и программных средств систем УКИ в институтах РАН, отраслевых институтах и научных центрах, вузах, промышленных предприятиях и коммерческих структурах
- Выявление перспективных направлений исследований и разработок технических и программных средств систем УКИ

Более подробную информацию можно найти на сайте <http://www.ipu-conf.ru/ktm>
e-mail: ktm-2008@mail.ru, тел. (495) 334-93-70, (495) 334-90-29, тел./факс (495) 334-93-61



О СПЕКТРЕ МАТРИЦЫ СМЕЖНОСТИ ОРГРАФА КОЛЬЦЕВОЙ СТРУКТУРЫ И ЕГО ПРИМЕНЕНИИ

Р.П. Агаев, С.В. Никифоров, Н.А. Андриюшина

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Установлено, что при удалении двух дуг спектр орграфа с кольцевой структурой, содержащего два гамильтоновых цикла, остается действительным тогда и только тогда, когда число вершин четно и «расстояние между дугами», удаленными из одного цикла, максимально. Рассмотрена возможность применения полученных результатов для оценки отказоустойчивости сетей кольцевой топологии.

ВВЕДЕНИЕ

Алгебраическая теория графов в последние годы все шире применяется в различных областях управления, социологии, комбинаторике, физике и др. Относительно недавно появился ряд работ (см. например, работу [1]), в которых изучается топология Интернета на уровне автономных систем для проектирования новых протоколов глобальной сети. Краткий обзор некоторых других работ, касающихся применения алгебраической теории графов в управлении и моделировании, был дан в статье [2].

Отметим, что в зависимости от поставленных задач в литературе изучается как спектр матрицы смежности, так и лапласовский спектр графа. Хотя монографии [3, 4] посвящены в основном спектру матрицы смежности, в них имеются разделы, относящиеся к лапласовскому спектру, характеризующему «древовидную» структуру графа.

В настоящей работе исследуется топология сети кольцевой структуры с помощью спектра матрицы смежности. Этот спектр характеризует циклическую структуру графа (см. например, главу 1 в книге [3]).

Мы рассматриваем орграф кольцевой структуры, содержащий два гамильтоновых цикла, в одном из которых удалена одна или две дуги. В статье [5] эта задача была частично решена для лапласовского спектра орграфа. Именно такой класс орграфов характеризует топологию сети кольцевой структуры. Для получения практических результатов применительно к сетям FDDI (Fiber Distributed Data Interface) необходимо получить ответы на следующие вопросы: будет ли спектр графа действительным при удалении одной дуги; когда при

удаленных из одного контура двух дугах спектр матрицы смежности остается действительным и как это связано с четностью числа узлов; как по спектральному радиусу определить число узлов в сегментированной сети?

1. ОСНОВНЫЕ ПОНЯТИЯ И ПРЕДВАРИТЕЛЬНЫЕ РЕЗУЛЬТАТЫ

Рассматривается ориентированный граф с множеством вершин $V = \{1, \dots, n\}$ и множеством дуг $E \subseteq V \times V$. Матрицу смежности соответствующего графа обозначим через A . Граф называют циркулянтным, если при определенной нумерации вершин его матрица смежности циркулянтная. Напомним, что матрица называется циркулянтной, если любая ее строка получается из предыдущей путем циклического сдвига на одну позицию вправо.

В данной работе изучаются циркулянтные графы, содержащие два «встречных» гамильтоновых цикла, в одном из которых удалена одна или две дуги.

Многочлен Чебышева второго рода $P_n(x)$, масштабированный на отрезке $[-2, 2]$, удовлетворяет рекуррентному соотношению: $P_n(x) = xP_{n-1}(x) - P_{n-2}(x)$ при $P_0(x) \equiv 1$ и $P_1(x) \equiv x$. Многочлены Чебышева часто представляют в тригонометрическом виде:

$$P_n(x) = \frac{\sin(n+1)\varphi}{\sin\varphi}, \quad (1)$$

где $x = 2\cos\varphi$, $\varphi \in (0, \pi)$.

В работе [5] (лемма 1) показано, что на интервале $(-2, 2)$ корни уравнения $P_n(x) + (-1)^p = 0$ имеют вид

$$2\cos\frac{\pi k}{n+1+(-1)^{k+p}}, \quad (2)$$

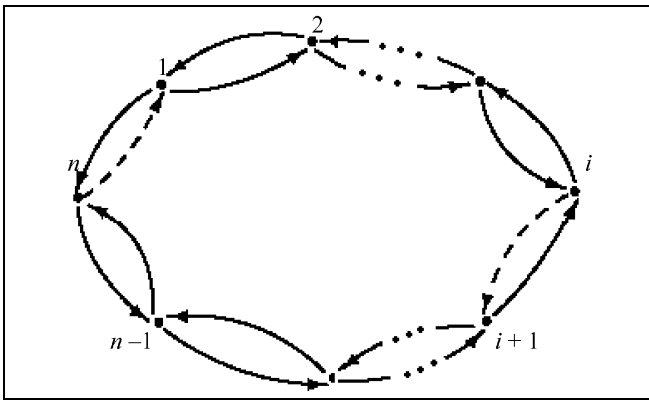


Рис. 1. Орграф, содержащий два гамильтоновых цикла, в одном из которых удалены две дуги

где $p \in \{1, 2\}$, $k = 1, \dots, n$ для нечетного n и $k = 1, \dots, n - 1, n + 1$ для четного n .

Утверждение. Если из циркулянтного графа с двумя гамильтоновыми циклами удалить одну дугу, то характеристический многочлен его матрицы смежности A

$$P_A(x) = P_n(x) - 1. \quad \blacklozenge \quad (3)$$

Многочлен (3) легко можно получить разложением определителя характеристической матрицы соответствующего графа по строке, соответствующей вершине, к которой смежна одна вершина. С другой стороны, спектр такого графа действителен и определяется формулой (2) при $p = 1$.

Рассмотрим орграф, в котором штриховые дуги удалены (рис. 1).

Обозначим через L_2 и A соответственно лапласовскую матрицу и матрицу смежности этого орграфа. Определим многочлен¹ $Z_n(x)$ следующим рекуррентным соотношением: $Z_n(x) = (x - 2)Z_{n-1}(x) - Z_{n-2}(x)$ при $Z_0(x) \equiv 1$ и $Z_1(x) \equiv x - 1$.

Задача о действительности спектра матрицы L_2 была рассмотрена в работе [5], где доказана следующая теорема.

Теорема 1 [5]. 1. Характеристический многочлен Δ_{L_2} матрицы L_2 равен $Z_i(\lambda)Z_{n-i}(\lambda) - (-1)^n$.

2. Если n четно, то все собственные значения матрицы L_2 действительны тогда и только тогда, когда $2i = n$. В этом случае собственные значения равны $4\cos^2 \frac{\pi k}{n}$, $4\cos^2 \frac{\pi k}{n+2}$, $k = 1, \dots, \frac{n}{2}$.

3. Если n нечетно, то все собственные значения матрицы L_2 действительны тогда и только тогда,

когда $i = \frac{n-1}{2}$ или $i = \frac{n+1}{2}$, и в этих случаях равны нулю (с кратностью 1) и $4\cos^2 \frac{\pi k}{n+1}$, $k = 1, \dots, \frac{n-1}{2}$ (все с кратностью 2). \blacklozenge

Доказательство теоремы основано на свойствах произведений многочленов Чебышева второго рода $P_i(x)P_j(x)$ при четных i и j .

Из теоремы 1 следует, что при удалении двух дуг e_1 и e_2 из одного гамильтонова цикла все собственные значения лапласовской матрицы останутся действительными тогда и только тогда, когда расстояние между удаленными дугами максимально. В случае четного n удаление дуг e_1 , e_2 и симметричных к ним дуг разбивает граф на две цепи $P_{n/2}$ с непересекающимися множествами вершин. Для случая нечетного n удаление этих дуг приводит к двум цепям $P_{(n-1)/2}$ и $P_{(n+1)/2}$.

Иначе обстоит дело с матрицей смежности. Если задача действительности лапласовского спектра в работе [5] связана со свойствами произведений $Z_i(x)Z_j(x)$ то решение аналогичной задачи для матрицы смежности связано со свойствами произведений $P_i(x)P_j(x)$. В силу соотношения $P_{2n}(x) = Z_n(x^2)$ в случае лапласовского спектра рассматриваются произведения $P_k(x)P_l(x)$ только для четных k и l , а в случае спектра матрицы смежности рассматриваются такие произведения при всевозможных k и l , что усложняет решение задачи.

2. ОСНОВНОЙ РЕЗУЛЬТАТ

Рассмотрим характеристическую матрицу смежности $\lambda I - A$ для орграфа, приведенного на рис. 1:

$$\lambda I - A = \begin{pmatrix} \lambda & -1 & & & & & & & -1 \\ -1 & \lambda & & & & & & & \\ & & \dots & & & & & & \\ & & & \lambda & 0 & & & & \\ & & & -1 & \lambda & & & & \\ & & (0) & & & \dots & & & \\ & & & & & & \lambda & -1 & \\ & & & & & & -1 & \lambda & n \end{pmatrix} \begin{matrix} 1 \\ 2 \\ \dots \\ i \\ i+1 \\ \dots \\ n-1 \\ n \end{matrix}$$

Нетрудно вычислить характеристический многочлен $C_A(\lambda)$ этой матрицы:

$$\begin{aligned} C_A(\lambda) &= \lambda(P_{i-1}P_{n-i}) + (-1)^3(-1)^2(-1)P_{i-2}P_{n-i} + \\ &+ (-1)^n(-1)(-1)^{n-1} = \lambda(P_{i-1}P_{n-i}) + \\ &+ P_{i-2}P_{n-i} - 1 = P_{n-i}(\lambda P_{i-1} - P_{i-2}) - 1 = \\ &= P_i P_{n-i} - 1, \end{aligned} \quad (4)$$

где P_k — многочлен Чебышева второго рода, масштабированный на отрезке $[-2, 2]$.

¹ В работе [5] (предложение 1) доказано, что $P_{2n}(x) = Z_n(x^2)$.



Из формулы (1) непосредственно следует, что корни² многочлена $P_n(x)$ на отрезке $[-2, 2]$

$$x_k = 2 \cos \frac{\pi k}{n+1}, \quad k = 1, \dots, n.$$

Рассмотрим задачу действительности спектра матрицы A , т. е. определим значения i , при которых уравнение $P_i(x)P_{n-i}(x) - 1 = 0$ при фиксированном n имеет только действительные корни.

Теорема 2. *Спектр матрицы A действителен тогда и только тогда, когда n четно и $i = n/2$.*

Доказательство см. в Приложении.

Если число узлов четно, то добавлением одного фиктивного узла можно получить нечетное число в качестве порядка матрицы смежности. В этом случае при удалении двух дуг спектр всегда будет содержать комплексные собственные значения.

3. ОЦЕНКА ОТКАЗОУСТОЙЧИВОСТИ СЕТЕЙ КОЛЬЦЕВОЙ ТОПОЛОГИИ

В § 2 мы установили, что при удалении двух дуг из одного гамильтонова цикла собственные значения соответствующей матрицы смежности действительны только тогда, когда удалены «диаметрально противоположные» дуги. Это свойство рассматриваемого орграфа может быть использовано для оценки отказоустойчивости сетей кольцевой топологии FDDI со скоростью передачи данных до 100 Мбит/с. Напомним, что сеть FDDI строится на основе двух оптоволоконных колец, которые образуют основной и резервный пути передачи данных между узлами сети (рис. 2).

Наличие двух колец дает возможность повышения отказоустойчивости в сети FDDI. В нормальном режиме данные проходят через все узлы и все участки кабеля только первичного кольца (направление — против часовой стрелки). Вторичное кольцо (направление — по часовой стрелке) при этом не используется. В случае отказа станции или же обрыва кабеля первичное кольцо объединяется со вторичным с помощью концентраторов и (или) сетевых адаптеров и образует единое кольцо. В этом случае сеть будет состоять из одного кольца, и следующий обрыв (или же отказ еще одного узла) приведет к сегментированию сети на два не сообщающихся кольца. При этом все станции и концентраторы, входящие в сеть, должны иметь двойное подключение DAS. Такие станции оснащены оптическим обходным переключателем.

В сети FDDI за мониторинг работы кольца и станций, обнаружение ошибок и их исправление отвечает протокол SMT (Station Management). Он

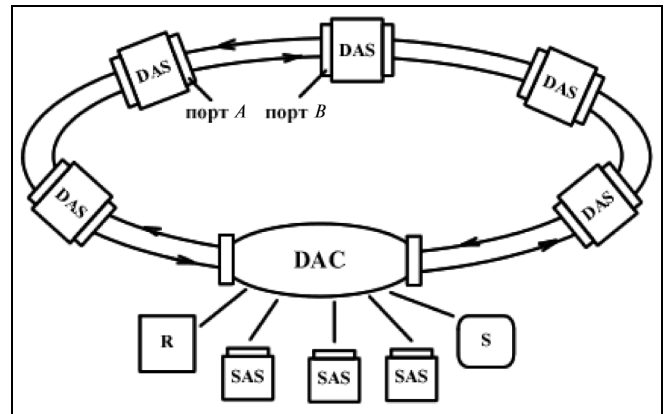


Рис. 2. Вариант топологии сети FDDI:

DAS — станция двойного подключения (Dual Attachment Station); SAS — станция одиночного подключения (Single Attachment Station); DAC — концентратор двойного подключения (Dual Attachment Concentrator); R — маршрутизатор (Router); S — сервер

также управляет другими уровнями: физическим уровнем стека³ протоколов FDDI — для устранения отказа сети из-за обрыва кабеля; уровнем управления доступом к среде MAC (Media Access Control) — для обнаружения логических ошибок (см., например, [6]).

Обычно в сетях FDDI, когда речь идет об обрыве линии связи, имеется в виду обрыв обоих колец. Но часто при прокладке оптического кабеля для большей надежности стараются разнести линии колец на некоторое расстояние. Кроме того, обрыв одной линии может быть обусловлен малым радиусом изгиба или же неисправностью оптического переключателя. В этом случае при обрыве кабеля одного кольца второе может взять на себя все функции первого. Внедрение алгоритма вычисления спектра топологии на серверной станции (для надежности такие серверы подключаются к двум концентраторам) позволяет оперативно отслеживать изменения топологии всей сети. Поскольку спектр топологии определяется с точностью до расстояния между неисправными участками, при обрыве двух связей можно оценить реконфигурированные сегменты сети.

Рассмотрим некоторые возможные неисправности сети с числом обрывов не больше двух, которые определяются спектром соответствующей матрицы смежности (доказательства некоторых утверждений в силу их простоты опущены):

— если 2 не является собственным значением, то это означает, что имеется хотя бы один обрыв линии связи;

² На самом деле, все корни многочлена $P_n(x)$ принадлежат отрезку $[-2, 2]$.

³ Стек протоколов — это иерархически организованный набор протоколов для организации взаимодействия узлов в сети.

— при обрыве двух линий связи чем меньше значение спектрального радиуса графа, тем ближе друг к другу размеры полученных сегментов, что соответствует наилучшему случаю неисправности;

— при фиксированном значении спектрального радиуса действительность спектра означает распад сети на два сегмента с одинаковым числом узлов;

— если спектр содержит комплексные собственные значения, то, как следует из теоремы 2, одно кольцо исправно (иначе не было бы комплексных корней), а неисправности приходится только на второе кольцо, и минимальное число обрывов равно двум.

ЗАКЛЮЧЕНИЕ

В работе было исследовано появление комплексных собственных значений в спектре матрицы смежности орграфа кольцевой структуры, содержащего два гамильтоновых цикла, в одном из которых удалены одна или две дуги. Полученные результаты позволяют оценить отказоустойчивость сетей кольцевой топологии FDDI. Предложенный подход особенно эффективен в том случае, когда узлы не пронумерованы в «удобном порядке» или же этот порядок изменяется в связи с реконfigurацией сети.

Из теоремы 2 следует, что при обрыве двух линий связи среди собственных значений матрицы смежности всегда присутствует пара комплексных чисел, за исключением случая, когда удалены «диаметрально противоположные» дуги. Как было отмечено в конце § 2, добавлением одного гипотетического узла можно получить нечетное число в качестве порядка матрицы смежности. В этом случае при обрыве двух линий связи спектр всегда будет содержать комплексные собственные значения. Однако остается вопрос: содержит ли спектр комплексные числа при удалении более двух линий связи? Можно легко проверить, что при удалении больше двух дуг, идущих подряд, спектр будет содержать комплексные числа. Отметим, что эта задача тесно связана с решением уравнения $P_a P_b \dots P_z = 1$, где $a + b + \dots + z = n$.

Орграф, рассмотренный в настоящей работе, получен из циркулянтного орграфа удалением двух дуг. В работах [7–9] был рассмотрен специальный класс циркулянтных графов, из которых удалены некоторые подмножества дуг, и для них подсчитаны числа остовных деревьев. Однако указанные работы посвящены лапласовскому спектру графа и его «древовидной» структуре. Исследование аналогичной задачи для описания циклических свойств представляется интересным и должно стать предметом дальнейших исследований.

ПРИЛОЖЕНИЕ

Доказательство теоремы 2. Если $2i = n$, то все корни характеристического многочлена (4) $P_i^2 - 1 = (P_i - 1)(P_i + 1)$ действительны и определяются выражением (2). Докажем, что во всех остальных случаях среди характеристических чисел есть хотя бы одна пара комплексных. Рассмотрим уравнение $P_i P_j = 1$ и предположим, что $i > j$. Для всевозможных значений i и j составим таблицу комбинаций (см. ниже).

Рассмотрим случай a из таблицы, представленный 10-ю вариантами. Обозначим через r четное число из i, j , если $i + j$ нечетно. В случае четности обоих этих чисел через r обозначим меньшее из них.

Докажем, что на интервале $(0; 2 \cos \frac{\pi r}{2(r+1)})$ значение $P_r P_q$, где $q \in \{i, j\}$, меньше единицы.

Положим $\varphi_\delta = \frac{\pi(r+\delta)}{2(r+1)}$, где $0 < \delta < 1$. Далее, учитывая представление (1),

$$|P_r P_q| \leq \frac{\left| \sin \left((r+1) \frac{\pi(r+\delta)}{2(r+1)} \right) \right|}{\sin^2 \frac{\pi(r+\delta)}{2(r+1)}} = \frac{\left| \sin \frac{\pi(r+\delta)}{2} \right|}{\sin^2 \frac{\pi(r+\delta)}{2(r+1)}} = \frac{\left| \sin \frac{\pi\delta}{2} \right|}{\sin^2 \frac{\pi(r+\delta)}{2(r+1)}}.$$

Рассмотрим производную

$$\begin{aligned} & \left(\frac{\sin \frac{\pi\delta}{2}}{\sin^2 \frac{\pi(r+\delta)}{2(r+1)}} \right)' = \\ & = \frac{\frac{\pi}{2} \cos \frac{\pi\delta}{2} \sin^2 \frac{\pi(r+\delta)}{2(r+1)} - \frac{\pi}{r+1} \sin \frac{\pi(r+\delta)}{2(r+1)} \cos \frac{\pi(r+\delta)}{2(r+1)} \sin \frac{\pi\delta}{2}}{\sin^4 \frac{\pi(r+\delta)}{2(r+1)}} = \\ & = \frac{\frac{\pi}{2} \cos \frac{\pi\delta}{2} \sin \frac{\pi(r+\delta)}{2(r+1)} - \frac{\pi}{r+1} \sin \frac{\pi\delta}{2} \cos \frac{\pi(r+\delta)}{2(r+1)}}{\sin^3 \frac{\pi(r+\delta)}{2(r+1)}} = \\ & = \frac{\left(\frac{\pi(r-1)}{2(r+1)} + \frac{\pi}{r+1} \right) \cos \frac{\pi\delta}{2} \sin \frac{\pi(r+\delta)}{2(r+1)} - \frac{\pi}{r+1} \sin \frac{\pi\delta}{2} \cos \frac{\pi(r+\delta)}{2(r+1)}}{\sin^3 \frac{\pi(r+\delta)}{2(r+1)}} = \\ & = \frac{\frac{\pi(r-1)}{2(r+1)} \cos \frac{\pi\delta}{2} \sin \frac{\pi(r+\delta)}{2(r+1)} + \frac{\pi}{r+1} \sin \frac{\pi\delta}{2} \frac{r(1-\delta)}{r+1}}{\sin^3 \frac{\pi(r+\delta)}{2(r+1)}}. \end{aligned}$$

Заметим, что числитель и знаменатель последней дроби положительны при $0 < \delta < 1$, т. е. на интервале $(0; 2 \cos \frac{\pi r}{2(r+1)})$

функция $\frac{\sin \frac{\pi\delta}{2}}{\sin^2 \frac{\pi(r+\delta)}{2(r+1)}}$ возрастает по δ .

Комбинации всевозможных значений i и j

i	$4q$	$4q + 1$	$4q + 2$	$4q + 3$
$4p$	b	a	a	a
$4p + 1$	a	c	a	d
$4p + 2$	a	a	b	a
$4p + 3$	a	d	a	c



Поскольку при $\delta = 0$ и $\delta = 1$ значения этого выражения равны соответственно 0 и 1, получаем, что на интервале

$$\left(0; 2 \cos \frac{\pi r}{2(r+1)}\right) \text{ справедливо неравенство } \frac{\sin \frac{\pi \delta}{2}}{\sin^2 \frac{\pi(r+\delta)}{2(r+1)}} < 1.$$

В случае *b* среди характеристических чисел есть хотя бы одна пара комплексных чисел, следует из предложения 1 см. работу [5], согласно которому $P_i(x)P_j(x) = Z_{i/2}(x^2)Z_{j/2}(x^2)$.

Поскольку $(i+j)/2$ — четное число и $i > j$, согласно п. 1 леммы 4 из работы [5] уравнение $P_i(x)P_j(x) = 1$ имеет комплексные корни.

В случае *c*. Положим $u_1^{(i)} = 2 \cos \frac{\pi(i+1)}{2(i+1)}$, $u_2^{(i)} = 2 \cos \frac{\pi(i-1)}{2i}$ — корни уравнения $P_i - 1 = 0$.

Пусть $x_1^{(i)}, x_2^{(i)} \geq x_1^{(j)}$ — наименьшие корни уравнения $P_i = 0$; положим $\varphi_{u_1^{(i)}} = \frac{\pi(i+1)}{2(i+1)}$, $\varphi_{u_2^{(i)}} = \frac{\pi(i-1)}{2i}$, $\varphi_\delta = \frac{\pi(i+1-\delta)}{2(i+2)}$, где $0 < \delta < 2/i$, $\varphi_0 = \varphi_{u_1}$, $\varphi_{2/i} = \varphi_{u_2}$. Тогда

$$\begin{aligned} \left| \frac{\sin(i+1)\varphi_\delta \sin(j+1)\varphi_\delta}{\sin^2 \varphi_\delta} \right| &\leq \frac{\left| \sin\left(\frac{\pi}{2}(j+1)\frac{i+1-\delta}{i+2}\right) \right|}{\sin^2 \frac{\pi i+1-\delta}{2i+2}} = \\ &= \frac{\left| \sin\left(\frac{\pi}{2}(j+1)\left(1-\frac{1+\phi}{i+2}\right)\right) \right|}{\sin^2 \frac{\pi}{2}\left(1-\frac{1+\phi}{i+2}\right)} = \frac{\left| \sin\left(\frac{\pi}{2}(j+1)\frac{1+\delta}{i+2}\right) \right|}{\sin^2 \frac{\pi 1+\delta}{2i+2}} \leq \frac{\left| \sin \frac{\pi j+1}{2i} \right|}{\cos^2 \frac{\pi}{2i}} \leq \\ &\leq \frac{\left| \sin \frac{\pi}{2} \frac{i-2}{i} \right|}{\cos^2 \frac{\pi}{2i}} = \frac{\cos \frac{2\pi}{2i}}{\cos^2 \frac{\pi}{2i}} = 1 - \operatorname{tg}^2 \frac{\pi}{2i} < 1 \end{aligned}$$

Заметим, что в случае *c*, если $u_1^{(j)} \geq x_2^{(i)}$, то $P_i(x)P_j(x) \leq 1$ на интервале $(0, x_2^{(i)})$.

Предположим, что $u_1^{(j)} < x_2^{(i)}$. Тогда из $\varphi_{u_1^{(j)}} = \frac{\pi(j+1)}{2(j+2)}$, $\varphi_{x_2^{(i)}} = \frac{\pi(i-1)}{2(i+1)}$ и $\varphi_{u_2^{(j)}} > \varphi_{x_2^{(i)}}$ получим $1 < \frac{i+1}{j+2} < 2$ и $j > \frac{i-3}{2}$.

Пусть $\varphi_\delta = \frac{\pi(i-1+\delta)}{2(i+1)}$. Тогда при $\delta = \frac{2j-i+3}{j+2} = \frac{2j+4}{j+2} - \frac{i+1}{j+2} = 2 - \frac{i+1}{j+2} < 1$ получим $\varphi_0 = \varphi_{x_2^{(i)}}$ и $\varphi_\delta = \varphi_{u_1^{(j)}}$.

На интервале $(0, 2 \cos \frac{\pi(i+1)}{2(i-1)})$ имеет место соотношение

$$\begin{aligned} \frac{|\sin(i+1)\varphi_\delta|}{\sin^2 \varphi_\delta} &= \frac{\left| \sin \frac{\pi}{2}(i-1+\delta) \right|}{\sin^2 \frac{\pi(i-1+\delta)}{2(i+1)}} = \frac{\left| \sin \frac{\pi \delta}{2} \right|}{\cos^2 \frac{\pi(2-\delta)}{2(i+1)}} \leq \\ &\leq \frac{\left| \sin \frac{\pi}{2}\left(2-\frac{i+1}{j+2}\right) \right|}{\cos^2 \frac{\pi}{i+1}} = \frac{\left| \sin \frac{\pi(i+1)}{2(j+2)} \right|}{\cos^2 \frac{\pi}{i+1}}. \end{aligned}$$

Заменим величину $\frac{i+1}{j+2}$ на заведомо меньшую

$$\frac{i+1-(2j-i+3)/2}{j+2-(2j-i+3)/2} = \frac{3i-2j-1}{i+1} = \frac{(i+1)+2(i-j)-2}{i+1} > 1.$$

Поскольку $(2j-i+3)/2 > 0$, выполняется неравенство $\frac{i+1}{j+2} > \frac{3i-2j-1}{i+1} > 1$.

Итак, получим

$$\begin{aligned} \frac{\sin \frac{\pi(i+1)}{2(j+2)}}{\cos^2 \frac{\pi}{i+1}} &\leq \frac{\sin \frac{\pi(3i-2j-1)}{2(i+1)}}{\cos^2 \frac{\pi}{i+1}} = \frac{\sin \frac{\pi(i+1+2(i-j)-2)}{2(i+1)}}{\cos^2 \frac{\pi}{i+1}} = \\ &= \frac{\cos \frac{\pi(i-j-1)}{i+1}}{\cos^2 \frac{\pi}{i+1}} < \frac{\cos \frac{2\pi}{i+1}}{\cos^2 \frac{\pi}{i+1}} = 1 - \operatorname{tg}^2 \frac{\pi}{i+1} < 1. \end{aligned}$$

Наконец, рассмотрим случай *d*. Представим многочлены $P_i(x)$ и $P_j(x)$ следующим образом:

$$P_k(x) = x^2 \varphi(x) + (-1)^{(k-1)/2} \left(\frac{k+1}{2}\right)x.$$

Тогда для произведения $P_i(x)P_j(x)$ в обоих вариантах случая *d* из таблицы получим

$$\begin{aligned} P_i(x)P_j(x) &= x^4 \psi(x) + (-1)^{4u/2} (-1)^{(4v+2)/2} \left(\frac{4u+2}{2}\right) \left(\frac{4v+4}{2}\right) x^2 = \\ &= x^4 \psi(x) + (-1)^{(4u+v+2)/2} \left(\frac{4u+2}{2}\right) \left(\frac{4v+4}{2}\right) x^2 = \\ &= x^4 \psi(x) + (-1)^{2(u+v)+1} (4u+2)(v+1)x^2, \end{aligned}$$

где u и v соответствуют значениям из таблицы. Поскольку первая производная последнего выражения в точке $x = 0$ равна нулю, а вторая производная отрицательна⁴, многочлен $P_i(x)P_j(x)$ в этой точке имеет локальный максимум, и уравнение $P_{n-1}(x)P_n(x) = 1$ имеет хотя бы одну пару комплексных корней, если $n \neq 2i$. Теорема доказана. ♦

ЛИТЕРАТУРА

1. *Vukadinovic D., Huang P., Erlebach Th.* On the Spectrum and Structure of Internet Topology Graphs // Innovative Internet Computing Systems / Second International Workshop on Innovative Internet Computing Systems (IICS). Lecture Notes in Computer Science. — Berlin, 2002. — Vol. 2346. — P. 83–95.
2. *Агаев Р.П., Чеботарев П.Ю.* Лапласовские спектры оргграфов и их приложения // Автоматика и телемеханика. — 2005. — № 5. — С. 47–62.
3. *Biggs N.* Algebraic Graph Theory. — Cambridge: Cambridge University Press, 1993. — 205 с.
4. *Цветкович Д., Дуб М., Захс Х.* Спектры графов: теория и применение. — Киев: Наукова думка, 1984. — 438 с.
5. *Агаев Р.П.* Об исследовании и применении лапласовских спектров оргграфов кольцевой структуры // Автоматика и телемеханика. — 2008. — № 2. — С. 19–25.
6. *Олифер В.Г., Олифер Н.А.* Компьютерные сети. — СПб.: Питер, 2006. — 958 с.
7. *Boesch F.T., Prodinger H.* Spanning tree formulas and Chebyshev polynomials // Graphs and Combinatorics. — 1986. — Vol. 2. — P. 191–200.
8. *Zhang Y., Yong X., Golin M.* The number of spanning trees in circulant graphs // Discrete Mathematics. — 2000. — Vol. 223. — P. 337–350.
9. *Zhang Y., Yong X., Golin M.* Chebyshev polynomials and spanning tree formulas for circulant and related graphs // Discrete Mathematics. — 2005. — Vol. 298. — P. 334–364.

☎ 334-88-69, mail: arpoje@rambler.ru, svrguiat@rambler.ru
Статья представлена к публикации членом редколлегии А.Т. Алескеровым. □

⁴ На интервале, содержащем точку x , график многочлена $P_i(x)P_j(x)$ является выпуклым вверх.

НЕДИФФЕРЕНЦИРУЕМЫЕ ОВРАЖНЫЕ ДВУМЕРНЫЕ ТЕСТ-ФУНКЦИИ¹

А.С. Рыков, М.Ю. Матвиенко

Государственный технологический университет «Московский институт стали и сплавов»

Приведено построение двумерных овражных тест-функций для испытания методов оптимизации, представляющих собой новый вид недифференцируемых тест-функций; к их особенностям относится наличие дна оврага с острым кусочно-линейным дном и кусочно-линейными склонами. Описан алгоритм конструирования овражных тест-функций с требуемыми свойствами.

ВВЕДЕНИЕ

Эффективность оптимизационных методов может изучаться как теоретически, так и экспериментально. Первый подход может применяться только для ограниченного числа задач, в то же время весьма удобно сравнивать методы с помощью тест-функций. Цель вычислительного эксперимента заключается в исследовании работоспособности метода, в испытании различных оптимизационных методов в одинаковых условиях и получении их сравнительных характеристик.

Отметим, что один метод может иметь несколько разных алгоритмических реализаций. Рассматривая алгоритмы, мы будем подразумевать их конкретные реализации.

Важен выбор критериев для сравнения алгоритмов. Обычно, когда в литературе представляется новый метод, некоторые из его свойств (например, число вычислений целевой функции, время вычислений, точность определения минимума) приводятся для анализа. Сравнение оптимизационных методов представляет собой многокритериальную проблему. Например, для сравнения алгоритмов и программ можно применять следующие критерии: 1) процессорное время; 2) точность решения; 3) число итераций; 4) робастность; 5) число вычислений целевой функции; 6) удобство применения; 7) требования к памяти; 8) число основных операций.

¹ Работа выполнена при финансовой поддержке РФФИ (гранты 07-07-00151 и 08-07-00055) и доложена на VII международной конференции «Идентификация систем и задачи управления» SICPRO'08, Москва, Институт проблем управления РАН, 28–31 января 2008 г.

Критерии 1, 3, 5 и 8 взаимосвязаны и могут применяться для оценки эффективности алгоритма.

В работах, представляющих новые алгоритмы, обычно подчеркивается их превосходство над известными методами. Иногда предпринимаются попытки сравнить алгоритмы одного или разных классов. Сравнение алгоритмов на основе вычислительного эксперимента при минимизации тест-функций может носить объективный характер. В этом случае важен выбор тест-функций, их свойства.

Существует множество различных тест-функций, которые широко используются для оценки вычислительных свойств градиентных алгоритмов оптимизации и методов прямого поиска [1–7].

Функция Розенброка $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ одна из самых известных и часто применяемых. Она имеет крутой параболический овраг вдоль кривой $x_2 = x_1^2$.

По нашему мнению, вытекающему из опыта решения оптимизационных задач, наибольшие вычислительные трудности вызывают овраги и хребты на поверхности минимизируемых функций. Среди овражных функций функции с недифференцируемым дном оврага, вероятно, наиболее трудные для минимизации, поскольку возникают дополнительные сложности в точках, где функция недифференцируема. На скорость сходимости методов влияет также крутизна склонов оврага.

Для минимизации недифференцируемых функций можно применять методы прямого поиска, поскольку они не используют производные [8]. Для исследования свойств методов прямого поиска важно иметь различные недифференцируемые овражные функции [8, 9].



В статье приведено построение овражных двумерных кусочно-линейных тест-функций для испытания методов оптимизации, представляющих собой новый вид недифференцируемых тест-функций. К особенностям этих тест-функций относится наличие дна оврага с острым кусочно-линейным дном и кусочно-линейными склонами. Описывается алгоритм конструирования тест-функций с требуемыми свойствами. При конструировании можно задавать угол изгиба (излома) дна оврага и крутизну склонов. Функции могут применяться для тестирования поисковых методов оптимизации.

2. КОНСТРУИРОВАНИЕ ДВУМЕРНЫХ НЕГЛАДКИХ ТЕСТ-ФУНКЦИЙ

Рассмотрим конструирование тест-функций для двумерного случая ($n = 2$). Пусть дно оврага конструируемой функции представляют собой некоторую кривую, симметричную относительно оси x_2 и проходящую через точку $(0, 0)^T$ (как в функции Розенброка). Кривая $x_2 = a_1|x_1|$ удовлетворяет этим требованиям, более того, она недифференцируема в точке $(0, 0)^T$. Параметр a_1 позволяет управлять углом изгиба дна, т. е. углом между осью x_2 и прямой $x_2 = a_1x_1$. Функция с таким дном оврага может быть аналитически записана в следующем виде: $f(x) = |x_2 - a_1|x_1||$, $x \in R^2$, $a_1 > 0$.

Легко заметить, что полученная запись соответствует члену $100(x_2 - x_1^2)^2$ функции Розенброка.

Построенная функция не пригодна для тестирования методов оптимизации, так как она не обладает единственным минимумом. Все дно оврага составляет множество точек минимума. В функции Розенброка единственность минимума достигается добавлением члена $(1 - x_1)^2$. Для нашей функции добавим член $a_2|1 - x_1|$.

Получили кусочно-линейную функцию, недифференцируемую вдоль дна оврага:

$$\begin{aligned} f(x) &= |x_2 - a_1|x_1|| + a_2|1 - x_1|, \\ x &\in R^2, \quad a_1 > 0, \quad a_2 > 0, \\ x^* &= \operatorname{argmin} f(x) = (1, a_1)^T, \quad f(x^*) = 0. \end{aligned} \quad (1)$$

Данную функцию назовем розенброкоподобной функцией или негладкой овражной функцией Рыкова [8, 9].

Продемонстрируем свойства полученной функции. Функция данного класса имеет шесть линейных областей. Дно оврага описывается уравнением $x_2 = a_1|x_1|$. Каждая функция содержит два параметра: угол излома (изгиба) дна оврага α (т. е. угол между прямыми $x_2 = a_1x_1$ и $x_2 = -a_1x_1$) и угол

между линиями уровня на противоположных склонах оврага β . Изучим зависимость между углами α , β и параметрами a_1 , a_2 . Параметр a_1 является тангенсом угла между прямой $x_2 = a_1x_1$ и осью x_1 . Этот угол равен $\pi/2 - \alpha/2$, где α — угол излома дна оврага. Отсюда

$$a_1 = \operatorname{tg}(\pi/2 - \alpha/2) = \operatorname{ctg}(\alpha/2). \quad (2)$$

Найдем градиенты во всех шести линейных областях, обозначая градиент функции $f(x)$ i -й области через $\nabla_i f(x)$, $i = 1, \dots, 6$. Область 1 описывается следующей системой неравенств:

$$\begin{cases} x_1 < 0, \\ x_2 < -a_1x_1. \end{cases}$$

Теперь легко получить вектор-градиент и квадрат его нормы:

$$\nabla_1 f(x) = (-a_1 - a_2, -1)^T, \quad \|\nabla_1 f(x)\|^2 = (a_1 + a_2)^2 + 1.$$

Для области 2:

$$\begin{cases} x_1 < 0, \\ x_2 > -a_1x_1, \end{cases}$$

$$\nabla_2 f(x) = (a_1 - a_2, 1)^T, \quad \|\nabla_2 f(x)\|^2 = (a_1 - a_2)^2 + 1.$$

Для области 3:

$$\begin{cases} 0 < x_1 < 1, \\ x_2 > a_1x_1, \end{cases}$$

$$\nabla_3 f(x) = (-a_1 - a_2, 1)^T,$$

$$\|\nabla_3 f(x)\|^2 = (a_1 + a_2)^2 + 1 = \|\nabla_1 f(x)\|^2.$$

Для области 4:

$$\begin{cases} x_1 > 1, \\ x_2 > a_1x_1, \end{cases}$$

$$\nabla_4 f(x) = (-a_1 + a_2, 1)^T,$$

$$\|\nabla_4 f(x)\|^2 = (-a_1 + a_2)^2 + 1 = \|\nabla_2 f(x)\|^2.$$

Для области 5:

$$\begin{cases} x_1 > 1, \\ x_2 < a_1x_1, \end{cases}$$

$$\nabla_5 f(x) = (a_1 + a_2, -1)^T,$$

$$\|\nabla_5 f(x)\|^2 = (a_1 + a_2)^2 + 1 = \|\nabla_1 f(x)\|^2 = \|\nabla_3 f(x)\|^2.$$

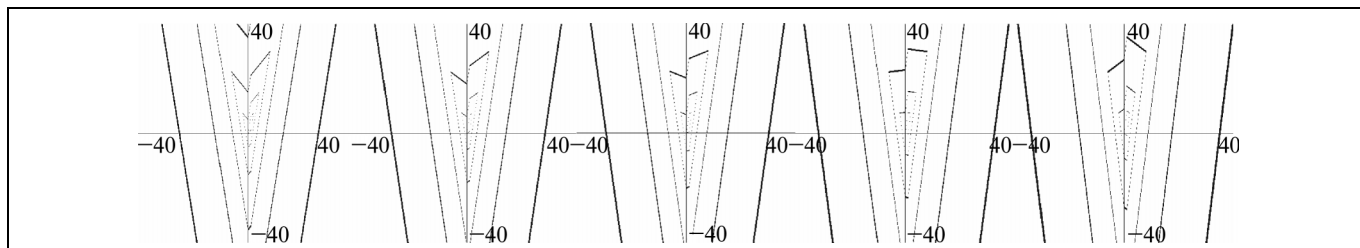
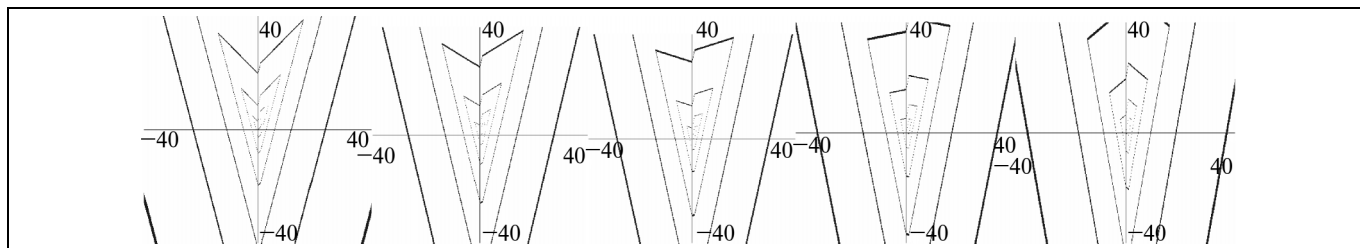
Для области 6:

$$\begin{cases} 0 < x_1 < 1, \\ x_2 < a_1x_1, \end{cases}$$

$$\nabla_6 f(x) = (a_1 - a_2, -1)^T,$$

$$\|\nabla_6 f(x)\|^2 = (a_1 - a_2)^2 + 1 = \|\nabla_2 f(x)\|^2 = \|\nabla_4 f(x)\|^2.$$

Сумма угла β между линиями уровня на противоположных склонах оврага и угла между векторами-градиентами на этих склонах равна π . Вычис-


 Рис. 1. Линии уровня тест-функции Рыкова для $\alpha = 30^\circ$ при $\beta = 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ$, координаты точки минимума (1; 3,732)

 Рис. 2. Линии уровня тест-функции Рыкова для $\alpha = 45^\circ$ при $\beta = 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ$, координаты точки минимума (1; 2,414)

лим скалярное произведение градиентов в областях 1 и 2, 3 и 6, 4 и 5:

$$\begin{aligned} (\nabla_1 f(x), \nabla_2 f(x)) &= (\nabla_3 f(x), \nabla_6 f(x)) = \\ &= (\nabla_4 f(x), \nabla_5 f(x)) = a_2^2 - a_1^2 - 1. \end{aligned}$$

Учитывая, что $(x, y) = \|x\| \|y\| \cos \vec{x}\vec{y}$, $x, y \in R^n$, где $\vec{x}\vec{y}$ обозначает угол между векторами x и y , получим следующее уравнение:

$$\frac{a_2^2 - a_1^2 - 1}{\sqrt{(a_1 + a_2)^2 + 1} \sqrt{(a_1 - a_2)^2 + 1}} = \cos(\pi - \beta). \quad (3)$$

После упрощений получим следующее биквадратное уравнение с неизвестным a_2 :

$$c_2 a_2^4 + c_1 a_2^2 + c_0 = 0, \quad (4)$$

где $c_0 = (a_1^2 + 1)^2(1 - t^2)$, $c_1 = (2a_1^2 - 2)t^2 - (2a_1^2 + 2)$, $c_2 = 1 - t^2$, $t = \cos(\pi - \beta)$.

Нами получен метод для конструирования недифференцируемых тест-функций (1), чьи свойства описываются углами α и β . Для заданных α и β коэффициент a_1 определяется по формуле (2), а

коэффициент a_2 — из множества решений биквадратного уравнения.

Для получения коэффициента a_2 надо из множества решений биквадратного уравнения (4) выбрать положительный корень. Отрицательные корни биквадратного уравнения не позволяют конструировать функцию с минимумом, а второй дополнительный положительный корень возникает после возведения в квадрат уравнения (3).

Формализуем полученный метод конструирования овражных тест-функций в виде алгоритма.

3. АЛГОРИТМ КОНСТРУИРОВАНИЯ ДВУМЕРНЫХ ОВРАЖНЫХ ТЕСТ-ФУНКЦИЙ РЫКОВА

1. Задать α — угол излома дна оврага, угол между прямыми $x_2 = a_1 x_1$ и $x_2 = -a_1 x_1$.
2. Задать β — угол между градиентами противоположных склонов оврага.
3. Вычислить a_1 по формуле $a_1 = \text{tg}(\pi/2 - \alpha/2)$.
4. Вычислить a_2 как положительный действительный корень биквадратного уравнения: $c_2 a_2^4 + c_1 a_2^2 + c_0 = 0$, где $c_0 = (a_1^2 + 1)^2(1 - t^2)$, $c_1 = (2a_1^2 - 2)t^2 - (2a_1^2 + 2)$, $c_2 = 1 - t^2$, $t = \cos(\pi - \beta)$.

 Связь между углами α и β (в градусах) и коэффициентами a_1 и a_2

α	β	a_1	a_2	α	β	a_1	a_2	α	β	a_1	a_2	α	β	a_1	a_2
30	30	3,732	0,400	60	30	1,732	1,094	90	30	1,000	1,984	120	30	0,577	2,860
	45		0,334		45		0,809		45		1,366		45		1,896
	60		0,300		60		0,665		60		1,052		60		1,401
	90		0,259		90		0,500		90		0,707		90		0,866
	120		0,223		120		0,376		120		0,475		120		0,535

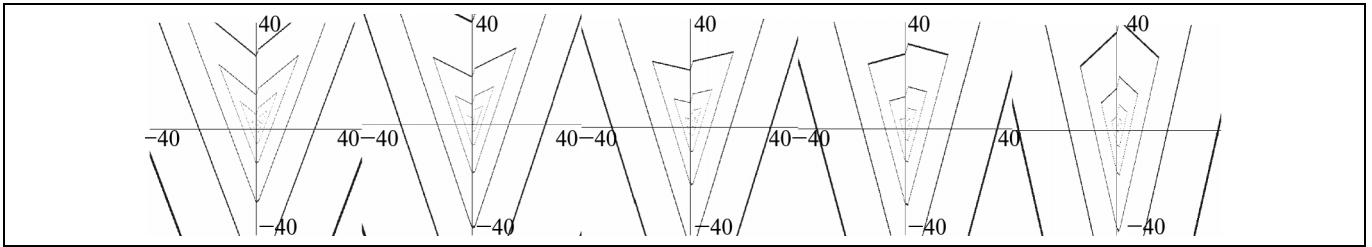


Рис. 3. Линии уровня тест-функции Рыкова для $\alpha = 60^\circ$ при $\beta = 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ$, координаты точки минимума (1; 1,732)

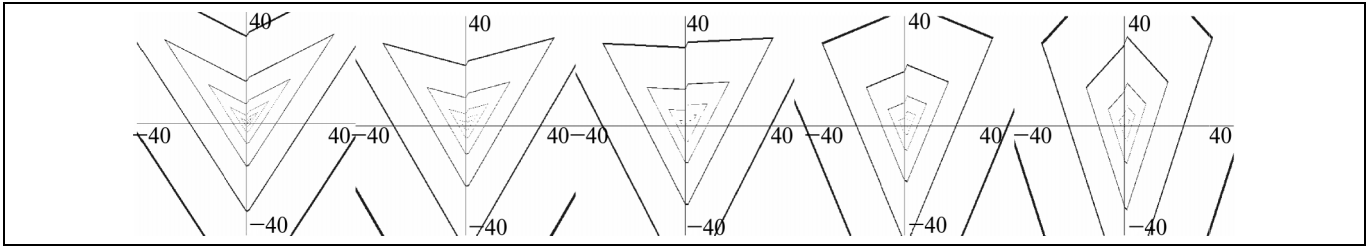


Рис. 4. Линии уровня тест-функции Рыкова для $\alpha = 90^\circ$ при $\beta = 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ$, координаты точки минимума (1; 1)

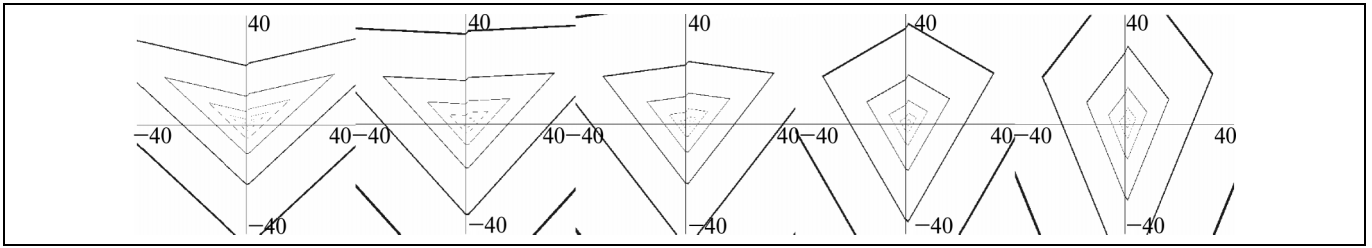


Рис. 5. Линии уровня тест-функции Рыкова для $\alpha = 120^\circ$ при $\beta = 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ$, координаты точки минимума (1; 0,577)

5. Записать полученную тест-функцию $f(x) = |x_2 - a_1|x_1|| + a_2|1 - x_1|$, $x \in R^2$, $a_1 > 0$, $a_2 > 0$,

$$x^* = \arg \min f(x) = (1, a_1)^T, \quad f(x^*) = 0.$$

Коэффициенты a_1 и a_2 двумерных овражных тест-функций (1) для различных значений α и β приведены в таблице. На рис. 1–5 линии уровня сконструированных с помощью предложенного алгоритма двумерных овражных тест-функций Рыкова (1) для различных значений α и β демонстрируют геометрические свойства тестовых функций.

ЗАКЛЮЧЕНИЕ

В статье приведено построение овражных двумерных кусочно-линейных тест-функций для испытания методов оптимизации. Описан первый шаг построения нового класса недифференцируемых тест-функций. Дальнейшее развитие данных тест-функций пойдет по пути построения многомерных вариантов овражных тестовых функций с многомерным дном оврага и наделением функций желаемыми свойствами.

ЛИТЕРАТУРА

1. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. — М.: Мир, 1985. — 509 с.
2. Ларичев О.И., Горвиц Г.Г. Методы поиска локального экстремума овражных функций. — М.: Наука, 1990. — 95 с.
3. Реклейтис Г., Рейвиндран А., Рэгдел К. Оптимизация в технике. Т. 1, 2. — М.: Мир, 1986.
4. Рыков А.С. Поисковая оптимизация. Методы деформируемых конфигураций. — М.: Физматлит: Наука, 1993. — 216 с.
5. Химмельблау Д. Прикладное нелинейное программирование. — М.: Мир, 1975. — 515 с.
6. Brooks S.H. A comparison of maximum-seeking methods // Operat. Res. — 1959. — Vol. 7, N 4. — P. 430–457.
7. Fletcher R. Practical methods of optimization. — N.-Y.: John Wiley & Sons, 1987. — 436 p.
8. Рыков А.С. Методы системного анализа: оптимизация. — М.: Экономика, 1999. — 255 с.
9. Rykov A.S., Kuznetsov A.G. Non-differentiable test function for comparison of direct search optimization techniques // Proc. of the AMSE International Conference on Information and Systems (ICIS). — Hangzhou (China): Zhejiang University, 1991.

e-mail: alexrykov@mail.ru, misha@webgrad.ru

Статья представлена к публикации членом редколлегии С.Д. Земляковым. □

УПРАВЛЯЕМЫЕ СИСТЕМЫ, РОБАСТНЫЕ ПО НАЧАЛЬНЫМ ДАННЫМ¹

Г.Н. Яковенко

Московский физико-технический институт (государственный университет), г. Долгопрудный

Введено понятие робастности по начальным данным: несколько управлений переводят фиксированное начальное состояние в одно и то же конечное, те же управления переводят любое другое начальное состояние в одно и то же конечное. Доказана теорема: управляемая система является робастной по начальным данным тогда и только тогда, когда она допускает максимальную группу симметрий по состоянию. Приведены примеры исследования конкретных управляемых систем на робастность по начальным данным.

ВВЕДЕНИЕ

Возможности систем с управлением зависят не только от внутренней структуры системы, в частности, от уравнений, определяющих функционирование системы, но и от ограничений, наложенных на управление. Наиболее распространено ограничение на значения допустимых управлений. Рассматриваются также ограничения интегрального типа, ограничения на управления как на элементы различных функциональных пространств. В настоящей работе под множеством допустимых управлений понимается такая совокупность управлений, что каждое управление из этой совокупности переводит за конечное время систему из одного фиксированного состояния в другое. Если применить допустимые управления к начальному состоянию, отличному от фиксированного, то возможны два варианта: разным допустимым управлениям будет соответствовать одно и то же конечное состояние (робастность по начальным данным); разные допустимые управления будут приводить в разные конечные состояния (отсутствие робастности по начальным данным). Основной результат работы — теорема, которая утверждает, что система является робастной по начальным данным тогда и только тогда, когда она допускает максимальную группу симметрий по состоянию. Рассмотрение

в работе проводится локально — в некоторой области пространства состояний. Функции, участвующие в построениях, достаточно гладкие.

1. РОБАСТНОСТЬ ПО НАЧАЛЬНЫМ ДАННЫМ

Рассматривается система с управлением

$$\begin{aligned} \dot{x} &= \varphi(x, u), \quad x \in \mathbb{R}^n, \quad u \in U \subset \mathbb{R}^m, \\ x(0) &= x_0. \end{aligned} \quad (1)$$

Вводится обозначение

$$x = f(t, x_0, u(\cdot)) \quad (2)$$

для общего решения системы (1), соответствующего некоторому управлению $u(\cdot)$. Систему (1) будем считать вполне управляемой в общепринятом смысле: $\forall x_0, \forall x_1, \exists T, \exists \tilde{u}(\cdot), x_1 = f(T, x_0, \tilde{u}(\cdot))$.

Выделим из множества допустимых управлений подмножество $M(x_0 \rightarrow x_1)$, каждый элемент которого переводит за конечное время систему из фиксированного начального состояния x_0 в фиксированное конечное x_1 :

$$\{\tilde{u}(\cdot) \in M(x_0 \rightarrow x_1)\} \leftrightarrow \{\exists T, x_1 = f(T, x_0, \tilde{u}(\cdot))\}.$$

Как показывают приведенные далее примеры, управления, принадлежащие некоторому классу $M(x_0 \rightarrow x_1)$, примененные к начальному состоянию x_0^* , отличному от x_0 , необязательно приведут в одно и то же конечное состояние.

Определение 1. Система (1) называется *робастной по начальным данным*, если совокупность управлений, принадлежащих произвольному классу

¹ Работа выполнена при финансовой поддержке РФФИ (проекты 05-01-00940, 07-01-00217) и доложена на VII международной конференции «Идентификация систем и задачи управления» SICPRO'08, Москва, Институт проблем управления РАН, 28–31 января 2008 г.



$M(x_0 \rightarrow x_1)$, примененная к произвольному состоянию x_0^* , приведет к одному и тому же конечному состоянию x_1^* :

$$\{f(T, x_0, u(\cdot)) = f(\tilde{T}, x_0, \tilde{u}(\cdot))\} \leftrightarrow \{f(T, x_0^*, u(\cdot)) = f(\tilde{T}, x_0^*, \tilde{u}(\cdot))\}. \quad \blacklozenge \quad (3)$$

Если система (1) не удовлетворяет условию (3), то назовем ее *неробастной по начальным данным*.

Помимо *робастности по начальным данным* используется и другая терминология: переносимость, мобильность, однократность управляемости [1].

Проиллюстрируем определение 1 на примере системы

$$\dot{x} = u + \tilde{u}x, \quad x \in \mathbb{R}^1, \quad (u, \tilde{u}) \in U \subset \mathbb{R}^2 \quad (4)$$

при различных ограничениях U на значения допустимых управлений.

Пример 1. При $U = \{u = \tilde{u}\}$ система (4) примет вид

$$\dot{x} = u(1 + x), \quad (5)$$

и общее решение (2) можно записать в явном виде

$$x = -1 + (1 + x_0) \exp \int_0^t u(\tau) d\tau. \quad (6)$$

Возьмем в качестве начальной точки $x_0 = 0$, а в качестве конечной $x_1 = 1$, и из решения (6) найдем условие того, что $u(\cdot) \in M(0 \rightarrow 1)$:

$$\exp \int_0^T u(\tau) d\tau = 2. \quad (7)$$

Если в качестве начальной взять другую точку x_0^* , то для любых управлений $u(\cdot)$ и моментов времени T , удовлетворяющих условию (7), конечная точка соответствующих траекторий будет одна и та же: $x_1^* = 1 + 2x_0^*$, т. е. система (5) — робастная по начальным данным.

Пример 2. При $U = \{\tilde{u} \equiv 1\}$ система (4) примет вид

$$\dot{x} = u + x, \quad (8)$$

и общее решение (2) запишется следующим образом

$$x = \left[x_0 + \int_0^t u(\tau) e^{-\tau} d\tau \right] e^t. \quad (9)$$

Приняв, как и в примере 1, в качестве начальной точки $x_0 = 0$, а в качестве конечной $x_1 = 1$, из решения (9) найдем условие того, что $u(\cdot) \in M(0 \rightarrow 1)$:

$$\left[\int_0^T u(\tau) e^{-\tau} d\tau \right] e^T = 1. \quad (10)$$

При другой начальной точке $x_0^* \neq 0$ управлениям, удовлетворяющим условию (10), соответствуют, вообще говоря, разные конечные точки $x_1^* = 1 + x_0^* e^T$. Расположение конечной точки x_1^* зависит от времени T перехода из состояния $x_0 = 0$ в состояние $x_1 = 1$. Если ограничиться только теми управлениями из множества $M(0 \rightarrow 1)$, которые переводят систему (8) из состояния $x_0 = 0$ в состояние $x_1 = 1$ за одно и то же время T^* , то при другой начальной точке $x_0^* \neq 0$ будет обеспечено совпадение точек $x_1^* = 1 + x_0^* e^{T^*}$. Таким образом, у системы (8) робастность по начальным данным носит условный характер.

Пример 3. При отсутствии ограничений на управление система

$$\dot{x} = u + \tilde{u}x \quad (11)$$

не является робастной по начальным данным. Для того чтобы в этом убедиться, ограничимся рассмотрением постоянных управлений и на фиксированном отрезке времени $t \in [0, 1]$. При постоянных управлениях:

$$x(t) = x_0 e^{\tilde{u}t} + u \frac{e^{\tilde{u}t} - 1}{\tilde{u}}. \quad (12)$$

Для того чтобы удовлетворить граничным условиям $x_0 = 0, x_1 = 1$, между постоянными управлениями u и \tilde{u} должна быть связь

$$u = \frac{\tilde{u}}{e^{\tilde{u}} - 1}, \quad (13)$$

которая находится из общего решения (12). Взяв другое начальное состояние, например, $x_0^* = 1$, из решения (12) с учетом связи (13) к моменту времени $t = 1$ получим множество $x_1^* = e^{\tilde{u}} + 1$, не совпадающих конечных точек, что и обосновывает то, что система (11) неробастная по начальным данным.

3. КРИТЕРИЙ РОБАСТНОСТИ ПО НАЧАЛЬНЫМ ДАННЫМ

Робастность по начальным данным тесно связана с групповыми свойствами систем с управлением и присуща только L -системам.

Определение 2 [1–5]. Система

$$\dot{x}^k = \sum_{l=1}^n \varphi_l^k(x) u^l(t), \quad k = \overline{1, n}, \quad u \in U \subset \mathbb{R}^n \quad (14)$$

называется L -системой, если для функций $\varphi_l^k(x)$ выполнены условия $\det \|\varphi_l^k(x)\| \neq 0, [X_i, X_j] = \sum_{i=1}^n C_{ij}^k X_k$,

$$C_{ij}^k = \text{const}, \quad i, j, k = \overline{1, n}, \quad \text{где } X_l = \sum_{k=1}^n \varphi_l^k(x) \frac{\partial}{\partial x^k},$$

$$l = \overline{1, n}, \quad [X_i, X_j] = \sum_{k=1}^n \{X_i \varphi_j^k(x) - X_j \varphi_i^k(x)\} \frac{\partial}{\partial x^k} -$$

коммутаторы операторов X_l .

С L -системой (14) связаны две n -параметрические группы преобразований пространства состояний \mathbb{R}^n [1–4]: группа движений вдоль траекторий системы (14) при различных управлениях

$$x = g(x_0^1, \dots, x_0^n, v^1, \dots, v^n) \quad (15)$$

и группа симметрий по состоянию

$$\hat{x} = \hat{x}(x, \tau), \quad x, \tau \in \mathbb{R}^n. \quad (16)$$

Группа движений (15) включает в себе все преобразования сдвига по траекториям (2) системы (14) при произвольных управлениях $u(\cdot)$ за произвольное время T . Группа симметрий (16) обладает следующим характеристическим свойством: подстановка любого решения $x(t)$ системы (14), соответствующего некоторому управлению $u(\cdot)$, в уравнения группы (16) $\hat{x}(t) = \hat{x}(x(t), \tau)$, $\tau \in \mathbb{R}^n$ при любом наборе параметров τ^1, \dots, τ^n приводит к решению $\hat{x}(t)$ системы (14), соответствующему тому же управлению. Указанные две группы (15), (16) и определяют робастность по начальным данным L -систем.

Теорема. *Вполне управляемая система (1) — является робастной по начальным данным в смысле определения 1 в том и только в том случае, если система является L -системой (14).* ♦

Доказательство см. в Приложении.

Вернемся к рассмотренным ранее примерам. Система (5) является L -системой с группой движений (15) $x = -1 + (1 + x_0)e^v$ и такой же группой симметрий (16) $\hat{x} = -1 + (1 + x_0)e^\tau$. Установленная в примере 1 робастность по начальным данным системы (5) — следствие ее принадлежности к L -системам. Системы (8) и (11) не являются L -системами, вследствие чего не обладают робастностью по начальным данным. Для погружения в класс L -систем добавим к системе (11) уравнение $\dot{y} = \tilde{u}$. Новая система — L -система

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ \tilde{u} \end{pmatrix} \quad (17)$$

по теореме обладает робастностью по начальным данным. Исходя из этого, можно найти дополнительные условия на управления $(u, \tilde{u}) \in M(x_0 \rightarrow x_1)$, чтобы системы (8) и (11) стали робастными по начальным данным.

Принадлежность $(u(\cdot), \tilde{u}(\cdot)) \in M(x_0 \rightarrow x_1)$ приводит в силу системы (17) к соотношению

$$\int_0^T \tilde{u}(\tau) d\tau = y_1 - y_0 = \text{const}, \quad (18)$$

которое и будет указанным дополнительным соотношением. В случае примера 2 ($\tilde{u} \equiv 1$), соотношение (18) принимает вид $T = \text{const}$, что, как было проверено вычислением, действительно приводит к робастности по начальным данным.

Приведем еще примеры L -систем — систем, обладающих робастностью по начальным данным.

Пример 4. Линейная стационарная система

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} Ax + Bu \\ 1 \end{pmatrix} = \begin{pmatrix} Ax & E \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ Bu \end{pmatrix},$$

где $x \in \mathbb{R}^n$, $y \in \mathbb{R}^1$, $u \in \mathbb{R}^m$, E — единичная матрица, A и B — числовые матрицы нужных размеров. Если убрать последнее уравнение, то оставшаяся система не будет обладать свойством робастности по начальным данным.

Пример 5. Движение летательного аппарата в горизонтальной плоскости при управлении креном

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} u \\ \cos x \\ \sin x \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos x & -\sin x \\ 0 & \sin x & \cos x \end{pmatrix} \begin{pmatrix} u \\ x \\ y \end{pmatrix}.$$

Если убрать последнее уравнение, то оставшаяся система не будет обладать свойством робастности по начальным данным.

Пример 6. Уравнение Риккати, погруженное в L -систему:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 1 & 2x & (x)^2 \\ 0 & 2 & 2x \\ 0 & 0 & e^y \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix}. \quad (19)$$

Если ограничиться первым уравнением — уравнением Риккати, то оно не будет обладать свойством робастности по начальным данным.

Пример 7. Уравнение Ньютона

$$m\ddot{s} = u(t) - \beta(\dot{s})^2,$$

определяющее одномерное движение управляемой точки при наличии аэродинамического сопротивления, пропорционального квадрату скорости, погружается в L -систему (19) при следующей специализации переменных:

$$x = m\dot{s}, \quad y = -\frac{2s\beta}{m}, \quad u^1 = u(t), \quad u^2 = 0, \quad u^3 = -\frac{\beta}{(m)^2}.$$



ЗАКЛЮЧЕНИЕ

Введенные определением 2 L -системы (14), кроме робастности по начальным данным, обладают и другими свойствами, присущими исключительно L -системам [5]. Например: с точностью до замен переменных вся информация о системе содержится в введенных в определении 2 постоянных C_{ij}^k ; дифференциальную L -систему (14) можно эквивалентно заменить конечной (без производных) системой (15); дифференциальные уравнения для сопряженных переменных в принципе максимума Л.С. Понтрягина можно заменить конечными уравнениями.

ПРИЛОЖЕНИЕ

Доказательство теоремы

Необходимость. Пусть система (1) робастная по начальным данным. Фиксируем начальное состояние системы. Не нарушая общности, будем считать его началом координат: $x_0 = 0$.

Каждому управлению $\tilde{u}(\cdot)$ и каждому интервалу времени $[0, \tilde{T}]$ соответствует состояние системы v — конечная точка траектории, начальная точка которой $x_0 = 0$:

$$v = f(\tilde{T}, 0, \tilde{u}(\cdot)). \quad (\text{П1})$$

Применение того же управления и того же интервала времени к другим начальным состояниям x_0 системы определит преобразование $x_0 \leftrightarrow x$ пространства состояний \mathbb{R}^n : $x = f(\tilde{T}, x_0, \tilde{u}(\cdot))$. В силу робастности по начальным данным каждое такое преобразование определяется не конкретным выбором управления $\tilde{u}(\cdot)$ и времени \tilde{T} , а только набором чисел $v = (v^1, \dots, v^n)$ — состоянием, в которое переводит пара $\{\tilde{u}(\cdot), \tilde{T}\}$ начало координат $x_0 = 0$. Действительно, если другой паре $\{\tilde{u}(\cdot), \tilde{T}\}$ соответствует в силу соотношения (П1) тот же набор чисел $v = (v^1, \dots, v^n)$, то по свойству (3) робастности по начальным данным каждое состояние x_0 перейдет в то же состояние, что и в случае пары $\{\tilde{u}(\cdot), \tilde{T}\}$. Таким образом, каждое преобразование пространства состояний \mathbb{R}^n , определяемое как сдвиг по траекториям системы для некоторой пары $\{\tilde{u}(\cdot), \tilde{T}\}$, принадлежит n -параметрической совокупности преобразований

$$x = \Phi(x_0, v), \quad (\text{П2})$$

где v — набор чисел $v = (v^1, \dots, v^n)$, задаваемый соотношением (П1). Покажем, что совокупность преобразований (П2) удовлетворяет всем групповым аксиомам.

Замкнутость относительно операции суперпозиции. Пусть пара $\{\tilde{u}(\cdot), \tilde{T}\}$ задает преобразование, которому в силу соотношения (П1) соответствует набор чисел \tilde{v} , а паре $\{\tilde{u}(\cdot), \tilde{T}\}$ — набор чисел \tilde{v} . Определим управление

$$u^*(t) = \begin{cases} \tilde{u}(t), & \text{если } t \in [0, \tilde{T}], \\ \tilde{u}(t - \tilde{T}), & \text{если } t \in [\tilde{T}, \tilde{T} + \tilde{T}], \end{cases}$$

интервал времени $[0, \tilde{T} + \tilde{T}]$ и в силу соотношения (П1) состояние

$$v^* = f(\tilde{T} + \tilde{T}, 0, u^*(\cdot)). \quad (\text{П3})$$

Очевидно, что если к пространству \mathbb{R}^n применить преобразование (П2), которому соответствует набор чисел \tilde{v} , а затем преобразование, которому соответствует набор чисел \tilde{v} , то полученное преобразование принадлежит совокупности (П3) и задается набором параметров v^* , определяемых состоянием (П3).

Наличие в совокупности (П2) тождественного преобразования. Тождественному преобразованию соответствует набор чисел $v^i = 0, i = \overline{1, n}$. Тождественное преобразование пространства \mathbb{R}^n реализуется любой парой $\{\tilde{u}(\cdot), \tilde{T}\}$, удовлетворяющей условию $f(\tilde{T}, 0, \tilde{u}(t)) = 0$, например, $\tilde{T} = 0$ и любое управление $\tilde{u}(\cdot)$.

Наличие обратных преобразований. Пусть паре $\{u(\cdot), T\}$ соответствует в силу соотношения (П1) состояние \tilde{v} . Так как система (1) вполне управляема, существует пара $\{\tilde{u}(\cdot), \tilde{T}\}$, которая переведет систему из состояния \tilde{v} в начало координат и будет реализовывать обратное преобразование. Набор параметров \tilde{v} , соответствующих паре $\{\tilde{u}(\cdot), \tilde{T}\}$, находится с помощью соотношения (18).

Выполнение групповых аксиом дает возможность утверждать, что совокупность преобразований (П2) есть группа (15), а вследствие этого исходная система (1) является L -системой (14) [1–5].

Достаточность. Пусть система (1) — L -система (14) и поэтому допускает группу симметрий (16). Пусть пара $\{\tilde{u}(\cdot), \tilde{T}\}$ переводит некоторое начальное состояние x_0 в конечное x_1 . В силу транзитивности группы симметрий (16) найдется такой набор параметров $\tilde{\tau}$, что соответствующее преобразование группы (16) переведет состояние x_0 в другое заданное состояние x_0^* . При этом состояние x_1 перейдет в конкретное состояние $x_1^* = \hat{x}(x_1, \tilde{\tau})$. Вследствие того, что группа симметрий (16) переводит решения системы (14) в ее же решения, для любой пары $\{\tilde{u}(\cdot), \tilde{T}\}$ перевод x_0 в x_1 влечет за собой перевод x_0^* в одно и то же конечное состояние x_1^* . Так как выбор состояний x_0 и x_0^* произволен, справедливо соотношение (3), т. е. система (14) робастная по начальным данным. Теорема доказана. ♦

ЛИТЕРАТУРА

1. Яковенко Г.Н. Однократность управляемости у групповых систем // Кибернетика и вычислительная техника. — 1985. — Вып. 65. — С. 39–43.
2. Яковенко Г.Н. Симметрии по состоянию в системах с управлением // Прикладная механика и математика: Межвед. сб. науч. тр. / МФТИ. — М., 1992. — С. 155–176.
3. Яковенко Г.Н. Обыкновенные дифференциальные уравнения и системы с управлением — сравнительный групповой анализ // Электронный журнал «Дифференциальные уравнения и процессы управления». — 2002. — 3. — С. 40–83. (<http://www.neva.ru/journal>).
4. Яковенко Г.Н. Дифференциальные уравнения с фундаментальными решениями: Софус Ли и другие. — М.: Физматкнига, 2006. — 112 с.
5. Яковенко Г.Н. Теория управления регулярными системами. — М.: БИНОМ. Лаборатория знаний, 2008. — 264 с.

☎ (495) 408-78-66, e-mail: yakovenko_g@mtu-net.ru

Статья представлена к публикации членом редколлегии С.Д. Земляковым. □

ОБОБЩЕННЫЕ ВЕРОЯТНОСТНЫЕ КРИТЕРИИ В ЗАДАЧЕ НЕЛИНЕЙНОЙ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

С.В. Соколов, П.А. Кучеренко

Ростовский государственный университет путей сообщения

Показана актуальность исследования новых и развития существующих методов стохастической нелинейной параметрической идентификации. Предложено решение задачи идентификации на основе использования обобщенных вероятностных критериев, явно зависящих от апостериорной плотности вероятности. Синтезирован алгоритм идентификации с использованием критерия минимума вероятности ошибки оценивания. Для иллюстрации эффективности предлагаемого подхода рассмотрен численный пример. Предложенный метод может быть весьма эффективно применен в самых различных областях связи, управления, метрологии и пр.

ВВЕДЕНИЕ

Существующие методы решения задачи стохастической параметрической идентификации требуют, как правило, для своей удовлетворительной реализации принятия ряда крайне упрощающих ограничений — линейности модели измерителя относительно параметров, необходимости нормального вида распределения (с нулевыми математическими ожиданиями и конечными дисперсиями) аддитивных помех наблюдаемых сигналов и др. В большинстве реальных ситуаций это оказывает значительное отрицательное влияние на качественные характеристики процедуры идентификации и, как следствие, снижает потенциально возможную точность получаемых оценок параметров [1–8]. Анализ литературы по данной тематике показал, что работы в этой области в основном посвящены вопросам, связанным с идентификацией параметров модели наблюдаемого стохастического объекта, в то время как вопросы, касающиеся определения неизвестных параметров самого наблюдателя (измерителя параметров состояния объекта), практически не освещены. Вместе с тем, задачи подобного рода очень часто возникают в самых различных областях телекоммуникации и связи, радионавигации, метрологии и пр. К числу наиболее распространенных из них можно отнести задачи определения неизвестных характеристик тракта передачи (времени распространения сиг-

налов, параметров среды передачи и др.), определения неизвестных или «шумящих» коэффициентов усиления (или других характеристик) приемника и пр.

Поэтому исследования, направленные на синтез новых и развитие существующих методов и алгоритмов идентификации, ориентированных на возможность оценки параметров наблюдателя и в значительной (или полной) мере устраняющих указанные недостатки разработанных методов, представляются весьма актуальными.

Для решения данной задачи далее рассмотрим подход, позволяющий избавиться от существующих ограничений разработанных методов и повысить потенциальную точность процедуры идентификации благодаря обобщенным вероятностным критериям, зависящим в общем случае нелинейно от апостериорной плотности распределения вектора состояния.

Вопросы идентификации по сложным статистическим критериям уже рассматривались в работах различных авторов [7, 8], но принятые в них критерии, основанные на минимизации среднеквадратической ошибки оценивания и определении апостериорного математического ожидания оцениваемых параметров (в том числе, с использованием статистической и дисперсионной линеаризации) или максимизации апостериорной плотности вероятности (АПВ) параметров наблюдаемого объекта, как это будет видно из последующих



построений, представляют собой частные случаи вероятностных критериев общего вида, зависящих нелинейно от АПВ и составляющих основу предлагаемого подхода.

1. ПОСТАНОВКА ЗАДАЧИ НЕЛИНЕЙНОЙ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

Пусть N -мерный вектор состояния нелинейного дискретного объекта \mathbf{x}_k задан в самом общем случае разностным уравнением

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{n}), \tag{1}$$

где $\mathbf{f}(\dots)$ — известная нелинейная вектор-функция с дифференцируемыми компонентами, для которых существуют обратные функции; \mathbf{x}_{k-1} — N -мерный вектор переменных состояния на $(k - 1)$ -м шаге времени; \mathbf{n} — N -мерный вектор шума с известной N -мерной невырожденной функцией плотности распределения вероятности $q(\mathbf{n})$.

Вектор наблюдения размерности M для k -го момента времени \mathbf{z}_k описывается следующим уравнением (в общем случае также нелинейным):

$$\mathbf{z}_k = \chi(\mathbf{C}_k, \mathbf{x}_k, \mathbf{w}), \tag{2}$$

где $\chi(\dots)$ — известная нелинейная вектор-функция наблюдения с дифференцируемыми компонентами, для которых существуют обратные функции; \mathbf{C}_k — вектор (или матрица) параметров наблюдателя соответствующей размерности, в общем случае нестационарный; \mathbf{w} — M -мерный вектор шума с известной M -мерной невырожденной функцией плотности распределения вероятности $g(\mathbf{w})$.

Для сокращения дальнейшей записи набор векторов сигналов наблюдения $\mathbf{z}_i, i = 1, \dots, k$, обозначим через \mathbf{z}_1^k .

В рассматриваемом общем нелинейном стохастическом случае задача идентификации текущего неизвестного параметра \mathbf{C}_k может быть сформулирована как задача нахождения его значения, удовлетворяющего некоторому вероятностному критерию оптимальности J . В качестве таких критериев, обеспечивающих максимальную (потенциально возможную) точность идентификации, следует использовать обобщенные вероятностные критерии, зависимость которых от апостериорной плотности вероятности переменных состояния $p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k)$ в общем случае нелинейна:

$$J = \int_X \Phi(p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k)) d\mathbf{x}_k = W(\mathbf{C}_k), \tag{3}$$

где Φ — известная нелинейная аналитическая функция, X — заданная область пространства состояний; $p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k)$ — АПВ вектора переменных

состояния объекта, W — известная нелинейная функция.

Различные вариации вида критериальной функции Φ позволяют охватить достаточно широкий класс критериев оптимальности [9]:

— для критерия минимума отклонения АПВ от заданной функции s : $\Phi(p) = (p - s)^2$, $\Phi(p) = -p \ln(s/p)$ (критерий Кульбака) и др.;

— для критерия максимума (минимума) вероятности существования параметров состояния в некоторой области пространства состояний: $\Phi(p) = \pm p$;

— для критерия максимума информации о параметрах состояния: $\Phi(p) = -p \left(\frac{\partial \ln p}{\partial \mathbf{x}_k} \right) \left(\frac{\partial \ln p}{\partial \mathbf{x}_k} \right)^T$ (критерий Фишера), $\Phi(p) = -p \log p$ (критерий Шеннона) и др.

Очевидно, что задача в данной постановке сводится к отысканию текущей апостериорной плотности $p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k)$ и последующему определению текущего значения искомого параметра \mathbf{C}_k , удовлетворяющего выбранному критерию оптимальности J .

2. СИНТЕЗ АЛГОРИТМА НЕЛИНЕЙНОЙ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

Известно [10], что многомерная апостериорная плотность вероятности вектора состояния \mathbf{x} для k -го момента времени $p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k)$ определяется выражением

$$p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k) = \frac{1}{h(\mathbf{C}_k)} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p(\mathbf{x}_{k-1} | \mathbf{z}_1^{k-1}; \mathbf{C}_{k-1}) p(\mathbf{x}_k | \mathbf{x}_{k-1}) d\mathbf{x}_{k-1} p(\mathbf{z}_k | \mathbf{x}_k),$$

$$h(\mathbf{C}_k) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p(\mathbf{x}_{k-1} | \mathbf{z}_1^{k-1}; \mathbf{C}_{k-1}) p(\mathbf{x}_k | \mathbf{x}_{k-1}) d\mathbf{x}_{k-1} p(\mathbf{z}_k | \mathbf{x}_k) \right) d\mathbf{x}_k, \tag{4}$$

где $p(\mathbf{x}_{k-1} | \mathbf{z}_1^{k-1}; \mathbf{C}_{k-1})$ — определенная на $(k - 1)$ -м шаге апостериорная плотность вероятности вектора состояния объекта; $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ — определяемая на текущем шаге алгоритма N -мерная условная плотность вероятности вектора \mathbf{x}_k ; $p(\mathbf{z}_k | \mathbf{x}_k)$ — определяемая на текущем шаге алгоритма функция правдоподобия для многомерного наблюдения.

Многомерная условная плотность $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ может быть получена из исходного уравнения объекта (1) при известном виде плотности распределения вероятности значений шума $q(\mathbf{n})$ (в предпо-

ложении их взаимной статистической независимости) [11]:

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}) = q(\gamma(\mathbf{x}_k, \mathbf{x}_{k-1})) | \mathbf{Y}_\gamma |,$$

$$\text{где } \mathbf{Y}_\gamma = \begin{pmatrix} \frac{\partial \gamma_{(1)}(\mathbf{x}_k, \mathbf{x}_{k-1})}{\partial x_{k(1)}} & \dots & \frac{\partial \gamma_{(1)}(\mathbf{x}_k, \mathbf{x}_{k-1})}{\partial x_{k(N)}} \\ \dots & \dots & \dots \\ \frac{\partial \gamma_{(N)}(\mathbf{x}_k, \mathbf{x}_{k-1})}{\partial x_{k(1)}} & \dots & \frac{\partial \gamma_{(N)}(\mathbf{x}_k, \mathbf{x}_{k-1})}{\partial x_{k(N)}} \end{pmatrix},$$

$$\gamma(\mathbf{x}_k, \mathbf{x}_{k-1}) = \begin{bmatrix} \gamma_{(1)}(\mathbf{x}_k, \mathbf{x}_{k-1}) \\ \vdots \\ \gamma_{(N)}(\mathbf{x}_k, \mathbf{x}_{k-1}) \end{bmatrix},$$

где \mathbf{Y}_γ — якобиан преобразования от вектора переменных \mathbf{n} к вектору переменных \mathbf{x}_k , зависимость которых в общем случае нелинейная и определяется выражением (1); $\gamma_{(i)}(\mathbf{x}_k, \mathbf{x}_{k-1})$, $i = 1, \dots, N$ — полученные в результате обратного преобразования соответствующих компонент $\mathbf{f}(\dots)$ однозначно определенные функции; $x_{k(i)}$, $i = 1, \dots, N$ — компоненты вектора состояния объекта \mathbf{x}_k .

Аналогичным образом из уравнения (2) можно определить входящую в выражение (4) функцию правдоподобия для многомерного наблюдения:

$$p(\mathbf{z}_k | \mathbf{x}_k) = g(\boldsymbol{\psi}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)) | \mathbf{Y}_\psi |,$$

$$\mathbf{Y}_\psi = \begin{pmatrix} \frac{\partial \psi_{(1)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)}{\partial z_{k(1)}} & \dots & \frac{\partial \psi_{(1)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)}{\partial z_{k(M)}} \\ \dots & \dots & \dots \\ \frac{\partial \psi_{(M)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)}{\partial z_{k(1)}} & \dots & \frac{\partial \psi_{(M)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)}{\partial z_{k(M)}} \end{pmatrix},$$

$$\boldsymbol{\psi}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k) = \begin{bmatrix} \psi_{(1)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k) \\ \vdots \\ \psi_{(M)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k) \end{bmatrix},$$

где \mathbf{Y}_ψ — якобиан преобразования от вектора переменных \mathbf{w} к вектору \mathbf{z}_k ; $\psi_{(j)}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)$, $j = 1, \dots, M$ — полученные в результате обратного преобразования соответствующих компонент $\boldsymbol{\chi}(\dots)$ однозначно определенные функции; $z_{k(j)}$, $j = 1, \dots, M$ — компоненты вектора наблюдения \mathbf{z}_k .

Так как АПВ $p(\mathbf{x}_{k-1} | \mathbf{z}_1^{k-1}; \mathbf{C}_{k-1})$ в правой части равенства (4) — известная функция (определенная на предыдущем шаге), рекуррентный алгоритм определения АПВ переменной состояния для k -го момента времени при наличии дискретных отсчетов сигнала наблюдения \mathbf{z}_1^k принимает вид:

$$p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k) = \frac{\Lambda(\mathbf{x}_k, \mathbf{C}_k)}{h^*(\mathbf{C}_k)}, \quad (5)$$

где

$$\Lambda(\mathbf{x}_k, \mathbf{C}_k) = \int \dots \int p(\mathbf{x}_{k-1} | \mathbf{z}_1^{k-1}; \mathbf{C}_{k-1}) \times \\ \times q(\gamma(\mathbf{x}_k, \mathbf{x}_{k-1})) | \mathbf{Y}_\gamma | d\mathbf{x}_{k-1} g(\boldsymbol{\psi}(\mathbf{z}_k, \mathbf{C}_k, \mathbf{x}_k)) | \mathbf{Y}_\psi |,$$

$$h^*(\mathbf{C}_k) = \int \dots \int \Lambda(\mathbf{x}_k, \mathbf{C}_k) d\mathbf{x}_k.$$

Тогда, с учетом выражения (5), обобщенный вероятностный критерий (3) можно окончательно представить следующим образом:

$$J = \int_X \Phi \left(\frac{\Lambda(\mathbf{x}_k, \mathbf{C}_k)}{h^*(\mathbf{C}_k)} \right) d\mathbf{x}_k. \quad (6)$$

Идентификация параметра \mathbf{C}_k предполагает минимизацию (максимизацию) критерия (6). Для этой цели возможно применение известных и широко применяемых методов оптимизации функций многих переменных, выбор которых определяется особенностями исследуемого объекта и его наблюдателя [12–14]. Другими словами, в конечном счете, выбор зависит от конкретного вида критериального выражения — аналитических свойств получаемой целевой функции (выпуклости, непрерывности, повторной дифференцируемости и др.), а также необходимого для реализации самого метода оптимизации уровня информации об этой функции. Так, в случае вида критериального выражения, близкого к квадратичному, целесообразно применять методы, использующие свойства сопряженных направлений (метод Давидона—Флетчера—Пауэлла, обеспечивающий в подобном случае относительно быструю сходимость за конечное число шагов, и т. п.), тогда как для случаев недифференцируемых функций неизвестной формы — методы прямой оптимизации (метод Нелдера—Мида и его модификации, метод Хука—Дживса и др.).

Эффективность предложенного подхода проиллюстрируем на следующем примере, имеющем самостоятельное теоретическое и практическое значение.

3. РЕШЕНИЕ ЗАДАЧИ НЕЛИНЕЙНОЙ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ НА ОСНОВЕ КРИТЕРИЯ МИНИМУМА ВЕРОЯТНОСТИ ОШИБКИ ОЦЕНИВАНИЯ

Исходя из физического смысла поставленной задачи идентификации — т. е. принимая во внимание вероятностный характер текущей ошибки оценивания переменных состояния объекта и очевидную связь ее значения со степенью точности получаемых оценок идентифицируемых параметров — воспользуемся для последующего решения задачи критерием минимума апостериорной плотности вероятности текущей ошибки оценивания



переменных состояния объекта в заданной области ее предельно допустимого изменения. Заметим, что упомянутые ранее и широко применяемые сегодня критерии, основанные на минимизации среднеквадратического отклонения ошибки, в силу неравенства Чебышева представляют собой лишь частные случаи предлагаемого критерия.

Итак, в качестве возможного варианта критерия J примем далее критерий минимума АПВ текущей ошибки оценивания σ_k переменных состояния объекта на выбранном интервале ее предельно допустимого изменения — от σ_{\min} до σ_{\max} , т. е.

$$\min_{\mathbf{C}_k} J = \min_{\mathbf{C}_k} \int_{\sigma_{\min}}^{\sigma_{\max}} \dots \int_{\sigma_{\min}}^{\sigma_{\max}} \rho(\sigma_k | \mathbf{z}_1^k) d\sigma_k,$$

где $\sigma_k = \mathbf{x}_k - \mathbf{x}_k^*$ — вектор текущей ошибки оценивания, \mathbf{x}_k^* — вектор текущих оценок переменных состояния объекта, $\rho(\sigma_k | \mathbf{z}_1^k)$ — многомерная условная апостериорная плотность вероятности ошибки оценивания.

Учитывая линейную зависимость значений текущей ошибки σ_k и переменной состояния x_k : $\sigma_k = x_k - x_k^*$, выразим АПВ ошибки оценивания $\rho(\sigma_k | \mathbf{z}_1^k)$ через определенную ранее АПВ переменной состояния $p(\mathbf{x}_k | \mathbf{z}_1^k; \mathbf{C}_k)$:

$$\rho(\sigma_k | \mathbf{z}_1^k) = p(\sigma_k + \mathbf{x}_k^* | \mathbf{z}_1^k; \mathbf{C}_k).$$

В этом случае минимизацию критерия идентификации можно представить следующим образом:

$$\begin{aligned} \min_{\mathbf{C}_k} J &= \min_{\mathbf{C}_k} \int_{\sigma_{\min}}^{\sigma_{\max}} \dots \int_{\sigma_{\min}}^{\sigma_{\max}} \rho(\sigma_k | \mathbf{z}_1^k) d\sigma_k = \\ &= \min_{\mathbf{C}_k} \int_{\sigma_{\min}}^{\sigma_{\max}} \dots \int_{\sigma_{\min}}^{\sigma_{\max}} p(\sigma_k + \mathbf{x}_k^* | \mathbf{z}_1^k; \mathbf{C}_k) d\sigma_k. \end{aligned} \quad (7)$$

Производя соответствующую замену переменных в полученном ранее выражении для АПВ параметров состояния (5) и обозначив критериальное выражение через $\Omega(\mathbf{C}_k)$, поиск минимума критерия (7) можно представить в следующем компактном виде:

$$\min_{\mathbf{C}_k} J = \min_{\mathbf{C}_k} \Omega(\mathbf{C}_k), \quad (8)$$

где

$$\begin{aligned} \Omega(\mathbf{C}_k) &= \int_{\sigma_{\min}}^{\sigma_{\max}} \dots \int_{\sigma_{\min}}^{\sigma_{\max}} p(\sigma_k + \hat{\mathbf{x}}_k | \mathbf{z}_1^k; \mathbf{C}_k) d\sigma_k = \\ &= \int_{\sigma_{\min}}^{\sigma_{\max}} \dots \int_{\sigma_{\min}}^{\sigma_{\max}} \left(\frac{\Lambda(\sigma_k + \mathbf{x}_k^*; \mathbf{C}_k)}{h^*(\mathbf{C}_k)} \right) d\sigma_k. \end{aligned}$$

Как было отмечено, идентификация текущего значения параметра, удовлетворяющего критерию (8), предполагает минимизацию функции $\Omega(\mathbf{C}_k)$ на основе известных методов оптимизации функций многих переменных, конкретное применение которых рассмотрим далее в процессе численного моделирования конкретной процедуры идентификации.

4. ПРИМЕР

Проиллюстрируем эффективность предложенного метода идентификации следующим примером.

Пусть стохастический дискретный объект задан нелинейным разностным уравнением

$$\begin{aligned} x_k &= f(x_{k-1}) + n = x_{k-1} + ax_{k-1}^2 + n, \\ x_1 &= 0,5, \end{aligned} \quad (9)$$

где n — белый гауссовский шум с нулевым средним и дисперсией $D_n = 0,025$, $a = -0,4$ — коэффициент уравнения объекта.

Переменные состояния заданного объекта наблюдаются измерителем, описываемым уравнением

$$z_k = \chi(c, x_k) + w = bcx_k^2 + w, \quad (10)$$

где $b = 0,35$ — коэффициент уравнения объекта, c — неизвестный искомым параметр наблюдателя (для рассматриваемого далее модельного примера выберем исходное значение $c = 1$ для всех k), w — белый гауссовский шум с нулевым средним и дисперсией $D_w = 0,15$.

Для последнего определения оценок воспользуемся достаточно эффективным в смысле вычислений рекуррентным нелинейным фильтром Калмана (предполагающим линеаризацию нелинейных функций $f(x_{k-1})$ и $\chi(c, x_k)$ в окрестности оценки). Линеаризованные уравнения объекта (9) и измерителя (10) принимают следующий вид:

$$\begin{aligned} x_k &= A_k x_{k-1} + B_k + n, \quad x_1 = 0,5, \\ z_k &= c(E_k x_k + P_k) + w, \end{aligned}$$

где $A_k = 1 - 0,8x_{k-1}^*$, $B_k = -0,4(x_{k-1}^*)^2$, $E_k = 0,7x_{k-1}^*$, $P_k = -0,35(x_{k-1}^*)^2$ — коэффициенты, полученные в результате линеаризации функций $f(x_{k-1})$ и $\chi(c, x_k)$ в окрестностях оценок переменной состояния объекта для $(k-1)$ -го шага x_{k-1}^* .

Текущие значения оценки переменной состояния объекта определялись посредством субоптимального фильтра Калмана [10], имеющего в рассматриваемом случае вид

$$\begin{aligned} x_k^* &= U(c) = A_k x_{k-1}^* + B_k + \\ &+ K_k (z_k - c(E_k (A_k x_{k-1}^* + B_k) + P_k)), \quad x_1^* = 0,7, \\ K_k &= c \frac{R_k}{D_w}, \quad R_k = \left(\frac{1}{R_{k-1} + D_n} + \frac{c^2}{D_w} \right)^{-1}, \quad R_1 = 0,5. \end{aligned}$$

Соответственно, функция $\Omega(c)$ для k -го шага алгоритма примет вид

$$\begin{aligned} \Omega(c) &= \int_{\sigma_{\min}}^{\sigma_{\max}} p(\sigma_k + \hat{x}_k | z_1^k; c) d\sigma_k = \\ &= \int_{\sigma_{\min}}^{\sigma_{\max}} \left(\frac{1}{h^*(c)} \int_{-\infty}^{+\infty} p(x_{k-1} | z_1^{k-1}) \frac{1}{D_n \sqrt{2\pi}} \times \right. \\ &\quad \times e^{-\frac{(\sigma_k + U(c) - f(x_{k-1}))^2}{2D_n}} dx_{k-1} \frac{1}{D_w \sqrt{2\pi}} \times \\ &\quad \left. \times e^{-\frac{(z_k - \chi(c, \sigma_k + U(c)))^2}{2D_w}} \right) d\sigma_k, \end{aligned} \quad (11)$$

$$\begin{aligned} h^*(c) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x_{k-1} | z_1^{k-1}) \times \\ &\quad \times \frac{1}{D_n \sqrt{2\pi}} e^{-\frac{(\sigma_k + U(c) - f(x_{k-1}))^2}{2D_n}} dx_{k-1} \times \\ &\quad \times \frac{1}{D_w \sqrt{2\pi}} e^{-\frac{(z_k - \chi(c, \sigma_k + U(c)))^2}{2D_w}} dx_k. \end{aligned}$$

Априорную плотность вероятности для первой итерации алгоритма выберем нормальной с дисперсией $D_0 = 0,45$ и математическим ожиданием $0,7$. Здесь интересно отметить, что отклонения среднего значения априорной плотности от начального значения переменной состояния не оказывают в дальнейшем существенного влияния на качество процедуры идентификации (алгоритм идентификации к ним устойчив).

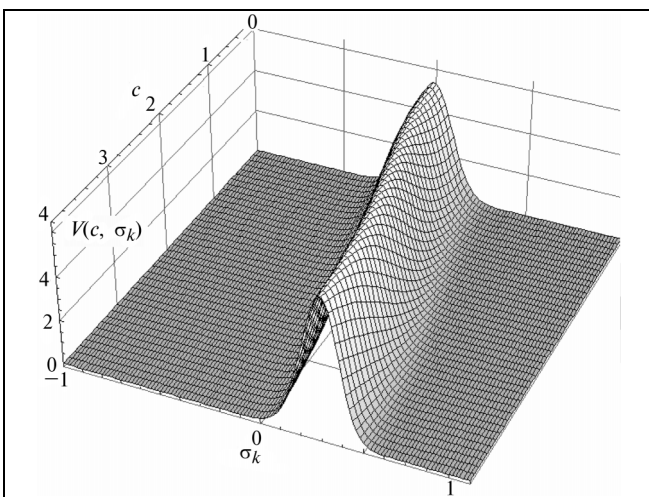


Рис. 1. График АПВ текущей ошибки оценивания для k -го шага алгоритма ($k = 150$)

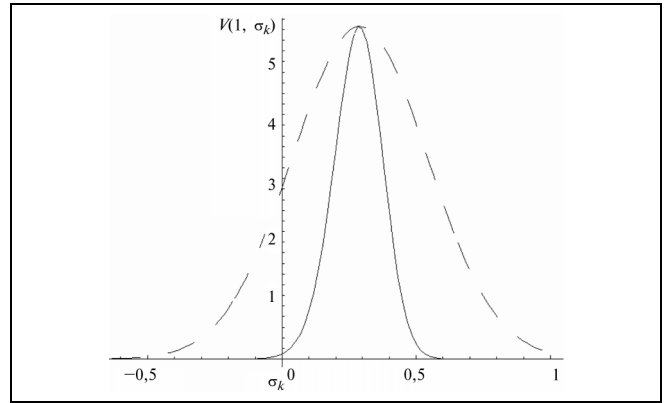


Рис. 2. Сечение вдоль оси σ_k функции $V(c, \sigma_k)$ при $c = 1$ и гауссовская плотность распределения с математическим ожиданием $0,29$ и дисперсией $0,064$ ($k = 150$)

На рис. 1 представлен полученный в результате моделирования график входящей в формулу (11) АПВ текущей ошибки оценивания для k -го шага алгоритма ($k = 150$), которая, являясь функцией текущей ошибки σ_k , зависит в то же время от значений искомого параметра c : $(p(\sigma_k + \hat{x}_k^* | z_1^k) = p(\sigma_k + U(c) | z_1^k) = V(c, \sigma_k))$.

Интересно, что характер получаемых в результате моделирования АПВ (как для переменной состояния нелинейного объекта, так и для ошибки оценивания σ_k) существенно негауссовский (ввиду нелинейности исходных уравнений объекта и измерителя). Отмеченное обстоятельство иллюстрируется приведенными на рис. 2 графиками, соответствующими сечению вдоль оси σ_k функции $V(c, \sigma_k)$ при $c = 1$ (непрерывная линия) и гауссовской плотности распределения с математическим ожиданием $0,29$ и дисперсией $0,064$ (штриховая линия).

Как видно из представленных графиков, указанные функции, имея похожие формы, в то же время, значительно различаются своими значениями практически на всем моделируемом интервале своих аргументов.

Интегралы в выражении (11) здесь и в дальнейшем определялись численно с помощью квадратурных формул с шагом $\Delta = 0,025$. Бесконечные пределы интегрирования по переменной состояния x были заменены на конечные, удовлетворяющие точностным требованиям к алгоритму оценки ($x_{\min} = -0,5$; $x_{\max} = 1,5$).

Эффективность работы алгоритма иллюстрируется приведенным на рис. 3 графиком зависимости критериального выражения $\Omega(c)$ от искомого параметра.

Очевидно, что минимум критериального выражения (и, следовательно, сечение построенной

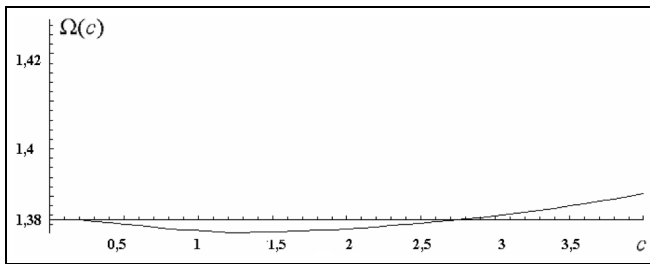


Рис. 3. Зависимость функции $\Omega(c)$ от искомого параметра ($k = 150$)

функции $V(c, \sigma_k)$ вдоль оси σ_k с наименьшей площадью) располагается около истинного значения искомого параметра $c = 1$.

Как показали результаты моделирования, вид приведенной на рис. 3 зависимости характерен для критериальных выражений, получаемых на различных итерациях алгоритма. Здесь важно отметить, что, будучи многоэкстремальными, критериальные выражения на различных шагах алгоритма (см. приведенный график) принимают свои наименьшие значения в районе истинного значения искомого параметра.

Для минимизации выбранного критерия на очередном шаге алгоритма — т. е. для однозначного определения численного значения текущей оценки искомого параметра — целесообразно, задав некоторый интервал возможных значений параметра c (в примере был выбран интервал $0 \leq c \leq 6$), воспользоваться одним из методов прямой минимизации. В данном случае применялся модифицированный симплексный метод прямой минимизации Нелдера—Мида — метод Бокса, обладающий достаточной вычислительной эффективностью и удобной программной реализацией [14].

Результаты компьютерного моделирования процедуры нелинейной параметрической идентификации показали, что при выборе числа дискретных значений сигнала наблюдения $k \geq 280$ отклонение оценки параметра c наблюдателя от его истинного значения не превышает 17 % от его значения.

Таким образом, результаты проведенных исследований подтверждают принципиальную возможность эффективной реализации метода нелинейной параметрической идентификации с критерием минимума АПВ текущей ошибки оценивания.

ЗАКЛЮЧЕНИЕ

Получено общее решение задачи нелинейной параметрической идентификации, обладающее рядом принципиально новых свойств. К их числу следует отнести:

— более высокий по сравнению с существующими методами уровень потенциальной точности

процесса идентификации благодаря обобщенным вероятностным критериям, зависящим в общем случае нелинейно от апостериорной плотности распределения вектора состояния и позволяющим охватить широкий класс условий оптимальности по точности;

— инвариантность к виду плотности распределения вероятности шума как объекта, так и наблюдателя;

— возможность применения метода для нелинейных объектов и наблюдателей, в том числе при нелинейной зависимости функции наблюдения от параметра.

В качестве основной проблемы практической реализации предлагаемого в работе подхода можно отметить некоторую вычислительную сложность определения плотностей вероятности многомерных случайных сигналов, однако, учитывая современные тенденции развития вычислительных средств, проблема подобного рода не представляется критической.

Предложенный метод нелинейной параметрической идентификации на основе обобщенных вероятностных критериев может быть весьма эффективно применен в самых различных областях связи, управления, метрологии и др.

ЛИТЕРАТУРА

1. Грон Д. Методы идентификации систем. — М.: Мир, 1979. — 302 с.
2. Теория моделей в процессах управления / Б.Н. Петров, Г.М. Уланов, И.И. Гольденблат, С.В. Ульянов. — М.: Наука, 1978. — 223 с.
3. Справочник по теории автоматического управления / Под ред. А.А. Красовского. — М.: Наука, 1987. — 712 с.
4. Штейнберг Ш.Е. Идентификация в системах управления. — М.: Энергоатомиздат, 1987. — 80 с.
5. Эйхофф П. Основы идентификации систем управления. — М.: Мир, 1975. — 683 с.
6. Ли Р. Оптимальные оценки, определение характеристик и управление. — М.: Наука, 1966. — 176 с.
7. Пугачев В.С., Казаков И.Е., Евланов Л.Г. Основы статистической теории автоматических систем. — М.: Машиностроение, 1974. — 400 с.
8. Пащенко Ф.Ф. Введение в состоятельные методы моделирования систем. Идентификация нелинейных систем. — М.: Финансы и статистика, 2007. — 288 с.
9. Хуторцев В.В., Соколов С.В., Шевчук П.С. Современные принципы управления и фильтрации в стохастических системах. — М.: Радио и связь, 2001. — 807 с.
10. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. — М.: Радио и связь, 1991. — 608 с.
11. Левин Б.Р. Теоретические основы статистической радиотехники. — М.: Радио и связь, 1989. — 455 с.
12. Измаилов А.Ф., Солодов М.В. Численные методы оптимизации. — М.: Физматлит, 2003. — 304 с.
13. Васильев Ф.П. Методы оптимизации. — М.: Факториал, 2002. — 824 с.
14. Банди Б. Методы оптимизации: Вводный курс. — М.: Радио и связь, 1988. — 128 с.

e-mail: pavelpost83@mail.ru

Статья представлена к публикации членом редколлегии Ф.Ф. Пащенко. □

АНАЛИЗ ФЕНОМЕНОВ КОЛЛЕКТИВИЗМА И ЭГОИЗМА В КОНТЕКСТЕ ОБЩЕСТВЕННОГО БЛАГОСОСТОЯНИЯ

П.Ю. Чеботарев, А.К. Логинов, Я.Ю. Цодикова, З.М. Лезина, В.И. Борзенко

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Исследована сравнительная выгодность основных социальных установок, к которым относятся коллективизм и эгоизм, в условиях демократического принятия коллективных решений.

ВВЕДЕНИЕ. ПРОБЛЕМА

Эгоизм участника голосования означает поддержку тех предложений, которые соответствуют его индивидуальным интересам. Коллективизм означает поддержку интересов группы, с которой участник себя идентифицирует. Проблема эгоиста состоит в том, что, имея лишь один голос, он часто не в состоянии эффективно защитить свои специфические интересы. Солидарно голосующая группа имеет больше шансов на успех, но, у ее участника тоже есть проблема, и суть ее в том, что чем больше и, соответственно, могущественнее группа, тем больше ее интересы могут расходиться с индивидуальными интересами участника. Поэтому вопрос: «Что выгоднее — играть за себя или защищать групповые интересы, не всегда совпадающие с собственными?» нетривиален и не имеет универсального ответа. Аналогичный вопрос встает перед политическими партиями и движениями: они должны решать, с какими другими политическими субъектами объединяться, с какими блокироваться, или же ни с кем не объединяться и не блокироваться, а участвовать в политическом процессе автономно.

Если группа мала и объединяет участников с близкими интересами, то коллективизм ее членов приближается к эгоизму. Если, напротив, группа является очень большой, в пределе совпадая со всем обществом, то коллективизм становится ва-

риантом альтруизма: действительно, каждый участник такой группы защищает общие интересы.

Проблемы соотношения эгоизма, альтруизма и рациональности изучались, в частности, в работах [1—3], а голосование как метод принятия решений о перераспределении социальных благ посредством сбора налогов и реализации социальных программ — в работах [4—7]. Как было показано в теории голосований (здесь следует в первую очередь отметить работы А.В. Малишевского (1970) из Института проблем управления РАН [8, с. 92—95] и Р. Маккельви (1976) из Калифорнийского технологического института [9]), демократическими решениями, которые принимаются голосующими-эгоистами, нетрудно манипулировать посредством формулирования предложений (см. также заметку [10]). Так, итог последовательности решений, принятых *подавляющим большинством голосов*, может быть невыгоден для *всех без исключения* голосующих! В отношении коллективистов, максимизирующих благо очень больших групп, такое манипулирование не проходит. Поэтому представляет большой интерес исследование следующего типа социальной динамики. Некоторое число участников объединяются с целью отстаивания своих групповых интересов. Предположим, что им удастся решить эту задачу успешно — т. е. средний участник группы добивается большего, чем если бы он ни с кем не объединялся. Поскольку группа принципиально открыта для присоединения новых членов, численность ее растет. Если группа



остается при этом успешной, то, вообще говоря, нет препятствий для ее расширения до размеров всего общества, и этому сопутствует приближение группового эгоизма к альтруизму. Разумеется, необходимым условием реализуемости подобной «социальной утопии» является успешность группы в защите интересов ее членов.

Аналізу возникающих при этом феноменов с помощью математических моделей посвящены несколько наших работ. В частности, в статьях [11–13], показано, что для широких областей в пространстве параметров изучаемой модели группа, действительно, может (причем именно в интересах своих членов) успешно конкурировать с индивидуальными участниками. Это свидетельствует в пользу принципиальной возможности нарастания «снежного кома» кооперации — тип социальной динамики, описанный выше. Вместе с тем, существуют условия, и они изучены, при которых группа проигрывает, а более выгодным оказывается индивидуальное поведение.

Для дальнейшего исследования выгодности и эффективных механизмов кооперативного поведения следует рассмотреть случай, когда группа конкурирует не только с индивидуальными участниками, но и с другими группами («партиями»). Случай конкуренции двух групп является предметом настоящей работы.

1. МОДЕЛЬ И СОДЕРЖАТЕЛЬНАЯ ПОСТАНОВКА ЗАДАЧИ

Модель, необходимая для изучения сравнительной выгодности коллективистской и эгоистической установок, должна быть достаточно богатой, чтобы отражать основные элементы моделируемой реальности, и достаточно простой, чтобы допускать исследование аналитическими методами.

В рамках модели должны формализовываться, в частности, следующие понятия: интересы участников, кооперация, интересы групп, защита интересов, благоприятность ситуации, вариативность среды, социальные возможности и вызовы, принятие и реализация коллективных решений.

Основные положения предложенной нами модели.

1. Текущее состояние каждого участника характеризуется значением количественного показателя, который может интерпретироваться как уровень благосостояния, капитал (далее для краткости будем придерживаться этой интерпретации), социальный вес, удовлетворенность, успешность и т. д. Этот показатель может принимать как положительные, так и отрицательные значения. Модель допускает задание начальных значений этого показателя для всех участников.

2. Интерес участника-эгоиста состоит в максимизации его капитала.

3. Участники могут объединяться в группы.

4. Успешность группы характеризуется убытками функциями от капиталов ее участников.

5. Изучается социальная динамика, определяемая коллективными решениями, которые принимаются путем голосования. Голосуют все участники.

6. Для принятия решений используются процедуры α -большинства. Параметр процедуры голосования α есть доля голосов, превышение которой необходимо и достаточно для принятия предложения. Так, $\alpha = 0,5$ отвечает процедуре простого большинства голосов. Если предложение не принято, сохраняется *статус-кво*, т. е. капиталы участников остаются неизменными.

7. Голосуя, участник-эгоист поддерживает те и только те предложения, которые приводят к увеличению его капитала (это конкретизация п. 2).

8. Группа голосует солидарно. Среди рассматриваемых критериев поддержки предложений группой: а) критерий α -большинства членов группы (с порогом α_1 , возможно, отличным от общего порога голосования); б) достаточно высокое (превышающее выбранный порог, например, порог 0) среднее приращение капитала членов группы.

9. Предложение моделируется вектором приращений капиталов всех участников.

10. В качестве простейшего механизма порождения предложений рассматриваются реализации случайного вектора с независимыми одинаково распределенными компонентами. Такие предложения интерпретируются как возможности, генерируемые стохастической средой — в целом благоприятной, нейтральной или неблагоприятной.

11. В качестве простейшего распределения одной компоненты вектора приращений капиталов рассматривается нормальное распределение $N(\mu, \sigma^2)$, параметры которого μ и σ характеризуют среду: $\mu > 0$, $\mu = 0$ и $\mu < 0$ характеризуют благоприятную, нейтральную и неблагоприятную среду соответственно; σ характеризует вариативность среды (разброс предложений).

12. Изучаются различные варианты модели, отличающиеся условиями присоединения участников к группам и выхода из них, условиями возможного выбывания «разорившихся» участников (т. е. участников, капиталы которых стали отрицательными), возможной зависимостью приращения капитала участника от текущего капитала и др. Рассматриваются также иные правила голосования, иные принципы формирования предложений и критерии поддержки их группами участников.

Основным предметом исследования является социальная динамика, определяемая коллектив-

ными решениями при сформулированных выше предположениях. Анализируются, в частности:

- сравнительная выгодность (как для участников, так и для общества в целом) различных социальных установок, к числу которых относятся эгоизм и коллективизм;
- зависимость результатов от параметров модели, в частности, от благоприятности и вариативности среды, порога голосования, принципов поддержки предложений группами, соотношения численности участников с разными установками и др.

Особое внимание уделяется интерпретации результатов в социальных терминах.

В данной работе рассматривается, главным образом, случай, когда участники делятся на три категории: эгоисты и члены двух групп («партий»).

2. «СНЕЖНЫЙ КОМ» КООПЕРАЦИИ

Одна из основных проблем социальной реальности может быть сформулирована следующим образом. Эгоистическое поведение обычно приводит социальных агентов к довольно печальному итогу. «Чистый» эгоист, попавший в беду или просто нуждающийся в помощи, как правило, помощи не получает; кроме того, сталкиваясь с мало-мальски организованным противодействием, он обычно оказывается бессилем. Альтруистическое поведение, вообще говоря, выгоднее для всего общества, но лишь до тех пор, пока оно характерно для абсолютного большинства участников. В противном случае альтруистам остается лишь чувство моральной правоты, а остальные дивиденды делятся между менее благородными участниками. Иными словами, альтруизм — это выгодное, но крайне неустойчивое равновесие. В современном обществе сколько-нибудь массовые его проявления ограничиваются сферой символических действий (любезность, предупредительность, иногда стилем напоминаящие Бобчинского и Добчинского) и поводов, всерьез не затрагивающих интересов «дающего» (пример — благотворительность). В качестве важных исключений необходимо отметить кодексы поведения в чрезвычайных ситуациях (не всегда выполняющиеся): спасение в первую очередь женщин и детей при кораблекрушении и т. п. Увы, мы видим, что с коммерциализацией постепенно утрачивают свой исключительный статус так называемые благородные профессии. И альтруизм остается прерогативой благородных людей, немногочисленных в любую эпоху.

По-видимому, нет изменений в сторону альтруизма и в международной политике. Как и раньше, внешнеполитическая деятельность каждого правительства направлена на защиту интересов своего

государства, что часто сочетается с довольно циничным отношением к интересам народов других стран. Аналогично во внутренней политике: представитель выборной власти, даже если он защищает интересы своих избирателей, то зачастую готов жертвовать интересами других жителей своей страны.

Приведенные политические примеры дают образчики корпоративного (кооперативного) поведения. Представитель государства (общности, корпорации) защищает интересы своей общности, конкурирующей с другими. Следует подчеркнуть, что, рассматривая весь диапазон социальных установок от эгоизма до альтруизма, можно построить примеры кооперативного поведения, относящиеся к каждому участку этого диапазона.

Как мы уже отмечали, самого пристального внимания заслуживает случай кооперативного поведения в рамках открытого и конкурентоспособного — и потому расширяющегося — объединения участников. По мере его расширения кооперативное поведение постепенно приближается к альтруистическому. Этот социальный механизм может быть назван «снежным комом» кооперации, а упоминание на него можно в шутку назвать «снежным комом»-мунизмом. В следующих разделах остановимся на результатах исследования модели, описанной в § 1.

3. КОГДА «СНЕЖНЫЙ КОМ» РАСТЕТ, А КОГДА ТАЕТ?

Прежде всего, рассмотрим общество, состоящее из социальных агентов двух типов: эгоистов и членов группы. Зададимся следующим вопросом: как успешность группы, эгоистов и общества в целом зависит от размера группы.

Формулы, полученные в работе [12], позволяют ответить на этот вопрос при любых значениях параметров модели. Приведем здесь приближенные выражения, которые менее громоздки, чем точные, и дают незначительную ошибку аппроксимации (за исключением случаев малых численностей, в которых мы используем точные выражения). В следующей теореме пара величин, заключенная в квадратные скобки, обозначает матрицу-строку или матрицу-столбец, одна величина в квадратных скобках — целую часть числа.

Теорема [12]. Пусть $M(\tilde{d}_\varepsilon)$ и $M(\tilde{d}_G)$ — математические ожидания приращений капиталов эгоиста и члена группы за один шаг соответственно. Предполагается, что группа поддерживает предложение тогда и только тогда, когда оно обеспечивает ей положительное суммарное приращение капитала. Тогда стандартная нормальная аппроксимация биноми-



ального распределения позволяет получить следующие приближенные значения:

$$M(\tilde{d}_\varepsilon) \approx [P_G Q_G] \left(\mu \begin{bmatrix} F_\gamma \\ F_\alpha \end{bmatrix} + \frac{\sigma f}{\sqrt{pql}} \begin{bmatrix} f_\gamma \\ f_\alpha \end{bmatrix} \right),$$

$$M(\tilde{d}_G) \approx [F_\gamma F_\alpha] \left(\mu \begin{bmatrix} P_G \\ Q_G \end{bmatrix} + \frac{\sigma f_G}{\sqrt{g}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right), \quad (1)$$

где $P_G = F\left(\frac{\mu\sqrt{g}}{\sigma}\right)$, $Q_G = 1 - P_G = F\left(-\frac{\mu\sqrt{g}}{\sigma}\right)$, $F_\theta = F\left(\frac{[\theta n] + 0,5 - pl}{\sqrt{pql}}\right)$, $f_\theta = f\left(\frac{[\theta n] + 0,5 - pl}{\sqrt{pql}}\right)$,

$\theta \in \{\alpha, \gamma\}$, $\gamma = \alpha - g/n$, $p = F(\mu/\sigma)$, $q = 1 - p$, $f = f(\mu/\sigma)$, $F(\cdot)$ и $f(\cdot)$ — стандартная нормальная функция распределения и соответствующая функция плотности, l — число эгоистов, g — число членов группы, $n = l + g$ — число участников. ♦

Типичный пример зависимости успешности участников от размера группы показан на рис. 1. В этом примере общество состоит из 1000 агентов, по горизонтальной оси отложена численность группы (от 0 до 1000); группа голосует, как указано в теореме. Параметры распределения предложений: $\mu = -0,8$, $\sigma = 30$. Порог голосования: $\alpha = 0,5$. По вертикальной оси отложено среднее (ожидаемое) приращение капитала участника группы, эгоиста и случайно выбранного члена общества за 1 шаг.

Заметим следующее. Каким бы ни был размер группы, она всегда имеет существенное преимущество перед эгоистами. Поэтому эгоистам выгодно присоединяться к группе независимо от ее размера. Участникам очень маленькой группы крайне

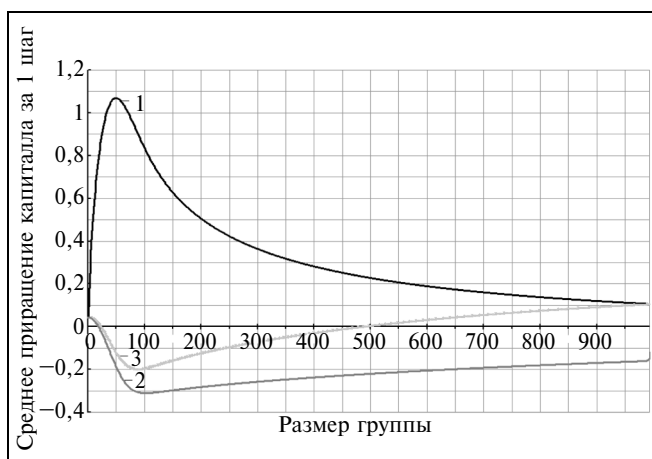


Рис. 1. Средние приращения капиталов члена группы (1), эгоиста (2) и случайно выбранного участника (3) за 1 шаг: 1000 участников — группа и эгоисты; $\mu = -0,8$, $\sigma = 30$, $\alpha = 0,5$

выгодно ее увеличение. Эгоистам и обществу в целом расширение малой группы, напротив, невыгодно. В данном примере размер группы, наиболее выгодный для ее участников, — 50 чел. Дальнейшее расширение группы приводит к понижению средних приращений капитала ее участников. Но, начиная с 88 членов группы, дальнейшее ее расширение становится выгодным для общества в целом, а начиная с 102 членов, оно выгодно уже и для эгоистов. Общество в целом входит в зону положительных средних приращений капитала при 488 членах группы и достигает такого же значения, как при отсутствии группы, когда в группе 649 членов.

Отсюда может быть сделан следующий вывод, важный для разработки алгоритмов управления. Если «правила игры» принимаются теми, кто заинтересован в максимизации суммарного капитала всего общества, то следует сделать группу открытой для свободного присоединения новых членов. Тогда правдоподобным сценарием является вхождение всех участников в группу и превращение группового эгоизма в вариант альтруизма. Отметим, что это «условный» альтруизм, потому что стоит кому-то начать голосовать за свои интересы, не совпадающие с общими, и все остальные члены группы тут же перестанут защищать его интересы. Иными словами, «все за одного» лишь до тех пор, пока «один за всех». Именно это условие обеспечивает устойчивость рассматриваемого «кооперативного» альтруизма, отличающую его от альтруизма «безусловного».

Если же «правила игры» определяются самими участниками, то каждый из них будет заинтересован, чтобы верхней границей размера группы было назначено то наименьшее, но не меньшее 50-ти, число, при котором гарантируется пребывание данного участника в группе. Если 50 участников группы уже набраны, то ее члены заинтересованы «подвести черту». Группа при этом представляет собой солидарную «элиту», напоминающую партию в однопартийной политической системе. Для общества в целом наличие такой группы невыгодно, но еще хуже для общества — допустить в группу определенное число членов сверх того, которое обеспечивает группе оптимум. В данном примере точке минимума суммарного общественного капитала соответствует группа в 88 членов.

Разумеется, кроме попыток проникнуть в группу, де-факто «элитную», есть другой, более естественный путь: образовать другую, конкурирующую группу. В § 4 мы обсудим, к чему это приводит.

4. ВТОРОЙ «СНЕЖНЫЙ КОМ»

Пусть группа достигла оптимального для нее размера, и дальнейшее ее расширение блокируется. Рассмотрим формирование второй группы при

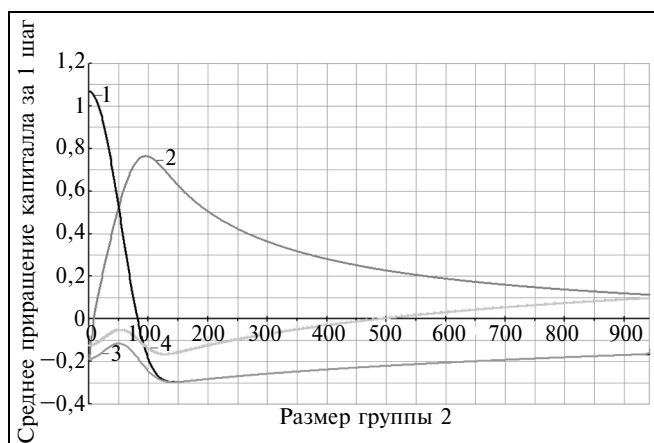


Рис. 2. Средние приращения капиталов члена группы (1 — группа 1, 2 — группа 2), эгоиста (3) и случайно выбранного участника (4) за 1 шаг: 1000 участников — 50 в группе 1, группа 2 и эгоисты; $\mu = -0,8$, $\sigma = 30$, $\alpha = 0,5$

этих условиях. Инструментами нашего анализа служат, с одной стороны, имитационная модель, реализованная с помощью компьютерной системы, с другой — аналитические выражения, аналогичные формулам (1) и не приводимые здесь из-за их громоздкости. На рис. 2 показана зависимость средних капиталов эгоистов, участников двух групп и общества от размера второй формирующейся группы.

Легко заметить, что рост группы 2 крайне невыгоден для группы 1: среднее приращение капитала участника группы 1 при этом стремительно убывает. Приращение капитала участника группы 2, напротив, растет и достигает максимума, когда в ней 96 членов. Этот максимум в данном примере составляет 72 % от максимума группы 1, наблюдаемого, когда группа 1 единственна. При дальнейшем расширении группы 2 путем присоединения к ней эгоистов преимущество членов группы 1 перед эгоистами полностью исчезает — это происходит, когда численность группы 2 приближается к 150. Продолжение присоединения эгоистов к группе 2 (и уменьшение ее «элитарности») выгодно для эгоистов, группы 1 и общества в целом. Группа 1 достигает минимума среднего приращения капитала, когда во второй группе 144 участника.

Таким образом, если единственной группе выгодно быть небольшой, то при конкуренции двух групп в данном примере побеждает группа, лидирующая по численности. Тем не менее, ее членам невыгодно, чтобы она превосходила меньшую группу более чем в 2—2,5 раза.

Отметим также следующую закономерность. При наличии двух групп эгоистам лучше всего, когда эти группы конкурируют «на равных»: сред-

нее приращение капитала эгоиста максимально при равенстве групп по численности. Политический прообраз данной ситуации — двухпартийность с равными по силе партиями. В рассматриваемом примере именно такая ситуация оптимальна для «неорганизованной» части общества. Механизм этого понятен: чем ближе партии по силе, тем важнее для них голоса «неорганизованных» граждан и, соответственно, тем сильнее эти голоса влияют на принимаемые решения. Если же одна из групп подавляюще превосходит другую, то в данном примере это даже чуть хуже для эгоистов, чем наличие одной группы суммарной численности.

5. ДВА «СНЕЖНЫХ КОМА», РАСТУЩИХ ВМЕСТЕ

В § 4 установлено, что для группы крайне невыгоден опережающий рост группы-конкурента. Вывод очевиден: при появлении второй группы первая должна не «замораживать» численность, а, наоборот, стараться сохранить превосходство в численности. Рассмотрим случай, когда группа 1 растет с той же абсолютной скоростью, что и группа 2, оставаясь на 50 членов больше. Соответствующий пример показан на рис. 3.

Как можно было предположить, группе 1 при такой тактике удастся сохранить свое доминирующее положение. Среднее приращение капитала ее участника при росте численности снижается в 5 раз, но остается значительно выше, чем приращения капиталов эгоистов и членов группы 2, не выходящие из отрицательной области. Общество при этом также выигрывает, так как лидирующая группа 1 становится менее «элитарной». Среднее

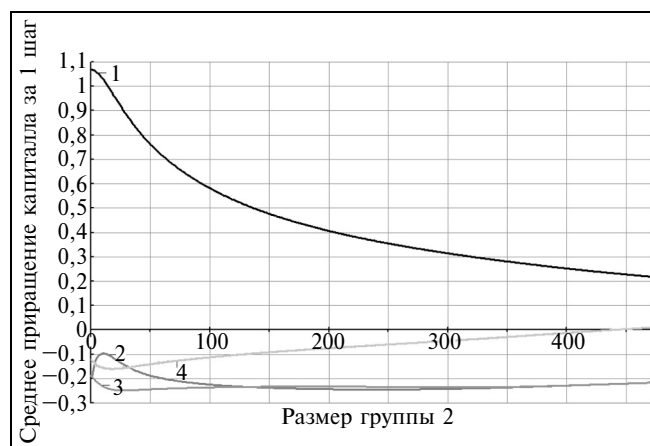


Рис. 3. Средние приращения капиталов члена группы (1 — группа 1, 2 — группа 2), эгоиста (3) и случайно выбранного участника (4) за 1 шаг: 1000 участников — группа 1 больше группы 2 на 50 участников; $\mu = -0,8$, $\sigma = 30$, $\alpha = 0,5$

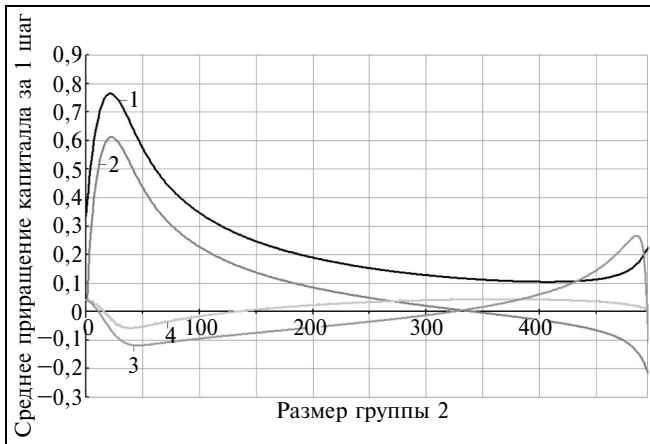


Рис. 4. Средние приращения капиталов члена группы (1 — группа 1, 2 — группа 2), эгоиста (3) и случайно выбранного участника (4) за 1 шаг:

1000 участников — группа 1 больше группы 2 на 5 участников;
 $\mu = -0,8, \sigma = 30, \alpha = 0,5$

приращение капитала участника группы 2 поначалу быстро растет, но очень скоро, достигнув максимума, примерно составляющего $-0,1$, при 10 членах группы, далее уменьшается и доходит до $-0,245$. На этом участке показатель группы 2 даже чуть ниже, чем у эгоистов и при дальнейшем росте группы 2 меняется слабо. По существу, группа 2 разоряется, и причина этого в том, что она практически всегда проигрывает при голосовании сильно превосходящей ее группе 1. Разница между ними в 50 голосов настолько существенна, что она лишь в редчайших случаях может компенсироваться голосами эгоистов.

Для сравнения, если разница между двумя группами составляет не 50, а 5 голосов, то динамика совершенно иная (рис. 4).

Помимо иного положения группы 2, здесь следует обратить внимание на кривую среднего приращения капитала эгоиста, а именно, на ее «взлет» и «падение» в правой части графика. Здесь мы наблюдаем проявление известного эффекта преимуществ «малой партии» при относительном равновесии сил больших. Эгоисты не являются сплоченной силой, но когда их мало, и ни одна из групп не может сама обеспечить большинство голосов, проходят в основном те предложения, которые поддержаны существенной частью эгоистов и, следовательно, выгодны для них. Этим объясняется «взлет» приращения среднего капитала эгоистов и их превосходство по отношению к обеим группам при низкой численности эгоистов. Но когда эгоистов остается меньше 5, все решения просто принимаются группой 1, поэтому ее кривая идет резко вверх, а кривые группы 2 и эгоистов — резко вниз. Характерно здесь и то, что минимум среднего при-

ращения капитала члена общества достигается при численности группы 2 (40 участников), примерно вдвое (но абсолютно — довольно умеренно) отличающейся от численностей, доставляющих максимум приращениям капиталов групп (22 и 23 участника для групп 1 и 2 соответственно — качественно примерно так же, как на рис. 1). Максимум общественного капитала наблюдается при наиболее острой конкуренции между группой 1, группой 2 и эгоистами — когда численность группы 2 равна 367. Когда в группе 2 становится 384 участника, показатель эгоиста «догоняет» показатель случайно выбранного члена общества. Полученные нами формулы позволяют выявить механизмы отмеченных здесь и описанных выше закономерностей, но объем настоящей статьи не позволяет остановиться на этом подробно.

6. ДВА «КЛУБКА ШЕРСТИ»

Рассмотрим теперь случай, когда перераспределение сил между двумя группами происходит внутри активной части общества, т. е. сумма численности двух групп и число эгоистов остаются неизменным. В примере, показанном на рис. 5, по горизонтальной оси откладывается численность группы 1, а численность группы 2 дополняет ее до 1000. Все общество состоит из этих двух групп и 500 эгоистов; значения остальных параметров: $\mu = -0,8, \sigma = 100, \alpha = 2/3$ (квалифицированное большинство).

Этот случай вызывает уже образ не снежного кома, а песочных часов или двух клубков шерсти с общей ниткой: то, что сматывается с одного клубка, наматывается на другой.

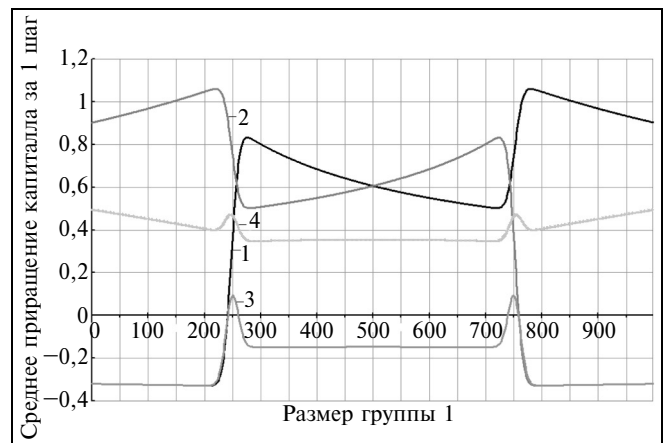


Рис. 5. Средние приращения капиталов члена группы (1 — группа 1, 2 — группа 2), эгоиста (3) и случайно выбранного участника (4) за 1 шаг:
 1500 участников, из них 1000 — в двух группах;
 $\mu = -0,8, \sigma = 100, \alpha = 2/3$

Прокомментируем эти немного неожиданные, на первый взгляд, результаты. Для принятия предложения нужен 1001 голос. Если в группе 1 более 780 членов, и она поддерживает предложение, то недостающие голоса (их нужно не более 220), как правило, обеспечивают эгоисты, а без поддержки группы 1 предложение не может быть принято. Поэтому в данном диапазоне группа 1 имеет высокие значения среднего приращения капитала. При увеличении численности группы 1 от 780 до 1000 среднее приращение капитала снижается, так как чем больше группа, тем реже она удовлетворена предложением и тем меньшими средними значениями приращения капитала выражается удовлетворенность. Если в группе 1 от 720 до 780 членов, то в случае, когда она поддерживает предложение, эгоисты часто не собирают недостающего числа голосов. Но если они его наберут, то предложение, весьма выгодное для них, будет принято. Поэтому на данном участке наблюдаем пик приращения капитала эгоистов. На отрезке [720, 750] численности группы 1 в случае, если она поддерживает предложение, недостающие голоса часто обеспечивает группа 2, что выгодно для нее. Поэтому приращение ее капитала здесь довольно значительно. На отрезке [750, 780] недостающие голоса все чаще дают эгоисты, поэтому приращение капитала группы 2 резко падает. На отрезке [280, 720] принимаются практически только предложения, одобренные обеими группами, причем меньшая группа обладает преимуществом, так как по закону больших чисел ее удовлетворенность в среднем выражается более высокими значениями приращения капитала. Этот участок более благоприятен для эгоистов, чем отрезок [780, 1000], так как здесь предложений, приносящих им отрицательное среднее приращение капитала, принимается меньше: ведь для их принятия здесь нужно согласие обеих групп, а на отрезке [780, 1000] — только группы 1. Участок [220, 280] симметричен участку [720, 780], а на участке [0, 220] принимаются фактически все предложения, выгодные группе 2. Поэтому здесь группа 1 и эгоисты находятся в неблагоприятном положении.

Отметим, что ситуация двухпартийности с равными по силе партиями обладает в этом примере большой устойчивостью. Действительно, если в одной группе 450, а в другой 550 членов, то участникам большей группы выгоден переход в меньшую, поскольку она имеет преимущество, и это приводит к сближению численностей групп. Вместе с тем, ситуация «равновесной двухпартийности» не является здесь наиболее выгодной для эгоистов и общества в целом. Дело в том, что из-за высокого порога голосования голоса эгоистов не позволяют одной из равных партий провести выгодное

ей предложение без одобрения его второй партией. Поэтому интересы эгоистов здесь практически не учитываются. Их голоса наиболее востребованы, когда численности групп примерно равны 750 и 250. Тогда для принятия предложения, выгодного в среднем для эгоистов, достаточно, чтобы оно было поддержано только большей группой, а для принятия предложения, в среднем невыгодного для них, нужна поддержка его обеими группами, что существенно менее вероятно. Поэтому эгоисты имеют максимумы приращения капитала, когда в первой группе 250 или 750 членов.

ЗАКЛЮЧЕНИЕ

Реальность несоизмеримо богаче рассмотренной простой модели; в частности, в реальности действует множество субъективных факторов, моделью не учитываемых. Вместе с тем, анализ модели позволяет выявить некоторые скрытые пружины социальных процессов. Модель представляет «базовый вариант», «нулевое приближение», в которых наиболее ярко проявляются многие значимые феномены, такие, например, как выгодность (в определенных условиях) двухпартийной системы с примерно равными партиями, преимущества малой партии в условиях конкуренции двух больших и ряд других. При известных соотношении сил и процедуре принятия решений в обществе, в парламенте, в коллективе может быть реализована рассмотренная модель, и ее анализ поможет понять специфику функционирования социальной системы. Результаты этого анализа не всегда легко спрогнозировать: так, даже немногие разобранные здесь примеры представляют довольно большое разнообразие типов функционирования систем, реализуемых при разных значениях параметров модели.

Центральная метафора настоящей статьи — «снежный ком» кооперации — относится к группе, эффективно защищающей интересы своих членов и открытой для присоединения новых участников. Расширяясь, такая группа становится все более альтруистичной.

Выясняется, однако, что для членов группы ее расширение с определенного момента может стать невыгодным, приводя к снижению средних приращений капитала. При этом интересы группы вступают в противоречие с интересами оставшейся части общества. Данный сценарий рассмотрен в § 3. Он имеет отношение ко многим реальным экономическим и политическим ситуациям, в частности, возможно, к остро обсуждаемой сейчас проблеме раздела природных богатств Арктики между претендующими на них государствами.



В случаях, когда члены группы не заинтересованы в ее расширении, естественным образом начинается формирование конкурирующей группы. При этом, как показано в § 4, для исходной группы крайне невыгодно пытаться «заморозить» свою оптимальную численность. Напротив, чтобы сохранить лидирующие позиции, она должна расти, причем не слишком медленно и не слишком быстро (см. § 5). При наличии двух больших групп близкой численности в выигрыше могут оказаться «неорганизованные» члены общества. Что касается общества в целом, для него наиболее выгоден консенсус, а если он невозможен, то наиболее выгодны ситуации, характеризующиеся реальной конкуренцией, в которую вовлечено максимально возможное число участников. Этот вывод необходимо учитывать при разработке алгоритмов управления в социально-экономических системах.

Наконец, в § 6 проанализированы некоторые закономерности «перетока» участников между группами. При этом парадоксальным образом может оказаться, что принадлежать к меньшей группе выгоднее, чем к большей. Там, где данный феномен проявляется, он обеспечивает устойчивость «двухпартийной» системы с практически равными партиями. Для представителей «неорганизованной» части общества выгоднее всего ситуации, в которых их голоса наиболее востребованы. Показано, что это не обязательно ситуации двухпартийности с равными партиями: некоторые неравновесные ситуации могут быть для эгоистов более выгодными.

Данная работа не содержит систематического описания всех типов социальной динамики, реализуемых в модели — в рамках небольшой статьи сделать это нельзя. Нашей задачей было рассмотрение лишь нескольких ярких примеров, имеющих отношение к реальности.

Возвращаясь к «снежному кому» кооперации, отметим, что есть немало случаев, когда уже вошедшие в альянс участники совсем не заинтересованы блокировать расширение этого альянса. В частности, это относится, судя по всему, к международному сотрудничеству по решению глобальных проблем, таких, как проблема изменения климата. Участники соответствующего сообщества реализуют совместные проекты в интересах альянса и не в ущерб третьим странам и весьма заинтересованы в расширении географии сотрудничества. В тех же ситуациях, когда альянсу («элите») выгодно ограничивать свой размер, принципиальная открытость альянса может, вообще говоря, устанавливаться нормативно, поскольку это отвечает интересам общества в целом. Во всех перечисленных случаях механизм «снежного кома» кооперации

может быть использован для построения устойчивых «условно-альтруистических» сообществ (о них шла речь в § 3). Этот механизм может быть особенно полезен при выстраивании структур гражданского общества, что ощущается сегодня как насущная необходимость (см., например, статью [14]). Так, ячейки гражданского общества выигрывают в случае вхождения в ассоциации, которые характеризуются определенным уровнем солидарности и развивают механизмы координации совместных действий.

ЛИТЕРАТУРА

1. *Margolis H.* Selfishness, altruism and rationality: a theory of social choice. — Cambridge: Cambridge University Press, 1982.
2. *Levine D.K.* Modeling altruism and spitefulness in experiments // *Review Economic Dynamics*. — 1998. — Vol. 1. — P. 593—622.
3. *Lindenberg S.* Social rationality versus rational egoism / *Handbook of sociological theory* / Turner J. (ed.). — New York: Kluwer Academic/Plenum, 2001. — P. 635—668.
4. *Romer T.* Individual welfare, majority voting, and the properties of a linear income tax // *J. Public Economics*. — 1975. — Vol. 4. — P. 163—185.
5. *Roberts K.W.S.* Voting over income tax schedules // *J. Public Economics*. — 1977. — Vol. 8. — P. 329—340.
6. *Kranich L.* Altruism and the political economy of income taxation // *J. Public Economic Theory*. — 2001. — Vol. 3, N 4. — P. 455—469.
7. *Galasso V., Profeta P.* The political economy of social security: a survey // *Eur. J. Political Economy*. — 2002. — Vol. 18, N 1. — P. 1—29.
8. *Миркин Б.Г.* Проблема группового выбора. — М.: Наука, 1974.
9. *McKelvey R.D.* Intransitivities in multidimensional voting models and some implications for agenda control // *J. Economic Theory*. — 1976. — Vol. 12. — P. 472—482.
10. *Чеботарев П.Ю., Кочубей Б.А.* Маленькие хитрости демократии // *Век XX и мир*. — 1990. — № 6. — С. 2—3.
11. *Стратегии при голосовании в стохастической среде: эгоизм и коллективизм* / В.И. Борзенко, З.М. Лезина, А.К. Логинов и др. // *Автоматика и телемеханика*. — 2006. — № 2. — С. 154—173.
12. *Чеботарев П.Ю.* Аналитическое выражение ожидаемых значений капиталов при голосовании в стохастической среде // *Автоматика и телемеханика*. — 2006. — № 3. — С. 152—165.
13. *Кооперация и эгоизм как стратегии голосования в стохастической среде* / П.Ю. Чеботарев, В.И. Борзенко, З.М. Лезина и др. // *Третья Междунар. конф. по проблемам управления. Пленарные докл. и избр. тр. / Ин-т пробл. упр.* — М., 2006. — С. 304—311.
14. *Жордан И.* Тот, кто будет первым // *На злобу*. — 2007. — № 9. (<http://www.nazlobu.ru/publications/print2188.htm>).

☎ (495) 334-88-69; e-mail: pchv@rambler.ru

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым. □

ПРИМЕНЕНИЕ ОБОБЩЕННЫХ МЕДИАННЫХ СХЕМ ДЛЯ ПОСТРОЕНИЯ НЕМАНИПУЛИРУЕМЫХ МЕХАНИЗМОВ МНОГОКРИТЕРИАЛЬНОЙ АКТИВНОЙ ЭКСПЕРТИЗЫ

В.Н. Бурков, М.Б. Исаков, Н.А. Коргин

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Механизмы многокритериальной активной экспертизы представлены в виде обобщенных медианных схем принятия коллективных решений в терминах семейств систем правых и левых коалиций, что позволило применить результаты теории коллективного выбора в области построения неманипулируемых механизмов.

ВВЕДЕНИЕ

Многие управленческие решения в современных условиях принимаются на основании результатов экспертизы. Нередко складываются ситуации, когда эксперты заинтересованы в результате экспертизы, что создает предпосылки для так называемого оппортунистического поведения с их стороны — искажения сообщаемой информации с целью влияния на результат в собственных интересах. В теории организационных систем подобные ситуации моделируются в рамках активной экспертизы. Механизмы активной экспертизы (правила определения итоговой оценки на основании заявок экспертов) называются неманипулируемыми, если они побуждают экспертов честно сообщать свое мнение.

Задачи экспертизы, при решении которых эксперты должны производить оценку по нескольким критериям одновременно, называются многокритериальными задачами экспертизы. Если эксперты, участвующие в подобной экспертизе, заинтересованы в ее результатах, то это будет задача многокритериальной активной экспертизы. Построение многокритериальных неманипулируемых механизмов активной экспертизы представляется достаточно актуальной задачей, так как обеспечение объективности результатов экспертизы всегда крайне важно. Отдельную сложность в задачах экспертизы по многим критериям составляет наличие нетривиальной связи между значениями различных критериев — на диапазон возможных значений отдельно взятого критерия могут влиять

значения остальных критериев. Наиболее простой пример многокритериальной экспертизы со связанными значениями критериев и поэтому применяемый в качестве иллюстрации в данной работе представляет собой задача распределения ресурсов, где каждый эксперт сообщает наилучшее с его точки зрения распределение ресурса между всеми потребителями. Другой пример — задача формирования бюджета или сметы на портфель проектов.

На данный момент задача построения неманипулируемого механизма активной экспертизы решена для однопиковых функций предпочтения экспертов и частных видов множества возможных сообщений (и наиболее предпочтительных с их точки зрения результатов экспертизы) экспертов:

отрезка (одномерного компакта) [1—3] — при моделировании экспертизы по одному критерию; двумерного симплекса [4, 5] — при моделировании экспертизы по двум критериям, связанным балансовым ограничением¹;

многомерного прямоугольника (декартова произведения одномерных отрезков) [5] — при моделировании многокритериальной экспертизы с несвязанными критериями, когда множество возможных значений каждого отдельно взятого критерия не зависит от конкретных значений остальных критериев.

Для более сложных видов множества возможных заявок экспертов задача построения немани-

¹ Под балансовым ограничением подразумевается ограничение типа равенства, когда сумма значений оценок по всем критериям должна быть постоянной.



пулируемых механизмов активной экспертизы остается нерешенной.

Задача построения неманипулируемых механизмов активной экспертизы в значительной мере пересекается с задачами поиска неманипулируемых функций коллективного выбора в теории коллективного выбора [3], в частности, функций, определяющих наилучший с точки зрения общества уровень коллективного блага² [6—8]. Один из важных результатов теории коллективного выбора заключается в *медианной схеме* (median voter scheme) — методе построения неманипулируемых функций коллективного выбора, в котором к n реальным голосующим членам общества (агентам) добавляется $n - 1$ «фантомный» агент, каждый из которых имеет собственный предпочтительный уровень коллективного блага. И все агенты, реальные и фантомные, упорядочиваются по значению их предпочтительного уровня коллективного блага. Результирующий уровень коллективного блага определяется как значение, предпочтительное с точки зрения агента, который находится в середине списка (служит медианой полученного упорядочения). Причем, варьируя значения предпочтительного уровня коллективного блага фантомных агентов, можно обеспечить любое из допустимых результирующих значений между минимальным и максимальным предпочтительными уровнями блага реальных агентов. Изначально данный метод был получен в работе [8] для одномерного случая, когда значение результирующего уровня коллективного блага и предпочтения агентов принадлежат пространству \mathcal{R}^1 . Затем полученные результаты были расширены на \mathcal{R}^m [7], т. е. когда множеством возможных значений альтернатив (предпочтительного уровня коллективного блага каждого из агентов) и результата коллективного выбора являются точки m -мерного пространства. Неманипулируемые функции коллективного выбора в пространстве \mathcal{R}^m были представлены в виде *обобщенных медианных схем* — набора из t одномерных медианных схем. Наконец, полученный результат был расширен на случай, когда соответствующее множество имело вид ограниченного замкнутого подмножества в пространстве \mathcal{R}^m [6, 9]. Для этого была предложена альтернативная формулировка медианных схем в терминах так называемых систем правых или левых коалиций агентов и введено дополнительное свойство — *свойство пересечения*. Обладание обобщенной медианной схемой свойством пересечения является необходимым и достаточным условием того, что результат общест-

венного выбора, полученный в рамках данной схемы, будет всегда принадлежать множеству возможных значений альтернатив. Соответственно, было показано, что функция коллективного выбора в некотором компактном подмножестве пространства \mathcal{R}^m неманипулируема тогда и только тогда, когда она представима в виде обобщенной медианной схемы, обладающей свойством пересечения.

Цель настоящей работы заключается в представлении механизмов многокритериальной активной экспертизы в виде обобщенных медианных схем с целью применения результатов теории коллективного выбора для построения неманипулируемых многокритериальных механизмов активной экспертизы.

1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ МНОГОКРИТЕРИАЛЬНОЙ АКТИВНОЙ ЭКСПЕРТИЗЫ

В качестве примера многокритериальной активной экспертизы рассмотрим задачу распределения ресурса, в которой каждый агент (эксперт) сообщает не собственную заявку — сколько ресурсов (например, денежных) ему требуется для его проекта³, а распределение ресурсов по всем проектам, наилучшее с его точки зрения. Число проектов может не совпадать с числом экспертов.

Приведем формальное описание данной задачи. Организатору экспертизы — центру (Π) — необходимо распределить ресурс в количестве C между набором проектов $M = \{1, \dots, m\}$. Результирующее распределение ресурсов $x = (x_1, \dots, x_m)$ (x_j — ресурс, выделяемый на проект j) по проектам определяется на основании сообщений экспертов: $s^i = (s^i_1, \dots, s^i_m)$ — сообщение эксперта i , $\forall i \in N$, $s^i_j \in A$, где $N = \{1, \dots, n\}$ — множество экспертов; $A = \{x \in \mathcal{R}^m \mid \sum_{j=1}^m x_j \leq C, x_j \geq 0, \forall j \in M\}$ — множество возможных значений результатов экспертизы, определяемое как неотрицательное подпространство m -мерного пространства действительных чисел \mathcal{R}^m , ограничиваемое бюджетным ограничением C . В данной статье речь пойдет о любых компактных подмножествах пространства \mathcal{R}^m . Однако в качестве примера, иллюстрирующего получаемые результаты, будет взято именно подобное множество.

Заинтересованность каждого эксперта в результате экспертизы отражается его многомерно однопиковой функцией предпочтения.

² Коллективное благо — благо, уровень которого влияет на полезность всех членов общества. Например, результат экспертизы в случае, когда все эксперты заинтересованы в нем, является коллективным благом для них.

³ Под проектом мы будем понимать работы, в выполнении которых заинтересован данный эксперт.

Определение 1 [6]. Функция предпочтения $u^i(x)$: $\mathfrak{R}^m \rightarrow \mathfrak{R}$ является *многомерно однопиковой* на множестве $A \subseteq \mathfrak{R}^m$, если:

- существует единственная точка $\tau^i = \arg \max_{y \in A} u^i(y)$ — точка «пики» эксперта i , $\forall i \in N \tau^i \in A$;
- $\forall z, z' \in \mathfrak{R}^m [z' \in \widehat{B}(\{z, \tau^i\}), z' \neq z] \Rightarrow [u^i(z') > u^i(z)]$, где $\widehat{B}(\{z, \tau^i\})$ — минимальный m -мерный прямоугольник, содержащий точки z и τ^i , определяемый следующим образом. Для произвольного множества $A \subseteq \mathfrak{R}^m$ обозначим через A_k проекцию данного множества на координатную ось $k \in M$, т. е. $A_k = \text{Proj}_k(A)$. Нижнюю и верхнюю границы множества A обозначим $\min A_k$ и $\max A_k$ соответственно. Тогда минимальный m -мерный прямоугольник, содержащий множество $A \subseteq \mathfrak{R}^m$, определяется как

$$\widehat{B}(A) = \prod_{k \in M} [\min A_k, \max A_k]. \blacklozenge$$

В рамках рассматриваемой модели точка τ^i трактуется как наиболее предпочтительный результат экспертизы (распределение ресурсов по проектам) с точки зрения эксперта i .

Примером однопиковой функции полезности эксперта может служить функция от нормы разности точки пика и результата экспертизы:

$$u^i(x) = \text{const} - \|\tau^i - x\|^\alpha, \quad \alpha > 0,$$

или квадратичная сепарабельная функция:

$$u^i(x) = \text{const} - \sum_{k=m} (\tau_k^i - x_k)^{2\alpha}, \quad \alpha > 0.$$

Организатор экспертизы (центр) определяет результаты экспертизы на основании сообщений экспертов по заранее объявленной процедуре (механизму) экспертизы $\pi(s): \mathfrak{R}^m \times \mathfrak{R}^n \rightarrow \mathfrak{R}^m$, где:

$s^i = (s_1^i, \dots, s_m^i)$ — сообщение эксперта i о распределении ресурса по всем проектам, причем $\forall i \in N s^i \in A$;

$s = (s^1, \dots, s^m)$ — вектор сообщений экспертов, $s \in A^n$;

$s_j = (s_j^1, \dots, s_j^n)$ — сообщения всех экспертов по j -му проекту.

Будем работать с механизмами активной экспертизы, удовлетворяющими следующим требованиям [1].

А.1. Функция $\pi(s)$ строго монотонна по всем переменным при $s \in A^n$.

А.2. Функция $\pi(s)$ непрерывна по всем переменным при $s \in A^n$.

А.3. Если обозначить $s^a = (a, \dots, a)$, $a \in A$, то $\pi(s^a) = a$ (условие единогласия).

Примером механизма, удовлетворяющего требованиям А.1—А.3 является механизм пропорционального распределения ресурса:

$$x = \pi(s): \forall j \in M x_j = \frac{1}{n} \sum_{i=1}^n s_j^i. \quad (1)$$

При применении данного механизма бюджет, выделяемый на каждый проект, будет определяться в результате усреднения сообщений всех экспертов по данному проекту. Отметим, что данный механизм является сепарабельным по критериям, т. е. в данном примере бюджет каждого проекта определяется независимо от бюджетов других проектов.

Определение 2. Механизм многокритериальной экспертизы $x = \pi(s)$ является *сепарабельным по критериям*, если $\forall k \in M x_k = \pi_k(s_k)$. \blacklozenge

Таким образом, модель многокритериальной активной экспертизы описывает выбор результирующего значения экспертизы из некоторого множества возможных значений A на основании сообщений экспертов (принадлежащих множеству A), заинтересованных в результате экспертизы (что отражается их однопиковой функцией полезности с точкой пика из множества A) по процедуре, отвечающей требованиям монотонности, непрерывности и единогласия.

Как было показано, например, в работе [1], применение подобных механизмов дает возможность экспертам манипулировать своими сообщениями, решая задачу максимизации собственной полезности. Вследствие этого возникает игра экспертов, так как результат экспертизы зависит от всех сообщений. В зависимости от выбранной концепции равновесия, в игре экспертов сформируется вектор их равновесных сообщений⁴ $s^*(\tau) \in A^n$, не обязательно совпадающий с точкой τ . В данной модели будем полагать, что равновесные сообщения определяются на основе концепции равновесия Нэша.

Мы еще не уточнили, как связано сообщение каждого эксперта с его точкой пика. Изначально, сообщение может быть произвольным, хотя мы и предположили, что область его допустимых значений совпадает с областью допустимых значений точек пиков экспертов и результата экспертизы — т. е. сообщение эксперта должно быть в тех же терминах, в каких определяется и результат экспертизы. Например, для задачи распределения ресурсов можно спросить у эксперта, сколько ресурсов

⁴ Равновесное сообщение определяется на основании принятых в модели предположений о поведении игроков (экспертов) и концепции равновесия. Отметим, что такого сообщения может не быть вообще или их может быть несколько. Если равновесий несколько, то необходимо ввести правило отбора равновесий, позволяющее из любого множества равновесий выбрать единственное.



нужно выделить на каждый из проектов, в то время как вопрос о его точке пика должен будет иметь другой вид — какое распределение наилучшее с точки зрения эксперта? Подобный механизм называется *непрямым* [2]. Для каждого непрямого механизма можно предложить *соответствующий ему прямой механизм*, определив его как $h(\tau) = \pi(s^*(\tau))$, $\forall \tau \in A^n$. Если в соответствующем прямом механизме сообщение достоверной информации является равновесием в доминантных стратегиях, то такой механизм называется *эквивалентным прямым механизмом*.

Сформулировав математическую модель многокритериальной активной экспертизы, перейдем к постановке задачи построения неманипулируемого многокритериального механизма активной экспертизы.

2. НЕМАНИПУЛИРУЕМОСТЬ МЕХАНИЗМОВ МНОГОКРИТЕРИАЛЬНОЙ АКТИВНОЙ ЭКСПЕРТИЗЫ: ОБЗОР РЕЗУЛЬТАТОВ

Прежде чем приступить к построению неманипулируемого механизма многокритериальной активной экспертизы, перечислим результаты, на которые мы будем опираться при решении данной задачи.

Первоначально, см. работу [1], для однокритериальной (одномерной) задачи активной экспертизы было доказано, что для любого механизма, удовлетворяющего требованиям А.1—А.3, существует эквивалентный прямой механизм. Соответствующая модель активной экспертизы описывается следующим образом. Пусть τ_i — точка пика i -го эксперта (его тип), $i \in N$ — множество всех экспертов, $\tau^i \in [d, D] \subset R^1$, $i \in N$, $-\infty < d < D < +\infty$. Экспертам известна процедура $\pi: [d, D]^n \rightarrow [d, D]$ определения результата экспертизы на основе их сообщений $s^i \in [d, D]$, $i \in N$: $x = \pi(s)$. Функции полезности экспертов $u(z)$ — однопиковые с точками пика τ^i , $i \in N$:

$$\forall z, z' \in [d, D] \cdot [z' \in (z, \tau^i), z' \neq z] \Rightarrow [u^i(z') > u(z)].$$

Тогда, если функция (процедура) выбора результатов экспертизы $\pi(s)$ удовлетворяет требованиям А.1—А.3, то для нее найдется эквивалентный прямой механизм $h_\pi(\tau)$, определяемый условием

$$x = \max_{i \in N} \min(s^i, w^{i-1}), \quad (2)$$

где эксперты упорядочены в порядке возрастания точек пика, а $w^k = \pi(s(k))$, и $s(k)$ определяется следующим образом:

$$s(k) = \begin{cases} k \text{ первых экспертов сообщают } d, k = \overline{0, n}; \\ (n - k) \text{ последних экспертов сообщают } D. \end{cases}$$

Элементы убывающей последовательности $w^k = \pi(s(k))$ могут рассматриваться как точки пиков $(n + 1)$ фантомных агентов.

Например, для механизма $x = \pi(s)$: $x = \frac{1}{n} \sum_{i=1}^n s^i$ последовательность точек w_k разбивает отрезок $[d, D]$ на n равных отрезков:

$$w^k = (kD + (n - k)d)/n, \quad k = \overline{0, n}. \quad (3)$$

Данный результат в значительной мере пересекается с результатами, полученными Э. Муленом [8], который доказал, что для задачи выбора уровня коллективного блага из упорядоченного множества A , в том числе, отрезка \mathcal{R}^1 , справедливо следующее утверждение.

Функция коллективного выбора на профиле⁵ однопиковых предпочтений над упорядоченным множеством неманипулируема тогда и только тогда, когда она является медианной схемой.

Все медианные схемы строятся по следующему принципу. Каждой группе агентов $S \subset 2^N \setminus \emptyset$ ставится в соответствие некоторая точка $a_S \in A$ (точнее, фантомный агент с точкой пика a_S), а результат выбора определяется как

$$h(\tau) = \inf_{S \subset N} (\sup_{i \in S} (a_S, \tau^i)).$$

Нетрудно показать, что условие (2) может быть получено подобным образом. Однако медианные схемы просто описывают весь класс прямых неманипулируемых механизмов, не позволяя определить механизм, оптимальный по какому-либо критерию или соответствующий произвольному непрямому механизму коллективного выбора. В то же время в работе [1] решались задачи построения прямых неманипулируемых механизмов активной экспертизы, оптимальных по различным критериям, в частности, по критерию максимальной погрешности процедуры экспертизы $\Delta_\pi = \max_{\tau \in [d, D]^n} |h_\pi(\tau) - \pi(\tau)|$, где функция $h_\pi(\tau)$ определялась выражением (2).

Важным результатом, полученным Э. Муленом, был тот факт, что значение функции коллективного выбора, являющейся медианной схемой, зависит только от точек пика функций предпочтения агентов, т. е. результат *определяется по точкам пика*⁶.

На основании результатов, полученных при решении описанной задачи однокритериальной активной экспертизы были предложены методы

⁵ Под профилем предпочтений подразумевается вектор функций предпочтений агентов.

⁶ «tops-only» в оригинале.

построения неманипулируемых механизмов активной экспертизы для случая двухкритериальной экспертизы, когда множество допустимых результатов экспертизы является двумерным симплексом $A = \{x \in A \subset \mathbb{R}^2: x_1 + x_2 = \text{const}\}$. Было показано [4], что подобные задачи экспертизы могут быть сведены к однокритериальным задачам.

Результаты, полученные Э. Муленом, были расширены К. Бордером и Д. Джорданом [7] на многомерные задачи коллективного выбора, когда множеством допустимых результатов выбора является пространство \mathbb{R}^m , или декартово произведение отрезков из пространства \mathbb{R}^1 : $A = \prod_{k \in M} A_k$, $A_k = [d_k, D_k] \subset \mathbb{R}^1$ — m -мерный (далее многомерный) прямоугольник. В работе [7] рассматривались так называемые звездообразные сепарабельные функции предпочтения агентов, обладающие одной точкой пика и определяемые следующим образом:

- существует единственная точка $\tau^i = \arg \max_{y \in A} u^i(y)$ — точка «пика» эксперта i , $\forall i \in N$, $\tau^i \in A$;
- $\forall z \in \mathbb{R}^m$, $\forall \lambda \in (0, 1)$, $u^i(\tau^i) > u^i(\lambda \tau^i + (1 - \lambda)z) > u^i(z)$;
- $\forall k \in M$, $\forall z_k, z'_k, \tilde{z}_{-k}, \hat{z}_{-k}$ [$u^i(z'_k, \hat{z}_{-k}) > u^i(z_k, \tilde{z}_{-k})$] \Rightarrow [$u^i(z'_k, \tilde{z}_{-k}) > u^i(z_k, \hat{z}_{-k})$].

При перечисленных допущениях было доказано, что

функция коллективного выбора на профиле звездообразных сепарабельных предпочтений над множеством \mathbb{R}^m неманипулируема тогда и только тогда, когда она представима в виде набора одномерных медианных схем.

Иными словами, многомерная задача коллективного выбора декомпозируется на набор одномерных, для каждой из которых подбирается своя медианная схема. Такой набор медианных схем получил название *обобщенной медианной схемы*. Причем результат коллективного выбора в каждой одномерной задаче определяется проекциями точек пика агентов на соответствующую ось. Формально, каждой коалиции агентов $S \subset 2^N \setminus \emptyset$ ставится в соответствие точка $a_S = (a_{1S}, \dots, a_{mS}) \in A$ коллективного выбора, а результат выбора определяется как $x = (x_1, \dots, x_m)$, где $x_k = \pi(\tau_k) = \inf_{S \subset N} (\sup_{i \in S} (a_{kS}, \tau_k^i))$, $\forall k \in M$.

Данный результат был использован в работе [5] для построения неманипулируемого механизма многокритериальной активной экспертизы для аналогичного вида множества допустимых значений результатов экспертизы — было показано, что подобные задачи многокритериальной экспертизы могут рассматриваться как набор задач однокрите-

риальной экспертизы — по каждому критерию экспертиза может быть проведена независимо.

Однако остается открытым вопрос о том, как решить задачу построения неманипулируемых функций коллективного выбора для множества допустимых результатов выбора более сложного вида, чем многомерный прямоугольник.

В качестве примера рассмотрим задачу распределения бюджета между тремя проектами ($m = 3$) на основании заявок двух экспертов ($n = 2$) в рамках модели, описанной в § 1. Результат экспертизы определяется на основании функции усреднения заявок экспертов по каждому проекту, определяемой выражением (1). Множество допустимых результатов экспертизы A — единичный куб — $x_i \in [0, 1]$, $\forall i \in \{1, 2, 3\}$. Точки пика функций предпочтения экспертов — $(1, 0, 0)$ и $(0, 1/2, 1/2)$. Легко получить, учитывая формулу (3), что результатом экспертизы будет точка $(1/2, 1/2, 1/2)$. Однако если мы наложим бюджетное ограничение на распределяемый ресурс $x_1 + x_2 + x_3 \leq 1$, то полученный нами ранее результат экспертизы не будет принадлежать множеству допустимых результатов экспертизы, так как бюджетное ограничение не будет выполнено. Иными словами, данный механизм не обладает *свойством реализуемости* — его применение не гарантирует, что полученный результат будет принадлежать множеству допустимых значений A даже при условии, что точки пиков всех экспертов принадлежат этому множеству.

Результаты К. Бордера и Д. Джордана [7] были расширены группой ученых под руководством С. Барберы, обзор результатов которой представлен в работах [6, 9], на более сложный вид множеств допустимых результатов коллективного выбора — произвольное компактное подмножество \mathbb{R}^m , при условии, что функции предпочтения агентов многомерно однопиковые в смысле определения 1. Для этого была предложена альтернативная формулировка обобщенных медианных схем — в терминах *систем правых и левых коалиций*. Приведем необходимые определения.

Определение 3 [6]. Система правых (левых)⁷ коалиций W определяет для каждой точки $z \in [d, D] \subset \subset \mathbb{R}^1$ набор коалиций $W(z)$ в соответствии со следующими требованиями.

1. Принцип суверенитета: для $\forall z \in (d, D)$ ($[d, D)$), $W(z) \neq \emptyset$, $\emptyset \notin W(z)$, и $W(d) = 2^N \setminus \emptyset$ ($W(D) = 2^N \setminus \emptyset$).
2. Монотонность коалиций: если $S \in W(z)$ и $S \subset S'$, то $S' \in W(z)$.
3. Монотонность результата: если $z' < (>) z$ и $S \in W(z)$, то $S \in W(z')$.

⁷ В данном определении и далее в скобках идет запись соответствующего условия для системы левых коалиций.



4. Полунепрерывность сверху: для $\forall S \subseteq N$, для $\forall z \in [d, D]$ и для любой последовательности $\{z^t\} \subseteq [d, D]$ такой, что $\lim_{t \rightarrow \infty} z^t = z$, верно $[\forall t, S \in W(z^t)] \Rightarrow [S \in W(z)]$. \blacklozenge

Иными словами, вводится функция $W(\cdot): [d, D] \rightarrow 2^N \setminus \emptyset$, ставящая в соответствие каждой точке $z \in [d, D]$ некоторое подмножество множества всех непустых коалиций агентов $2^N \setminus \emptyset$. Для минимального многомерного прямоугольника, содержащего множество допустимых результатов выбора A , $\widehat{B}(A) = \prod_{k=1}^m A_k$, можно определить *семейство систем правых коалиций* R — набор $\{R_k\}_{k=1}^m$, где R_k является системой правых коалиций. Аналогичным образом можно определить *семейство систем левых коалиций* L — набор $\{L_k\}_{k=1}^m$. Используя данные понятия, вводится следующее формальное определение обобщенной медианной схемы.

Определение 4 [6]. Пусть задано множество допустимых результатов выбора A и семейство систем правых (левых) коалиций R (L) на множестве $\widehat{B}(A)$. Тогда обобщенная медианная схема $x = h(\tau)$, порожденная совокупностью (A, R) (соответственно (A, L)), определяется следующим образом: для $\forall \tau \in A^n$ и любого $k \in M$

$$h_k(\tau_k) = \max\{z_k \in A_k \mid \{i \in N \mid \tau_k^i \geq z_k\} \in R_k(z_k)\},$$

$$(h_k(\tau_k) = \min\{z_k \in A_k \mid \{i \in N \mid \tau_k^i \leq z_k\} \in L_k(z_k)\}). \blacklozenge$$

Иными словами, в качестве результата экспертизы по каждому критерию выбирается максимальное (минимальное) значение $z_k \in A_k$, для которого эксперты, чьи точки пика лежат справа (слева) от данной точки, составляют коалицию, удовлетворяющую требованиям выбранной системы правых (левых) коалиций для данной точки. Связь между системами левых и правых коалиций, порождающими одну и ту же медианную схему (для одного критерия), определялась следующим образом: $L^*(z) = \{S \in 2^N \mid \forall z' > z, \forall S' \in R(z'), S \cap S' \neq \emptyset\}$, т. е. для любой точки $\forall z' > z$ найдется хотя бы один агент, входящий одновременно и в правую для z , и в левую для z' коалиции (и это верно для всех соответствующих правых и левых коалиций), удовлетворяющие требованиям соответствующих систем коалиций, порождающих один и тот же механизм.

Обобщенные медианные схемы определяются на множестве $\widehat{B}(A)$, т. е. на m -мерном прямоугольнике. Поэтому любой прямой неманипулируемый механизм коллективного выбора, определенный в терминах системы правых или левых коалиций для

произвольного m -мерного прямоугольника B , останется таковым для любого множества допустимых результатов выбора A , такого что $\widehat{B}(A) = B$. Однако, как было показано ранее, не для всех A верно, что для $\forall \tau \in A^n$, $h(\tau) \in A$. Для проверки реализуемости прямого неманипулируемого механизма для порождающего его семейства систем правых коалиций $R = \{R_k\}_{k=1}^m$ и семейства систем левых коалиций $L^* = \{L_k^*\}_{k=1}^m$ было сформулировано следующее свойство пересечения.

Определение 5 [6]. Семейство систем правых коалиций $R = \{R_k\}_{k=1}^m$, определенное на множестве $\widehat{B}(A)$, обладает *свойством пересечения* для A , если для $\forall y \in \widehat{B}(A) \setminus A$ и любого конечного подмножества $\{z^1, \dots, z^T\} \subset A$

$$\bigcap_{i=1}^T \left\{ \left[\bigcup_{k \in M^+(y, z^i)} l_k(y_k) \right] \cup \left[\bigcup_{k \in M^-(y, z^i)} r_k(y_k) \right] \right\} \neq \emptyset, \quad (4)$$

для любой коалиции $r_k(y_k) \in R_k(y_k)$, где $k \in \bigcup_{i=1}^T M^-(y, z^i)$ и любой коалиции $l_k(y_k) \in L_k^*(y_k)$, где $k \in \bigcup_{i=1}^T M^+(y, z^i)$. \blacklozenge

Множества $M^+(y, z) = \{k \in M \mid z_k > y_k\}$ и $M^-(y, z) = \{k \in M \mid z_k < y_k\}$ определяют для любой пары векторов $y, z \in \widehat{B}(A)$ набор координатных осей, для которых значение соответствующей компоненты вектора z строго больше y и наоборот.

Качественно, наличие свойства пересечения у семейства систем правых коалиций означает, что любая точка $y \in \widehat{B}(A) \setminus A$ может быть выбрана в результате применения прямого неманипулируемого механизма только в том случае, если найдется хотя бы один агент, чья точка пика также не принадлежит множеству A . Если семейство систем правых коалиций обладает свойством пересечения на множестве A , то порождаемая им обобщенная схема удовлетворяет свойству пересечения на этом же множестве.

Определение свойства пересечения достаточно сложно на первый взгляд, но может быть упрощено — в частности, условие (4) достаточно проверять лишь для одного *решающего* множества $\widehat{S}(y) \subset A$ для каждой точки $y \in \widehat{B}(A) \setminus A$. Мы не будем подробно излагать общий алгоритм построения решающего множества, описанный в работе [9], приведя лишь результат, актуальный для рассматриваемых нами задач:

для выпуклых компактных множеств A для $\forall y \in \widehat{B}(A) \setminus A$ $\widehat{S}(y) = \{z^1, \dots, z^m\} \subset A$, где для $\forall t = 1, \dots, m$

$z_k^t = y_k$, если $k = t$, а z_k^t для $k \neq t$ любые такие, что $z^t \in A$.

Введя все необходимые определения, приведем основной результат, полученный группой С. Барберы:

функция коллективного выбора на профиле многомерно однопиковых предпочтений над множеством допустимых результатов голосования $A \subseteq \mathbb{R}^m$ неманипулируема тогда и только тогда, когда она представима в виде обобщенной медианной схемы, удовлетворяющей свойству пересечения.

По сути, было доказано, что переход от множества допустимых результатов голосования, представимых в виде многомерного прямоугольника, к более сложным множествам (единственное требование, предъявляемое к ним — компактность и «полноразмерность»⁸) не изменяет формы представления неманипулируемых функций коллективного выбора, а лишь добавляет новое условие — свойство пересечения. Удовлетворение этому свойству какого-либо набора одномерных медианных схем гарантирует, что для любого профиля функций предпочтения агентов результат коллективного выбора будет принадлежать множеству допустимых значений. Там же было доказано, что для новой формулировки медианной схемы результаты, полученные Э. Муленом [8], К. Бордером и Д. Джорданом [7], остаются в силе.

Перечислив необходимые результаты, полученные в работах [6, 9] можно перейти к решению задачи построения неманипулируемых механизмов многокритериальной активной экспертизы. Для этого требуется представить задачу однокритериальной и многокритериальной активной экспертизы в терминах семейств систем правых и левых коалиций и проанализировать получаемые медианные схемы на предмет выполнения свойства пересечения.

3. ФОРМУЛИРОВКА МЕХАНИЗМА МНОГОКРИТЕРИАЛЬНОЙ АКТИВНОЙ ЭКСПЕРТИЗЫ В ТЕРМИНАХ ЛЕВЫХ И ПРАВЫХ КОАЛИЦИЙ

Для представления неманипулируемых механизмов однокритериальной активной экспертизы определим понятие медианной схемы на основании определения 4.

⁸ Иначе говоря, множество A — замкнутая область, т. е. $A = \text{cl}(\text{int}(A))$, где cl и int означают замыкание и внутренность соответственно.

Определение 6. Медианная схема $x = h(\tau)$ на множестве $A \subset \mathbb{R}^1$ порождена системой правых (левых) коалиций $R(z)$ ($L(z)$), если для $\forall t \in A^n$

$$h(\tau) = \max\{z \in A \mid \{i \in N \mid \tau^i \geq z\} \in R(z)\}$$

$$(h(\tau) = \min\{z \in A \mid \{i \in N \mid \tau^i \leq z\} \in L(z)\}). \blacklozenge$$

Опираясь на данное определение, сформулируем требования к системам правых и левых коалиций, определяющих прямые неманипулируемые однокритериальные механизмы активной экспертизы, эквивалентные процедурам экспертизы, удовлетворяющим условиям А.1—А.3. Обозначим s_S — вектор сообщений экспертов, участников коалиции S , $s_{N/S}$ — вектор сообщений остальных экспертов.

Лемма. Пусть задана функция $\pi: [d, D]^n \rightarrow [d, D]$, удовлетворяющая условиям А.1—А.3. Тогда функция $R_\pi: [d, D] \rightarrow 2^N \setminus \emptyset$, определяемая как

$$R_\pi = \{S \subset N \mid \pi(s_S, s_{N/S}) \geq z, \quad s_S = (D, \dots, D), \\ s_{N/S} = (d, \dots, d)\}, \quad \forall z \in (d, D], \quad R_\pi(d) = 2^N \setminus \emptyset, \quad (5)$$

является системой правых коалиций. \blacklozenge

Доказательства леммы и утверждений 1—3 (см. далее) вынесены в Приложение.

Качественно, запись (5) означает, что для любой коалиции $S \subset N$ максимальный результат экспертизы, который гарантированно может обеспечить данная коалиция, определяется на основе вектора сообщений, в котором все участники коалиции сообщают максимально возможное значение параметра D , а эксперты, не входящие в эту коалицию, — минимально возможное значение d .

Покажем, что определяемая выражением (5) функция R_π порождает прямой механизм, эквивалентный механизму $\pi(s)$.

Утверждение 1. Пусть на множестве $[d, D]^n$ задан механизм активной экспертизы $\pi(s)$, удовлетворяющий требованиям А.1—А.3. Тогда система правых коалиций R_π определяемая выражением (5), порождает эквивалентный ему прямой механизм. \blacklozenge

Для отдельных примеров механизмов экспертизы формула (5) принимает более наглядный вид. В частности, для механизма

$$x = \pi(s): x = \sum_{i=1}^n \alpha_i s^i, \quad \forall i \alpha_i > 0, \quad \sum_{i=1}^n \alpha_i = 1,$$

система правых коалиций записывается следующим образом: $\forall z \in (d, D]$,

$$R_\pi(z) = \left\{ S \subset N \mid \sum_{i \in S} \alpha_i \geq \frac{z-d}{D-d} \right\},$$

$$\forall z \in (d, D], \quad R_\pi(d) = 2^N \setminus \emptyset.$$



В частности, для механизма усреднения (1) эта запись приобретает вид:

$$R_\pi(z) = \left\{ S \subset N \mid \#S = r \geq \frac{z-d}{D-d} \right\}, \forall z \in (d, D],$$

$$R_\pi(d) = 2^N \setminus \emptyset.$$

Другими словами, возможности коалиции экспертов влиять на результат экспертизы определяются отношением их численности к общему числу экспертов. Например, если трое экспертов определяют в результате экспертизы точку на отрезке $[0, 1]$, то для выбора любой точки из диапазона $[0, 1/3]$ достаточно одного эксперта, чья точка пика будет расположена не левее этой точки, для точки из полуинтервала $(1/3, 2/3]$ — двух, для точки из полуинтервала $(2/3, 1]$ — трех.

Таким образом, мы определили вид системы правых коалиций, порождающей интересующие нас прямые неманипулируемые механизмы однокритериальной активной экспертизы. В соответствии с работой [6] можно определить условия на систему левых коалиций $L^*(z)$, порождающую тот же механизм, что и заданная система правых коалиций $R(z)$.

Утверждение 2. Для системы правых коалиций $R_\pi(z)$, определяемой выражением (5), система левых коалиций L_π^* , порождающая тот же механизм, записывается следующим образом:

$$L_\pi^* = \{S \subset N \mid \pi(s_S, s_{N/S}) \leq z, s_S = (d, \dots, d),$$

$$s_{N/S} = (D, \dots, D)\}, \forall z \in [d, D], L_\pi^* = 2^N \setminus \emptyset. \blacklozenge (6)$$

Для наглядности приведем вид условия (6) для механизма экспертизы $x = \pi(s): x = \sum_{i=1}^n \alpha_i s^i$:

$$L_\pi(z) = \left\{ S \subset N \mid \sum_{i \in S} \alpha_i \leq n \frac{D-z}{D-d} \right\}, \forall z \in [d, D],$$

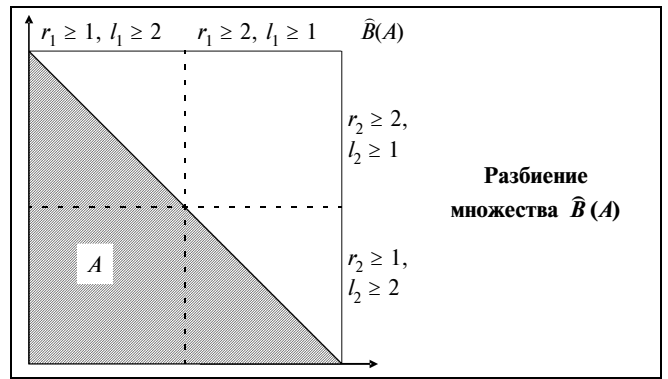
$$L_\pi(D) = 2^N \setminus \emptyset.$$

А для механизма (1):

$$L_\pi(z) = \left\{ S \subset N \mid \#S = l \leq \frac{D-z}{D-d} \right\}, \forall z \in [d, D],$$

$$L_\pi(D) = 2^N \setminus \emptyset.$$

Полученные в утверждениях 1 и 2 системы правых и левых коалиций для однокритериальных задач активной экспертизы могут быть использованы для построения многокритериальных неманипулируемых механизмов активной экспертизы в соответствии с определением 4 — путем определения семейств правых или левых коалиций. По сути, эти семейства коалиций разбивают все множество $\hat{B}(A)$ на n^m m -мерных прямоуголь-



ников вида $B_{s_1^{i_1}, \dots, s_m^{i_m}} = \prod_{k \in M} [\pi_k(s_k^{i_k-1}), \pi_k(s_k^{i_k})]$.

Для $\forall z \in \text{int } B_{s_1^{i_1}, \dots, s_m^{i_m}}$ все требования к правым и левым коалициям неизменны. На границах требования меняются — на нижних границах по каждому критерию изменяются требования к правым коалициям, на верхних — к левым. На рисунке приведен пример такого разбиения для задачи, сформулированной в § 1 при $n = 2, m = 2$.

В данной статье мы не исследуем подробно вопрос, для каких механизмов многокритериальной экспертизы можно построить эквивалентные прямые механизмы, ограничиваясь лишь формулировкой следующего утверждения.

Утверждение 3. Пусть задан многокритериальный механизм активной экспертизы $\pi(s)$, удовлетворяющий свойствам A.1—A.3, определенный на множестве A^n , где $A = \prod_{k \in M} A_k, A_k = [d_k, D_k] \subset \mathbb{R}^1$.

Тогда эквивалентный ему прямой механизм $h_\pi(\tau)$ существует тогда и только тогда, когда механизм $\pi(s)$ является сепарабельным по критериям. Причем прямой механизм $h_\pi(\tau)$ порождается семейством систем правых коалиций $R_\pi = \{R_k\}_{k=1}^m$, определяемым следующим образом: для $\forall k \in M$

$$R_k(z_k) = \{S \subset N \mid \pi_k(s_k^S, s_k^{N/S}) \geq z_k,$$

$$s_k^S = (D_k, \dots, D_k), s_k^{N/S} = (d_k, \dots, d_k)\},$$

$$\forall z \in (d_k, D_k], R_k(d_k) = 2^N \setminus \emptyset. \blacklozenge (7)$$

Нами приведен класс механизмов многокритериальной активной экспертизы, для которых можно построить эквивалентные прямые механизмы, если множество допустимых результатов голосования — многомерный прямоугольник. Утверждения 1—3 составляют основной результат данной работы — прямые неманипулируемые механизмы активной экспертизы представлены как обобщенные медианные схемы, сформулированные в терминах семейств правых и левых коалиций, что

позволяет проверять наличие у них свойства пересечения для различных множеств допустимых результатов экспертизы.

4. ПРИМЕР ПРИМЕНЕНИЯ АЛГОРИТМА ПРОВЕРКИ РЕАЛИЗУЕМОСТИ ПРЯМОГО НЕМАНИПУЛИРУЕМОГО МЕХАНИЗМА МНОГОКРИТЕРИАЛЬНОЙ АКТИВНОЙ ЭКСПЕРТИЗЫ

Проиллюстрируем применение алгоритма проверки выполнения свойства пересечения для сформулированной в § 1 задачи распределения ресурса на основании экспертных оценок по механизму усреднения, определяемого выражением (1), на

множестве $A = \{x \in A \subset \mathfrak{R}^m \mid \sum_{j=1}^m x_j \leq C, x_j \geq 0, \forall j \in M\}$.

Определим, для какого числа проектов эквивалентный прямой механизм реализуем — результат экспертизы остается в области допустимых значений. Из утверждения 3 получаем, что семейство систем правых коалиций, порождающее неманипулируемый прямой механизм, эквивалентный механизму (1), имеет следующий вид:

$$R_k(z) = \left\{ S \subset N \mid \#S = r \geq n \frac{z}{C} \right\}, \quad \forall z \in (0, C],$$

$$\forall k \in M, R_k(0) = 2^N \setminus \emptyset, \quad (8)$$

семейство левых коалиций:

$$L_k(z) = \{S \subset N \mid \#S = l \leq n(C - z)/C\}, \quad \forall z \in [d, C],$$

$$\forall k \in M, L_k(C) = 2^N \setminus \emptyset.$$

Для рассматриваемого множества A для любой точки $y \in \widehat{B}(A) \setminus A$ точки решающего множества $\widehat{S}(y) \subset A$ определяются следующим образом [9]: $\forall t \subset M z^t \in \widehat{S}(y): z_k^t = y_k$ если $k = t$, а $z_k^t = 0$ для $k \neq t$.

То, что $y \in \widehat{B}(A) \setminus A$ означает, что $\sum_{j=1}^m y_j > C$. Для

каждой точки $y \in \widehat{B}(A) \setminus A$ и соответствующего ей решающего множества $\widehat{S}(y) = \{z^1, \dots, z^m\} \subset A$ для $\forall t \in M, M^+(y, z^t) = \emptyset$, а $M^-(y, z^t) = M \setminus t$. Поэтому выражение (4) принимает следующий вид:

$$\bigcap_{t=1}^m \left\{ \bigcup_{k \in M \setminus \{t\}} r_k(y_k) \right\} \neq \emptyset,$$

что эквивалентно следующей записи:

$$\sum_{j=1}^m r_j(y_j) \geq nC^{-1} \sum_{j=1}^m y_j > n \geq (m-1)n + 1.$$

Из полученного условия видно, что для $\forall n \geq 2$ прямой многокритериальный механизм активной

экспертизы, порождаемый механизмом (8), реализуем только для $m \leq 2$. Однако «ослабление» бюджетного ограничения может увеличить число критериев, для которых данный механизм будет реализуем. Рассмотрим множества допустимых результатов экспертизы следующего вида:

$$A = \left\{ x \in A \subset \mathfrak{R}^m \mid \sum_{j=1}^m x_j \leq C, x_j \in [0, 1], \forall j \in M \right\},$$

где $C \in [1, m]$. В смысле задачи распределения ресурсов, с помощью множества подобного вида моделируется ситуация, когда на каждый проект разрешается выделять долю средств из общего объема финансирования C не более чем $1/C$, причем при $C = m$ множество A превращается в m -мерный прямоугольник. Можно показать, что прямой механизм активной экспертизы, порождаемый механизмом (8), удовлетворяет свойству пересечения на множестве A при $C \geq m - 1$.

Качественно это означает, что чем больше проектов рассматривается при распределении бюджета, тем меньшую долю средств позволено любому эксперту выделять на отдельный проект.

Более подробное рассмотрение зависимости вида прямого неманипулируемого механизма, а, точнее, обобщенных медианных схем, порождающих его, и вида множеств допустимых значений результатов экспертизы, на которых данный механизм обладает свойством реализуемости, представляет собой отдельную задачу, интересной, как с теоретической, так и практической точки зрения.

ЗАКЛЮЧЕНИЕ

В данной статье проиллюстрирована возможность применения обобщенных медианных схем для построения прямых неманипулируемых механизмов многокритериальной активной экспертизы. Для этого механизм представляется в терминах семейств систем левых и правых коалиций, что позволяет анализировать его на предмет выполнения свойства пересечения. Полученные утверждения, определяющие вид семейств систем левых и правых коалиций, позволяют наглядно представить прямые неманипулируемые механизмы активной экспертизы. Применение данного подхода представляется достаточно перспективным для практической реализации неманипулируемых механизмов экспертизы. Однако за рамками статьи остались такие важные вопросы, как оптимальность построенного механизма по какому-либо критерию, точнее, вопрос сохранения оптимальности механизма при переходе от множества допустимых результатов экспертизы, описываемого многомерным прямоугольником, к множеству, вписанному в него, для которого механизм обла-



дает свойством пересечения. Отдельной перспективной и актуальной с практической точки зрения является задача поиска связи вида множества допустимых результатов экспертизы с множеством допустимых механизмов — прямых неманипулируемых механизмов, обладающих свойством реализуемости. И, как отмечалось, за рамками данной статьи осталась задача построения прямого неманипулируемого механизма активной экспертизы, эквивалентного произвольному непрямоу, так как задача решена лишь для узкого класса механизмов экспертизы — сепарабельных по критериям. И, даже для данного класса, следует анализировать сохранение свойства эквивалентности при изменении вида множества допустимых результатов экспертизы.

ПРИЛОЖЕНИЕ

Доказательство леммы. Для доказательства проверим, что функция R_π , определяемая выражением (5), удовлетворяет всем пунктам определения 3.

1. Из записи (5) следует, что $R_\pi(d) = 2^N \setminus \emptyset$, и для $\forall z \in (d, D]$, $R_\pi(z) \neq \emptyset$, так как хотя бы для $S = N$ $\pi(s_S, s_{N/S}) = D \geq z$ из условия единогласия А.3. Из этого же условия следует, что для $S = \emptyset$ $\pi(s_S, s_{N/S}) = d$, следовательно, для $\forall z \in (d, D]$ $\emptyset \in R_\pi(z)$.

2. Если $S \in R_\pi(z)$ и $S \subset S'$, из условия строгой монотонности А.1 и условия единогласия А.3 следует, что $\pi(s_{S'}, s_{N/S'}) \geq \pi(s_S, s_{N/S}) \Rightarrow S' \in R_\pi(z)$.

3. Если $z' < z$ и $S \in R_\pi(z)$, то $\pi(s_{S'}, s_{N/S'}) > z' \Rightarrow S \in R_\pi(z')$.

4. Из условия непрерывности А.2. следует, что если для любой последовательности $\{z^i\} \subset [d, D]$ такой, что $\lim_{i \rightarrow \infty} z^i = z$ и для $\forall S \subset N$, таких, что $\pi(s_S, s_{N/S}) \geq z^i$; т. е. $\pi(s_S, s_{N/S}) \geq z^i$; верно, что $\pi(s_S, s_{N/S}) \geq z \Rightarrow [S \in W(z)]$.

Следовательно, определяемая выражением (5) функция R_π является системой правых коалиций. Лемма доказана. ♦

Доказательство утверждения 1. Упорядочим экспертов по возрастанию точек пика τ . Обозначим $w^i = \pi(s(i))$, где $s(i): \{\forall j > i, s_j = D, \forall j \leq i, s_j = d\}$; т. е. $w^{i-1} = \pi(s_{S_i}, s_{N/S_i})$, где коалиция S_i состоит из всех экспертов с индексами i и выше. Тогда условие $\{i \in N | \tau^i \geq z\} \in R_\pi(z)$ эквивалентно условию $\{i \in N | \min(\tau^i, w^{i-1}) \geq z\}$. Соответственно, в определении схемы «среднего голоса» в терминах правых коалиций $h(\tau) = \max\{z \in [d, D] | \{i \in N | \tau^i \geq z\} \in R_\pi(z)\}$ можно перейти от максимума по z к максимуму по i и записать: $h(\tau) = \max_{i \in N} \min(\tau^i, w^{i-1})$.

Данная запись эквивалентна условию (2); т. е., в соответствии с результатами работы [1], прямой неманипулируемый механизм, порожденный системой правых коалиций (5), эквивалентен механизму $\pi(s)$. Утверждение 1 доказано.

Доказательство утверждения 2. Очевидно, что для точек d и D выражение (6) верно, так как

$$R_\pi(d) = 2^N \setminus \emptyset, L_\pi^* = \{N\} \Rightarrow \forall S \in R_\pi(d), \forall S' \in L_\pi^*(d) S \cap S' \neq \emptyset$$

$$\text{и } L_\pi^*(D) = 2^N \setminus \emptyset, R_\pi(D) = \{N\} \Rightarrow \forall S \in R_\pi(D),$$

$$\forall S' \in L_\pi^*(D) S \cap S' \neq \emptyset.$$

Функция $\pi(S): \pi(s_S, s_{N/S}), s_S = (d, \dots, d), s_{N/S} = (D, \dots, D), S \subset N$ задает разбиение отрезка $[d, D]$ не более чем на 2^n от-

резков. Введем упорядочение на множестве всех подмножеств N по возрастанию функции $\pi(S): \forall i, j \in 2^n, j > i, \pi(S_j) \leq \pi(S_i)$. Тогда рассматриваемые нами системы правых и левых коалиций можно записать как $\forall i R(\pi(S_i)) = \{S_j, j \geq i\}, \forall z \in (\pi(S_i), \pi(S_{i+1}))$, $R(z) = R(\pi(S_{i+1}))$, если $\pi(S_i) = \pi(S_{i+1})$, то $R(\pi(S_{i+1})) = R(\pi(S_i)) = \{S_j, j \geq i\}$, и $\forall i L^*(\pi(S_i)) = \{N \setminus S_j, j \leq i\} \forall z \in (\pi(S_{i-1}), \pi(S_i))$, $L^*(z) = L^*(\pi(S_i))$, если $\pi(S_i) = \pi(S_{i+1})$, то $L^*(\pi(S_{i-1})) = L^*(\pi(S_i)) = \{N \setminus S_j, j \leq i\}$.

Тогда для всех $\forall r \geq i + 1$ и $\forall l \leq i$ ситуация, что $S_r \cap N \setminus S_l = \emptyset$ возможна только в том случае, когда $S_r \subset S_l$, чего не может быть в силу свойств монотонности функции $x = \pi(s)$ и введенного нами упорядочения. Следовательно, система левых коалиций $L^*(z)$ определяемая выражением (6) удовлетворяет записи (4). Утверждение 2 доказано. ♦

Доказательство утверждения 3. Докажем, что прямой механизм, эквивалентный многокритериальному механизму активной экспертизы, удовлетворяющему условиям А.1—А.3 и сепарабельному по критериям, порождается семейством систем правых коалиций, определяемым записью (7). Из определения 2 следует, что результат голосования по каждому критерию определяется лишь сообщениями экспертов по данному критерию — $x = \pi(s): \forall k \in M x_k = \pi_k(s_k)$; т. е. мы можем рассмотреть m однокритериальных задач экспертизы, для каждой из которых предложив систему правых коалиций в соответствии с выражением (5) — тем самым получив запись (7).

Покажем необходимость сепарабельности по критериям механизма активной экспертизы для существования эквивалентного ему прямого механизма. Из определения 4 видно, что любой прямой механизм, порождаемый семейством систем правых или левых коалиций, является сепарабельным по критериям. Из выражения (7) видно, что это возможно только в том случае, если механизм $\pi(s)$ является сепарабельным по критериям. Утверждение 3 доказано.

ЛИТЕРАТУРА

1. *Большие системы: моделирование организационных механизмов* / В.Н. Бурков, Б. Данев, А.К. Еналеев и др. — М.: Наука, 1989. — 245 с.
2. *Новиков Д.А., Петраков С.Н.* Курс теории активных систем. — М.: СИНТЕГ, 1999. — 108 с.
3. *Петраков С.Н.* Механизмы планирования в активных системах: неманипулируемость и множества диктаторства. — М.: ИПУ РАН, 2001. — 135 с.
4. *Иващенко А.А., Коргин Н.А., Новиков Д.А.* Модели и методы оценки эффективности портфеля проектов // Системы управления и информационные технологии. — 2005. — № 3 (20). — С. 92—98.
5. *Иващенко А.А., Коргин Н.А., Новиков Д.А.* Неманипулируемые механизмы экспертизы при неограниченных множествах возможных сообщений экспертов // Изв. Тульского гос. ун-та. — 2005. — Вып. 8. — Ч. 2. — С. 159—165.
6. *Barbera S., Masso J., Serizawa S.* Strategy-proof voting on compact ranges // Games and Behavior. — 1998. — Vol. 25. — P. 272—291.
7. *Border K., Jordan J.* Straightforward elections, unanimity and phantom voters // Review of Economic Studies. — 1983. — Vol. 50. — P. 153—170.
8. *Moulin H.* On strategy-proofness and single-peakedness // Public Choice. — 1980. — Vol. 35. — P. 437—455.
9. *Barbera S., Masso J., Neme A.* Voting under Constraints // Journ. Econ. Theory. — 1997. — Vol. 76. — P. 298—321.

☎ (495) 334-79-00; e-mail: nkorgin@ipu.ru

Статья представлена к публикации членом редколлегии Д.А. Новиковым. □

МНОГОКРИТЕРИАЛЬНАЯ ОЦЕНКА КОНКУРЕНТОСПОСОБНОСТИ И РАНЖИРОВАНИЕ ЭКОНОМИЧЕСКИХ СУБЪЕКТОВ, ФУНКЦИОНИРУЮЩИХ В РАЗНОРОДНЫХ УСЛОВИЯХ ПРИ НАЛИЧИИ НЕОПРЕДЕЛЕННОСТИ

Б.М. Кувшинов, А.С. Челядин, В.И. Ширяев

Южно-Уральский государственный университет, г. Челябинск

Предложен подход к построению многокритериальной оценки конкурентоспособности и ранжированию предприятий, основанный на оценке их потенциальных возможностей. Для построения таких оценок привлечен аппарат распознавания образов. Рассмотрены особенности решения задачи оценки потенциала предприятия и сравнения полученных интервальных оценок предприятий в условиях неопределенности исходной информации.

ВВЕДЕНИЕ

Задача оценки и сопоставления функционирования экономических предприятий всегда возникает у инвесторов при выборе объекта финансирования, а также у руководителей организаций в ходе планирования их развития. Существует достаточно много методик оценки конкурентоспособности, но, как правило, они позволяют оценить либо меру фактических результатов деятельности предприятия (показатели прибыльности, рентабельности) [1, 2], либо меру потенциальных результатов как оценку доступных предприятию ресурсов [3]. Все они опираются в основном на методы многокритериального ранжирования [4–8]. Большой интерес представляет оценка конкурентоспособности, которая могла бы служить мерой перевода доступных ресурсов в фактический результат. Подобная трактовка понятия «конкурентоспособность» встречается в работе [2], в соответствии с которой последняя может быть также названа эффективностью управления предприятием. Процедура оценки конкурентоспособности и ранжирования в этом смысле должна содержать механизм «масштабирования», т. е. приведение результатов

деятельности предприятий к единой системе, инвариантной относительно объемов производства, условий работы и ресурсов, которыми располагают предприятия, и позволять оценивать эффективность использования предприятиями своих ограниченных ресурсов в сложившихся рыночных условиях. Настоящая работа продолжает исследования [9–14].

1. ПОСТАНОВКА ЗАДАЧИ ПОСТРОЕНИЯ ОЦЕНКИ

Результат, достигнутый предприятием за определенный период времени, обычно представляется набором значений показателей по множеству заранее заданных количественных критериев. Такому набору может быть поставлено в соответствие число. Ресурсы, условия функционирования, а также стратегия поведения предприятия на рынке также представляется множеством показателей.

Пусть u_i — вектор ресурсов, которыми располагает i -е предприятие, s_i — вектор показателей, характеризующих стратегию поведения i -го предприятия, c_i — вектор экономических условий функционирования i -го предприятия, $i = \overline{1, N}$, где N — число анализируемых предприятий. Введем



скалярную величину p_i , которая будет характеризовать результат, которого могло бы достичь i -е предприятие в сложившихся экономических условиях, с учетом имеющихся ресурсов и выбранной стратегии поведения. Назовем ее потенциалом i -го предприятия.

Для потенциала можно записать зависимость:

$$p_i = \psi_i(u_i, c_i, s_i), \quad i = \overline{1, N},$$

где ψ — отображение, $\psi: R^M \rightarrow R$, M — суммарная размерность векторов u_i , c_i и s_i . Совокупность векторных оценок (u_i, c_i, s_i) для i -го предприятия обозначим через вектор r_i и будем называть его вектором параметров состояния.

Оценку потенциала можно получить, сравнив статистические результаты работы предприятия с фактическими данными о работе других предприятий, имевших в своем распоряжении аналогичные ресурсы и находившиеся в близких условиях. Предприятия, располагающие близкими ресурсами и находящиеся в близких условиях, должны иметь близкие оценки потенциала.

Таким образом, оценки потенциала p_i будут строиться в виде:

$$p_i = f(\{r_j, q_j\}, \quad j = \overline{1, N}), \quad (1)$$

где $f: R^{(M+1)N} \rightarrow R$ — искомое отображение; q_j — скалярная оценка результата, достигнутого j -м предприятием (построение оценки q_j будет описано в § 2). Зависимость (1) строится с учетом информации о фактических результатах и параметрах состояния всех предприятий.

Итоговая скалярная оценка конкурентоспособности i -го предприятия строится как соотношение между фактическим результатом и потенциальным: $e_i = g(q_i, p_i)$, $i = \overline{1, N}$, где $g: R^2 \rightarrow R$. Такая оценка учитывает отличия условий функционирования предприятий и соотносит их с достигнутым результатом, приводя, таким образом, оценки к единой системе измерений.

В общем случае, когда число параметров состояния велико, нахождение зависимости (1) становится трудной задачей. Здесь предлагается ввести некоторые упрощения при построении ее аппроксимации, которые будут рассмотрены позднее.

2. ОЦЕНКА ФАКТИЧЕСКИХ РЕЗУЛЬТАТОВ

Оценка фактического результата функционирования предприятий представляет собой многокритериальную задачу. В качестве ее решения мо-

жет быть взята любая функция ценности в зависимости от используемых критериев и предпочтений между ними.

Так как построение функции ценности на практике для критериев, число которых велико, — трудоемкая задача, можно положить, что замещения критериев не зависят от их значений и выполняется условие взаимнезависимости по предпочтению [15]. В этом случае оценить фактический результат можно с помощью распространенной [16, 17] функции ценности, представляющей собой взвешен-

ную сумму частных критериев: $q_i = \sum_{j=1}^L \alpha_j x_{ij}$, где

$x_{i,j} \in R^M$, $i = \overline{1, N}$ — исходные параметры (значения критериев), характеризующие результаты работы i -го предприятия, α_j — весовые коэффициенты, определяющие значимость частных критериев, $\sum_{j=1}^L \alpha_j = 1$, $\alpha_j > 0$.

Чтобы избежать главного недостатка этого метода — подверженности ошибкам определения коэффициентов α_i — положим [7, 8] $\alpha_i \in [\alpha_{i\min}, \alpha_{i\max}]$, $i = \overline{1, L}$.

Значения критериев, используемых для оценки фактического результата функционирования предприятия, должны быть нормированы.

3. ОЦЕНКА ЗНАЧИМОСТИ ЧАСТНЫХ КРИТЕРИЕВ

Человеку трудно напрямую оценивать значения важности критериев, даже если они представлены в интервалах [18], поэтому для оценки интервалов предлагается использовать процедуру, предложенную в работе [19].

Эксперт должен расставить все критерии по местам в порядке возрастания их значимости, оставляя между ними пустые места, если это необходимо. Если один критерий находится на одну позицию выше другого, то критерии будут иметь разные веса, и эта разница может быть выбрана как единица оценки интервалов между весами u . Если между критериями остается одно пустое место, разница между весами будет составлять $2u$; если два пустых места, то $3u$. Дальнейшая процедура подсчета весов состоит из следующих шагов.

1. Определение ненормализованных весов k_r для каждой позиции расстановки множества критериев. Пусть $e_r = e'_r + 1$, $\forall r = 1, \dots, \bar{L} - 1$, где \bar{L} —

число позиций в расстановке критериев, e'_r — число пустых мест между позициями r и $r + 1$, тогда

$$k_r = 1 + (z - 1)(e_0 + \dots + e_{r-1}) / \sum_{j=1}^{\bar{L}-1} e_j$$

где z — значение, которое показывает, во сколько раз самый «верхний» критерий важнее самого «нижнего». Это значение определяется экспертом напрямую.

2. Определение нормализованных весов критериев:

$$\alpha_i = 100 k'_i / \sum_{j=1}^n k'_j,$$

где k'_i — ненормализованный вес, соответствующий позиции, которую занимает i -й критерий. ♦

Авторами предлагается изменить метод для получения границ интервалов значимости весовых коэффициентов следующим образом. Для получения нижней (верхней) границы весового коэффициента $\alpha_{i\min}$ ($\alpha_{i\max}$) для i -го критерия он сдвигается на одну позицию вниз (вверх), значения α_i при этом пересчитываются. Таким образом, для получения границ по всем критериям пересчет значений весовых коэффициентов производится $2L$ раз.

4. ОЦЕНКА ПОТЕНЦИАЛА

Поскольку оценки фактического результата интервальные, то оценка потенциала также интервальная: $p_i = \varphi(\{\alpha_j\}, i = \overline{1, L})$, $p_i \in [p_{i\min}, p_{i\max}]$ (рис. 1).

Неопределенность в значениях параметров состояния (ресурсах и условиях функционирования) можно интерпретировать как неопределенность в

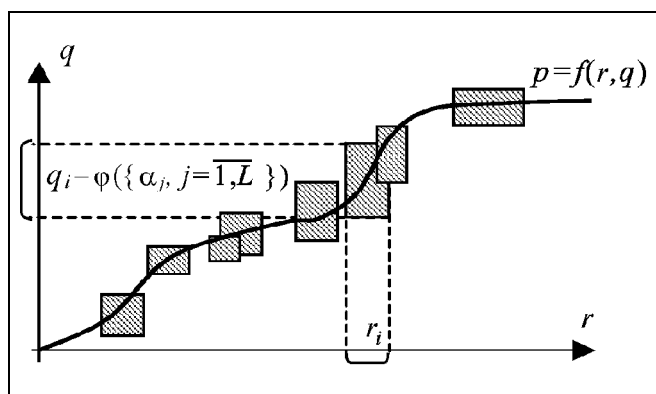


Рис. 1. Особенности постановки задачи оценки потенциала: интервальные оценки результатов q_i , p_i и неопределенность значений ресурсов и условий работы r_i

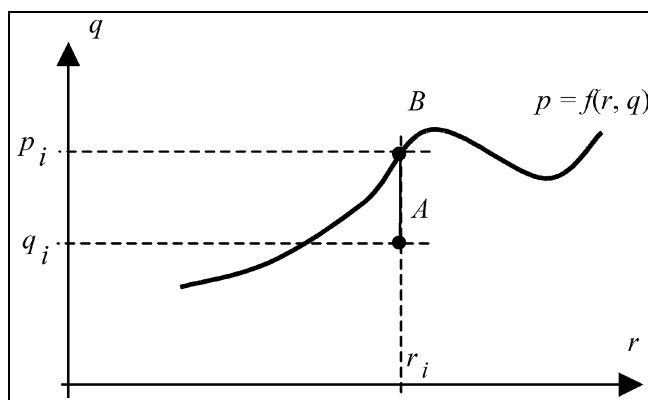


Рис. 2. Интерпретация отклонения фактического результата предприятия от его потенциала

оценке фактических результатов, т. е. от описания состояния предприятия в форме (r_j, q_j) , $r_{ij} \in [r_{ij\min}, r_{ij\max}]$, $q_j \in [q_{j\min}, q_{j\max}]$, где $j \in \overline{1, M}$ — индекс координаты пространства параметров, по которой значение задано в виде интервала, можно перейти к описанию вида (r_j, q_j) , $q_j \in [q'_{j\min}, q'_{j\max}]$, где r_{ij} принимает конкретное значение, а $[q'_{j\min}, q'_{j\max}] \supseteq [q_{j\min}, q_{j\max}]$ — границы интервала для q_j , расширенные с учетом возможной ошибки в выбранном значении r_{ij} .

Число параметров состояния, как правило, достаточно велико и сравнимо с числом рассматриваемых предприятий: $M \geq \sqrt{N}$ (N — число рассматриваемых объектов; M — число используемых параметров), а функциональный класс, в котором следует искать аппроксимацию $p = f(r, q)$, заранее не известен. Чтобы избежать недообучения/переобучения в этих условиях, для аппроксимации необходимо использовать класс функций, сложность которых будет адаптивно настраиваться на сложность представленных для анализа данных.

Таким образом, процедура оценки потенциала предприятий должна породить в пространстве размерности $(M + 1)$ семейство аппроксимаций $p = f(r, q)$, которое должно быть устойчиво к неопределенности в исходных данных.

5. ИТОГОВАЯ ОЦЕНКА КОНКУРЕНТОСПОСОБНОСТИ

Отклонение значения q_i от p_i интерпретируется как отличие эффективности использования предприятием своих ресурсов в сложившихся условиях от нормальной. Другими словами, предприятие, фактически располагающееся в координатах (r, q) в точке A , могло бы оказаться в точке B (рис. 2). Соответственно, эффективность перевода ресур-



сов в результат или конкурентоспособность i -го предприятия e_i можно оценить величиной $e_i = g(q_i, p_i) = q_i/p_i = q_i/f(r_i, q_i)$, где $e_i \in [e_{i\min}, e_{i\max}]$.

6. ИСПОЛЬЗОВАНИЕ АДАПТИВНОЙ КЛАССИФИКАЦИИ ДЛЯ ОЦЕНКИ ПОТЕНЦИАЛА

Множество анализируемых объектов можно представить как совокупность точек $\{(r_i, q_i) \in R^{M+1}, i = \overline{1, N}\}$ в многомерном пространстве параметров состояния и фактических результатов работы предприятий, причем положение каждой точки может быть задано с ошибкой. Информации о работе предприятий недостаточно для построения аналитической зависимости $p = f(r, q)$. Однако эти данные можно использовать для того, чтобы выделить несколько линий уровня этой зависимости: $d(q^k)$, где $q^k, k \in \overline{1, K-1}$ — некоторые фиксированные значения для q , представляющих собой геометрическое место точек, являющихся пересечением искомой функциональной зависимости $p = f(r, q)$ с гиперплоскостью $q = q^k$ (верхний индекс обозначает номер линии уровня).

Таким образом, при проецировании информации из $(M+1)$ -мерного пространства параметров состояния и результатов работы предприятий в M -мерное пространство параметров их состояния, фактические результаты работы каждого предприятия q_i заменяются на информацию о том, ниже какой линии уровня находится этот результат (к какому классу $C(q_i) \in \overline{1, K}$ он принадлежит). Обратное, информация о линии уровня задается значениями параметров состояния предприятий и их классификацией по фактическому результату:

$$\{(r_i, C(q_i)), i = \overline{1, N}\}. \quad (2)$$

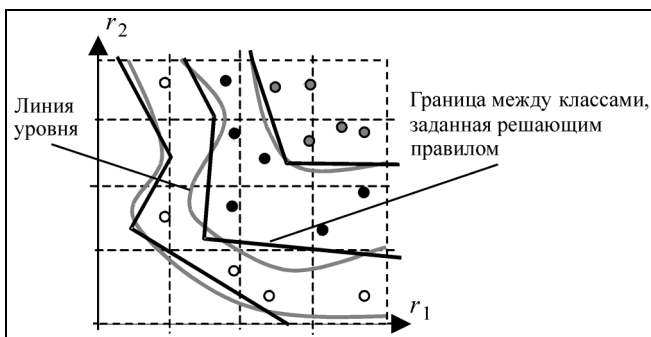


Рис. 3. Аппроксимация линий уровня функциональной зависимости $p = f(r, q)$ решающими правилами

Такое упрощение исходной информации позволяет применить методы распознавания образов для выполнения аппроксимации. Информация вида (2) представляет собой обучающую выборку, по которой можно решить стандартную задачу дискриминантного анализа. Полученные в результате решающие правила $D^k(r), k = \overline{1, K-1}$, будут задавать правило отнесения любой точки $r \in R^M$ к одному из K классов. Таким образом будут определены границы между подмножествами объектов, принадлежащих к разным классам (рис. 3).

Для аппроксимации линий уровня функциональной зависимости $p = f(r, q)$ предлагается применять кусочно-линейные (комитетные) решающие правила [20] и алгоритм их построения, позволяющий максимизировать функционал согласованности априорного описания объектов [12, 13] путем изменения информации о классах, к которым должны быть отнесены объекты.

Таким образом, класс потенциала оценивается с помощью решающего правила по значениям ресурсов и условий функционирования каждого предприятия. Конкретное числовое значение потенциала предприятия определяется с помощью процедуры максимизации функционала согласованности априорного описания объектов.

7. СОПОСТАВЛЕНИЕ ИТОВОГЫХ ОЦЕНОК

По полученным интервальным оценкам предприятия необходимо расставить по местам в рейтинге. Процедура упорядочения должна удовлетворять следующему свойству: место оценки $q_A \in [q_{A\min}, q_{A\max}]$, $A \in \overline{1, M}$, в рейтинге выше, чем место оценки $q_B \in [q_{B\min}, q_{B\max}]$, $B \in \overline{1, M}$, тогда и только тогда, когда эти оценки находятся в отношении доминирования π . Отношение π можно задать следующим образом:

— оценка q_A доминирует над оценкой q_B , если интервал значений оценки q_A целиком выше, чем интервал значений оценки q_B , т. е. $q_A \pi q_B: q_{A\min} > q_{B\max}$;

— оценка q_A доминирует над оценкой q_B , если при любой допустимой комбинации коэффициентов α_j значение оценки q_A больше, чем значение оценки q_B , т. е. $q_A \pi q_B: \forall \alpha = (\alpha_1, \dots, \alpha_L) \in \Omega$:

$$\sum_{j=1}^L \alpha_j x_{Aj} > \sum_{j=1}^L \alpha_j x_{Bj}$$

Эти отношения задают слишком жесткие ограничения, при которых место каждого предприятия в рейтинге может изменяться в широком диапазо-

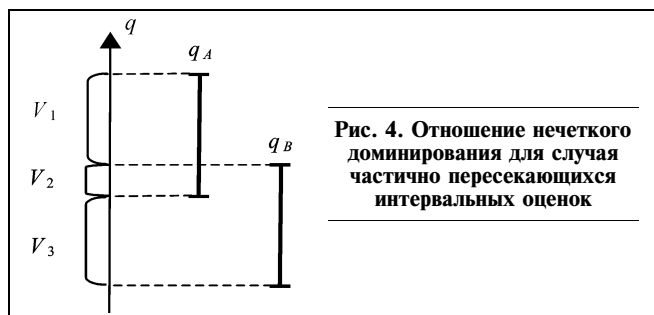


Рис. 4. Отношение нечеткого доминирования для случая частично пересекающихся интервальных оценок

не. В то же время, практическую ценность при построении рейтинга представляет строгое упорядочивание совокупности произвольного набора интервальных оценок. С этой целью, по аналогии с работой [8], предлагается ввести отношение нечеткого доминирования ρ :

$$q_A \rho q_B: \frac{V_1 + \min(V_2, 0)}{V_1 + V_2} + \max\left(\frac{(V_2 + V_3 + \max(V_3, 0))(V_2 + \min(V_3, 0))}{2(V_1 + V_2)(V_2 + V_3)}\right) > c, \quad (3)$$

где $V_1 = q_{A\max} - q_{B\max}$, $V_2 = q_{B\max} - q_{A\min}$, $V_3 = q_{A\min} - q_{B\min}$, c — некоторая константа, $c \in [0,5; 1]$.

Левая часть отношения (3), по сути, определяет вероятность того, что значение оценки q_A из интервала $[q_{A\min}, q_{A\max}]$ окажется больше значения

оценки q_B из интервала $[q_{B\min}, q_{B\max}]$ при случайном выборе уровней значимости $\alpha = (\alpha_1, \dots, \alpha_L) \in \Omega$ для построения каждой из этих оценок. На рис. 4 проиллюстрирован частный случай, когда интервалы $[q_{A\min}, q_{A\max}]$ и $[q_{B\min}, q_{B\max}]$ частично пересекаются. В этом случае формула (3) сводится к более простому виду:

$$q_A \rho q_B: \frac{V_1}{V_1 + V_2} + \frac{(V_2 + V_3 + V_3)V_2}{2(V_1 + V_2)(V_2 + V_3)} > c.$$

Такое отношение позволяет получить строгое упорядочение даже при широких интервалах изменения коэффициентов важности критериев. В случае $c = 1$ интервал для оценки q_A лежит целиком выше интервала для оценки q_B . На практике константа c сначала должна выбираться близкой к единице и уменьшаться, если результат ранжирования неприемлем (слишком много равнозначных предприятий).

8. ПРИМЕР ПОСТРОЕНИЯ РЕЙТИНГА

Приведем пример построения рейтинга предприятий по конкурентоспособности для десяти предприятий — операторов сотовой связи России. Пусть получены оценки фактического, потенциального результата, а также оценка конкурентос-

Таблица 1

Оценки для операторов (Оп.) сотовой связи РФ

	Оп. 1	Оп. 2	Оп. 3	Оп. 4	Оп. 5	Оп. 6	Оп. 7	Оп. 8	Оп. 9	Оп. 10
q_{\min}	0,977	0,790	0,430	0,237	0,356	0,764	0,417	1	0,390	0,759
q_{\max}	0,989	0,836	0,528	0,402	0,466	0,825	0,477	1	0,453	0,829
p_{\min}	0,961	0,814	0,334	0,478	0,402	0,802	0,567	0,692	0,443	0,774
p_{\max}	0,981	0,844	0,442	0,633	0,467	0,842	0,683	0,813	0,524	0,795
e_{\min}	0,995	0,958	1,025	0,734	0,953	0,922	0,794	1,186	0,928	0,964
e_{\max}	1,015	0,991	1,101	0,917	1,026	0,992	0,886	1,307	0,952	1,054

Таблица 2

Результат вычисления левой части отношения ρ

	Оп. 1	Оп. 2	Оп. 3	Оп. 4	Оп. 5	Оп. 6	Оп. 7	Оп. 8	Оп. 9	Оп. 10
Оп. 1	—	1	0	1	0,707	1	1	0	1	0,451
Оп. 2	0	—	0	1	0,293	0,750	1	0	1	0,121
Оп. 3	1	1	—	1	0,999	1	1	0	1	0,937
Оп. 4	0	0	0	—	0	0	0,420	0	0	0
Оп. 5	0,292	0,706	0,0001	1	—	0,851	1	0	1	0,292
Оп. 6	0	0,249	0	1	0,148	—	1	0	0,745	0,061
Оп. 7	0	0	0	0,579	0	0	—	0	0	0
Оп. 8	1	1	1	1	1	1	1	—	1	1
Оп. 9	0	0	0	1	0	0,254	1	0	—	0
Оп. 10	0,548	0,878	0,062	1	0,707	0,938	1	0	1	—



Результаты сравнения и ранги предприятий

	Оп. 1	Оп. 2	Оп. 3	Оп. 4	Оп. 5	Оп. 6	Оп. 7	Оп. 8	Оп. 9	Оп. 10	Ранг
Оп. 1	—	1	0	1	1	1	1	0	1	0	4
Оп. 2	0	—	0	1	0	1	1	0	1	0	6
Оп. 3	1	1	—	1	1	1	1	0	1	1	2
Оп. 4	0	0	0	—	0	0	0	0	0	0	10
Оп. 5	0	1	0	1	—	1	1	0	1	0	5
Оп. 6	0	0	0	1	0	—	1	0	1	0	7
Оп. 7	0	0	0	1	0	0	—	0	0	0	9
Оп. 8	1	1	1	1	1	1	1	—	1	1	1
Оп. 9	0	0	0	1	0	0	1	0	—	0	8
Оп. 10	1	1	0	1	1	1	1	0	1	—	3

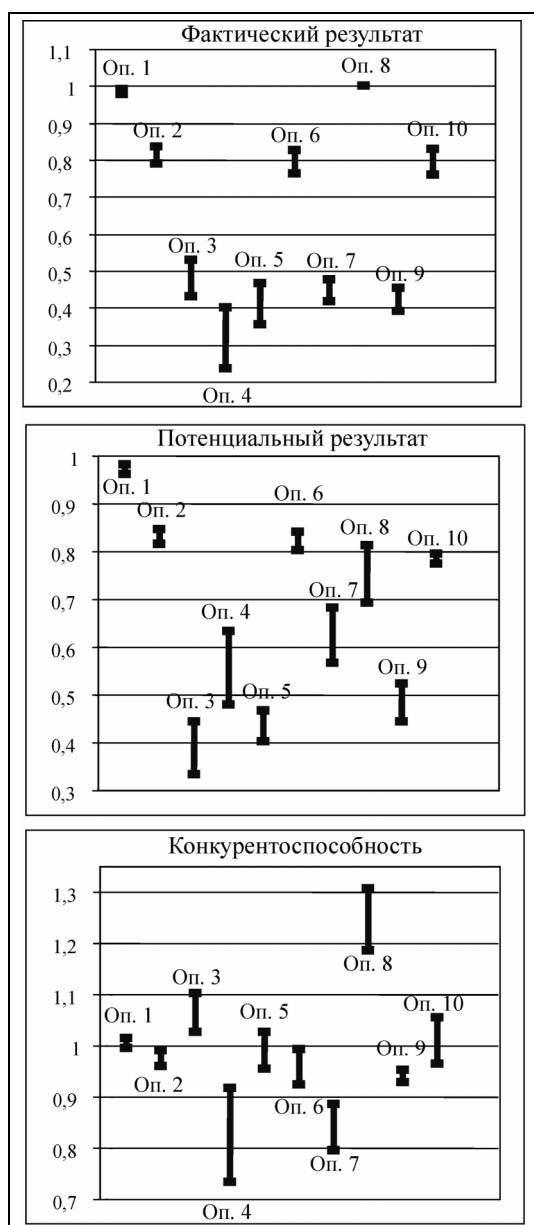


Рис. 5. Результаты ранжирования предприятий — операторов сотовой связи РФ

пособности (результаты представлены в табл. 1 и изображены на рис. 5).

Расположение предприятий (см. рис. 5) с учетом рыночных условий и доступных ресурсов отличается от первоначального расположения по фактическому результату.

Вычислим левую часть отношения ρ для всех пар объектов по конкурентоспособности (табл. 2).

Положив $c = 0,55$, получим полную ранжировку предприятий по конкурентоспособности (при $c > 0,55$ появляются несравнимые по отношению ρ объекты). Результаты их сравнения и ранги представлены в табл. 3.

ЗАКЛЮЧЕНИЕ

Предлагаемый метод оценки конкурентоспособности и ранжирования позволяет использовать информацию по уже состоявшемуся развитию большого числа предприятий, действующих в различных условиях, для выявления общих закономерностей влияния условий среды и доступных ресурсов на результаты деятельности, а также привлекать для ранжирования большое число параметров, описывающих состояние каждого предприятия, и сопоставлять результаты их работы по совокупности критериев. Учитывается разница в доступных ресурсах и условиях работы при сопоставлении результатов деятельности различных предприятий.

Основной процедурой, определяющей такие свойства ранжирования, является процедура адаптивной классификации. По сути, она представляет собой аналог процедуры аппроксимации многомерной функциональной зависимости результатов работы предприятий от ресурсов и условий их функционирования, эффективный при использовании неточной и неполной информации о ранжируемых предприятиях.

Полученные в результате сравнительные оценки конкурентоспособности могут использоваться для решения ряда задач поддержки принятия решений в управлении как отдельным предприятием, так и группой предприятий [10]. Методика может применяться для сравнения экономических субъектов различного уровня [1].

ЛИТЕРАТУРА

1. *Фатхутдинов Р.А.* Конкурентоспособность: экономика, стратегия, управление. — М.: Инфра-М, 2000. — 312 с.
2. *Buckley P., Christopher L., Prescott K.* Measures of International Competitiveness: critical survey // *Journal of Marketing Management.* — 1988. — 4 (2). — P. 175–200.
3. *Бородин К.Г.* Конкурентоспособность в рыночной экономике. — М.: ТЕИС, 2005. — 125 с.
4. *Ногин В.Д.* Принятие решений в многокритериальной среде: количественный подход. — М.: Физматлит, 2005. — 176 с.
5. *Брахман Т.Р.* Многокритериальность и выбор альтернативы в технике. — М.: Радио и связь, 1984. — 288 с.
6. *Ларичев О.И.* Свойства методов принятия решений в многокритериальных задачах индивидуального выбора // *Автоматика и телемеханика.* — 2002. — № 2. — С. 146–158.
7. *Подinovский В.В.* Анализ решений при множественных оценках коэффициентов важности критериев и вероятностей значений неопределенных факторов и целевой функции // *Автоматика и телемеханика.* — 2004. — № 11. — С. 141–158.
8. *Овезгельдыев А.О., Петров К.Э.* Оценка и ранжирование альтернатив в условиях интервальной неопределенности // *Кибернетика и системный анализ.* — 2005. — № 4. — С. 148–153.
9. *Блинов А.Б., Кувшинов Б.М., Ширяев В.И.* Оценка результатов и эффективности работы предприятий на региональных рынках (на примере операторов сотовой связи) // *Менеджмент в России и за рубежом.* — 2006. — № 5. — С. 22–31.
10. *Блинов А.Б., Кувшинов Б.М., Челядин А.С., Ширяев В.И.* О ранжировании экономических субъектов при многих критериях // *Искусственный интеллект.* — 2006. — № 4. — С. 133–137.
11. *Кувшинов Б.М., Шапошник И.И., Ширяев В.И., Ширяев О.В.* Использование комитетов в задачах распознавания образов с неточными экспертными оценками // *Изв. РАН. Теория и системы управления.* — 2002. — № 5. — С. 87–94.
12. *Система* классификации многопараметрических объектов для задач распознавания образов с неточной априорной информацией / *Б.М. Кувшинов, О.В. Ширяев, Д.В. Богданов и др.* // *Информационные технологии.* — 2001. — № 11. — С. 37–43.
13. *Ширяев В.И., Кувшинов Б.М.* Использование адаптивных методов распознавания образов в задачах принятия решений // *Искусственный интеллект.* — 2002. — № 4. — С. 526–533.
14. *Метод* многокритериального ранжирования предприятий, действующих в разнородных условиях / *А.Б. Блинов, Б.М. Кувшинов, А.С. Челядин, В.И. Ширяев* // *Вестник компьютерных и информационных технологий.* — 2007. — № 4. — С. 27–39.
15. *Кини Р.Л., Райфа Х.* Принятие решений при многих критериях: предпочтения и замещения. — М.: Радио и связь, 1981. — 560 с.
16. *Айвазян С.А.* К методологии измерения синтетических категорий качества жизни населения // *Экономика и математические методы.* — 2003. — Т. 39, № 2. — С. 33–53.
17. *Карминский А.М., Пересецкий А.А., Петров А.Е.* Рейтинги в экономике. — М.: Финансы и статистика, 2005. — 240 с.
18. *Kahneman D., Slovic P., Tversky A.* Judgment under uncertainty: heuristics and biases. — Cambridge: Cambridge University Press, 1982.
19. *Figueira J., Roy B.* Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure // *European Journal of Operational Research.* — 2002. — N 139. — P. 317–326.
20. *Khachai M.Yu., Mazurov V.D., Rybin A.I.* Committee constructions for solving problems of selection, diagnostics, and prediction // *Proc. of the Steklov Institute of Math.* — 2002. — N 1. — P. 67–101.

e-mail: aschel@ems.ru

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым. □

XVI МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ «ПРОБЛЕМЫ УПРАВЛЕНИЯ БЕЗОПАСНОСТЬЮ СЛОЖНЫХ СИСТЕМ»

Институт проблем управления им. В.А. Трапезникова РАН
г. Москва, 17 декабря 2008 г.

Предполагается рассмотреть: проблемы и методы оценки безопасности различного типа; механизмы управления безопасностью; правовое регулирование вопросов безопасности; формирование структур систем управления безопасностью; теорию и методы принятия решений, связанные с безопасностью; прогнозирование и моделирование процессов управления безопасностью; планирование и стратегическое управление в системах обеспечения безопасности; методы построения средств информационной поддержки принятия решений в системах управления безопасностью; системы управления силами и средствами при управлении безопасностью. Продолжительность работы конференции - 1 день.

Заявки на участие в конференции принимаются по адресу: 117997 Москва, ГСП-7, Профсоюзная ул., 65, Институт проблем управления, лаб. 20, Оргкомитет международной конференции; тел. (495) 334-89-59, e-mail: Conf20@ipu.rssi.ru



УПРАВЛЕНИЕ ВЫБОРОМ ПРОИЗВОДИТЕЛЯ С УЧЕТОМ ВЛИЯНИЯ КОНЪЮНКТУРНОЙ ИНФОРМАЦИИ

А.А. Байкин, Е.Ю. Иванов, О.В. Исаева

Алтайский государственный университет, г. Барнаул

Приведены результаты исследования модифицированной с учетом цены информации модели выбора производителя, проведенные для некоторых видов производственных функций (линейной, Леонтьева, мультипликативной, постоянной эластичности замены, Филлипова). Изложены оценки характера и степени влияния дополнительной конъюнктурной информации на изменение структуры и объемов используемых фирмой ресурсов.

ВВЕДЕНИЕ

Большинство экономико-математических моделей, описывающих взаимодействие экономических субъектов как на макро-, так и на микроуровне, абстрагированы от учета информационных воздействий. Так, из восемнадцати аксиоматических предположений, на которых строится классическая *economics*, пять гипотез (Н1, Н3, Н13, Н14, Н17) [1] непосредственно связаны с информацией. Обобщив предположения этих гипотез о совершенстве информации о существующих товарах, а также о возможностях выбора и ценах, выделим информацию:

- о технологических характеристиках ресурса;
- о цене ресурса;
- о текущем собственнике ресурса и его месторасположении.

Такая информация появляется с возникновением любого производственного ресурса и может иметь собственную сферу обращения.

Допуская, что производителю хорошо известно качество представленных на рынке ресурсов и возможности их вовлечения в производство (т. е. принимая гипотезы Н13, Н14), будем называть информацию о цене ресурса и его месторасположении конъюнктурной информацией о рынке факторов производства или просто *конъюнктурной информацией*.

Влияние различных экономических воплощений информации на рыночное равновесие отмечалось многими исследователями, указывающими, в частности, на необходимость коррекции «закона единой цены» с учетом стоимости конъюнктурной

информации [2, 3, с. 351—354]. Наиболее заметно упущение информационных аспектов стало проявляться при исследовании случаев «фиаско рынка», одним из которых является рынок с асимметричным распределением информации между потребителями и продавцами [4]. На подобных рынках для достижения большей эффективности необходимо понести дополнительные издержки на приобретение более качественных ресурсов, фактически — это издержки на приобретение конъюнктурной информации. Таким образом, формализация влияния информационных факторов для традиционных микроэкономических моделей становится актуальной проблемой современной экономической теории.

1. КЛАССИЧЕСКАЯ ЗАДАЧА ВЫБОРА ПРОИЗВОДИТЕЛЯ И ЕЕ МОДИФИКАЦИЯ С УЧЕТОМ КОНЪЮНКТУРНОЙ ИНФОРМАЦИИ

Производство обычно представляется как процесс вовлечения некоторого набора благ-факторов производства и преобразование их в благо-продукт. Число используемых в производстве факторов всегда ограничено сверху, исходя из свойства редкости всех существующих благ [5].

Теория фирмы предполагает, что производственный процесс можно формализовать с помощью производственной функции (ПФ), описывающей технологию, т. е. правило преобразования используемых факторов производства в конечный продукт. Поскольку трудно привести детальное описание всех технологических процессов, происходящих на каждом предприятии, имеющееся многообразие технологий может быть представлено

множеством X , содержащим набор векторов x , соответствующих технологически возможным преобразованиям затрачиваемых ресурсов в выпускаемую продукцию, иными словами, технологические ограничения формализуются как $x \in X$. Допуская существование функции, которая дает максимальный физический объем выпуска продукции при всех технически возможных комбинациях физических ресурсов и при данном уровне свободно распространяемого технического знания о связи между ресурсами и выпуском, можно записать ПФ:

$$f(x_1, \dots, x_n) = 0, \quad (1)$$

где f — некоторая действительная функция, определенная на множестве возможных выпусков чистой продукции, состоящем из векторов чистой продукции, которую производитель в состоянии выпускать, а $x = (x_1, \dots, x_n)$ — эффективный вектор существующих на рынке благ, общее число которых n .

Необходимо отметить, что ПФ фирмы определяется на том же самом множестве товаров и услуг, что и функция полезности потребителя. Тем самым любые производимые фирмой блага являются теми же самыми благами, которые покупаются конечными потребителями. Кроме того, подобная формализация деятельности фирмы одинаково подходит как для строгой дополняемости производственных ресурсов (фиксированные технологические коэффициенты, задающие строгие пропорции использования ресурсов), так и для процессов с широкими возможностями замещения (один из ресурсов произвольно замещается каким-либо дополнительным количеством других).

В дальнейшем для упрощения рассуждений перейдем от общей формы ПФ (1) к ее частному представлению, для которого предполагается, что каждый производитель одновременно выступает и продавцом своей продукции, выпуская при этом только одно благо (подробнее см. работу [3, с. 54–55]). Такая частная ПФ имеет форму:

$$x_n = y(x_1, \dots, x_{n-1}), \quad (2)$$

где y — функция, описывающая технологические ограничения эффективного производства, n — порядковый номер блага, выпускаемого данной фирмой. Тогда технологические ограничения записываются как $x_n \leq y(x_1, \dots, x_{n-1})$.

Из выражения (2) получаем еще одну широко распространенную трактовку термина «производственная функция», которая характеризует функцию y как зависимость объема выпуска фирмы от

затрат первичных ресурсов. Основные свойства ПФ вида (2) описаны, в частности, в работе [6].

Оптимальный производственный выбор для фирмы в терминах *economics* определяется только тогда, когда производитель определил цели или критерий оптимальности. Наиболее известное, но не единственное предположение о целях деятельности предприятий заключается в *максимизации фирмой прибыли* — в этом случае фирма стремится получить максимально возможную разницу между доходами и расходами. Вместе с тем существует ряд других интересных гипотез о целях деятельности фирмы (см., например, работу [7]).

Задача оптимизации, соответствующая выбранной цели деятельности фирмы, может рассматриваться на различных временных интервалах, что вызывает необходимость поиска принципиально различных решений для каждой из этих задач в зависимости от выбранного периода времени. Основной различий между кратковременным и долгосрочным периодами деятельности фирмы будем традиционно считать возможность изменения объемов использования всех производственных ресурсов. Тогда *краткосрочный период времени* — период производства, в течение которого все ресурсы, за исключением одного, неизменные, и, следовательно, весь прирост объема производства связан с приростом использования именно данного фактора, *долгосрочный период времени* — период, в течение которого производитель может изменить все факторы производства выпускаемой продукции. Долгосрочный период обычно рассматривают как последовательно сменяющиеся друг друга краткосрочные периоды.

Принимая в качестве цели деятельности производителя классическую гипотезу максимизации прибыли, рассмотрим три случая задачи оптимизации производственного выбора (выбора фирмы):

- 1) максимизация чистой прибыли (разницы между доходами и расходами фирмы) при заданной ПФ;
- 2) максимизация выпуска продукта при заданном уровне затрат;
- 3) минимизация затрат при заданном уровне выпуска продукта.

Переходя к интерпретации производственной функции как зависимости объема выпуска от затрат первичных ресурсов, сформулируем следующую модель для долгосрочного периода. Пусть $n - 1$ — число факторов производства, $x = (x_1, \dots, x_{n-1})$ — некоторая технология производства, определяющая объемы вовлекаемых в производство ресурсов, евклидово пространство с неотрицательными координатами X — рынок факторов производства: $X = \{x: x \geq 0\}$. Тогда формализация первого



случая задачи оптимизации приводит к следующей классической задаче на условный экстремум:

$$\begin{cases} \max \sum_{i=1}^n p_i x_i = \max \left(p_n x_n - \sum_{i=1}^{n-1} p_i x_i \right) \\ x_n = y(x_1, \dots, x_{n-1}), \quad x_i \geq 0, \end{cases} \quad (3)$$

где $p = (p_1, \dots, p_n)$ — вектор цен всех существующих благ, n — порядковый номер блага, выпускаемого рассматриваемой фирмой. В дальнейшем задача (3) будет рассматриваться при нахождении выбора фирмы в долгосрочном периоде, а ее решение будем обозначать $x^* = (x_1^*, \dots, x_{n-1}^*)$.

Теперь формализуем задачу производителя для краткосрочного периода в условиях заранее заданного объема оборотных средств, которые фирма может потратить на приобретение переменного фактора. Пусть C — объем оборотных средств, доступных предприятию, множество факторов производства, которые может приобрести производитель на рынке ресурсов, понеся в данный временной период полные затраты C , назовем множеством *финансово доступных* комбинаций производственных факторов¹ $B: B = \{x: px \leq C\}$, где $p = (p_1, \dots, p_{n-1})$ — вектор цен на ресурсы. Имеем:

$$\begin{cases} \max_{x \in X} u(x), \\ px \leq C, \quad x \geq 0. \end{cases} \quad (4)$$

Задача (4) в дальнейшем будет рассматриваться при нахождении выбора фирмы в краткосрочном периоде.

Формализация третьего случая задачи оптимизации производственного выбора фирмы приводит к двойственной задаче по отношению ко второму случаю, поэтому отдельно рассматриваться не будет.

Аналитическое решение задач (3) и (4) может быть получено методом неопределенных множителей Лагранжа для соответствующей дифференцируемой функции и позволяет определить функции спроса на ресурсы $x^* = x^*(p, C)$. Полученные таким образом решения, как уже отмечалось, основаны на равномерном распределении информации среди всех участников рынка и на действии закона единой цены. Реальный рынок зачастую демонстрирует различную конъюнктурную осведомленность производителей (что особенно характерно для малых и средних предприятий) и существование ценовых интервалов, внутри которых содержатся цены реальных транзакций производ-

ственных ресурсов. С учетом цены конъюнктурной информации, модифицированная модель выбора фирмы в краткосрочном периоде может быть формализована в следующем виде:

$$\begin{cases} \max_{x \in X} u(x), \\ p^* x \leq C, \quad x \geq 0, \end{cases} \quad (5)$$

где $p^* = (p_1 - \Delta p_1, \dots, p_{n-1} - \Delta p_{n-1})$.

В этом случае цена фактора производства складывается из собственно цены этого фактора и цены конъюнктурной информации об этом факторе. Таким образом, информированный производитель фактически уменьшает цену производственного фактора, например, путем повышения эффективности его использования либо путем экономии транзакционных издержек на его приобретение. Подобное снижение цены одного из ресурсов можно интерпретировать как увеличение оборотных средств фирмы благодаря расширению множества доступных комбинаций производственных факторов, что, в свою очередь, влияет на выбор производителя в краткосрочном периоде времени. Иными словами, решения задач (3) и (4) представляют собой выбор «неинформированного»² производителя, решение же задачи (5) — это новый выбор фирмы с учетом полученной производителем информации.

Попытаемся определить, как изменится решение задач (3) и (4) при изменении исходного вектора цен ресурсов p в некоторый вектор p^* , где $p_n = p_n^*$ — цена выпускаемого продукта. Решая задачу (5) методом множителей Лагранжа, получаем оптимальный набор ресурсов $x^{**} = x^{**}(p^*, C)$ для модифицированной задачи.

Заметим, что в точке локального рыночного равновесия предельная норма технической замены одного ресурса другим равна отношению их рыночных цен

$$\frac{\frac{\partial y}{\partial x_i} \Big|_{x_i = x^{**}}}{\frac{\partial y}{\partial x_j} \Big|_{x_j = x^{**}}} = \frac{p_i - \Delta p_i}{p_j - \Delta p_j}, \quad \forall i = \overline{1, n-1}, \\ j = \overline{1, n-1}, \quad i \neq j. \quad (6)$$

Рассматривая задачу (3) и полагая преобразование вектора цен всех благ p в вектор $p^* = (p_1 - \Delta p_1, \dots, p_{n-1} - \Delta p_{n-1}, p_n)$, получаем задачу для инфор-

¹ Количественной оценкой доступности ресурсов служит понятие изокосты.

² Точнее, производителя, обладающего одинаковой конъюнктурной информацией со всеми прочими производителями.

мированного производителя в долгосрочном периоде:

$$\begin{cases} \max \left(p_n x_n - \sum_{i=1}^{n-1} p_i^* x_i \right), \\ x_n = y(x_1, \dots, x_{n-1}), x_i \geq 0. \end{cases} \quad (7)$$

Дальнейшие общие рассуждения относительно выбора информированного производителя затруднительны, так как абсолютная величина выражения (6) не позволяет качественно интерпретировать полученное соотношение. Однако можно явным образом рассмотреть различные типы производственных функций, наиболее адекватно описывающих частные случаи на конкретном предприятии. Поэтому далее проанализируем некоторые наиболее популярные виды производственных функций [8, с. 1–3] для моделирования поведения фирмы в условиях изменившейся информированности.

Предварительно отметим, что без потери общности задачи (5) и (7) можно рассматривать как выбор фирмы в условиях двухресурсного производства. Пусть тогда ресурс x_1 будет переменным, а остальные ресурсы, информированность о которых не изменяется, будут представлены агрегированным фактором x_2 . Подобная двухпродуктовая модель позволит не только упростить математические выкладки, но и наглядно продемонстрировать полученные результаты. Геометрически решение $x^* = (x_1^*, x_2^*)$ можно интерпретировать как точку касания изокванты производственной функции с одной из изокост, являющейся линией уровня функции издержек.

2. МОДИФИЦИРОВАННАЯ ЗАДАЧА ВЫБОРА ПРОИЗВОДИТЕЛЯ ДЛЯ ЛИНЕЙНОЙ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ

В случае использования фирмой в качестве ресурсов совершенных субститутов производственный процесс можно описать линейной производственной функцией вида $y(x_1, \dots, x_{n-1}) = \sum_{i=1}^{n-1} \alpha_i x_i$, где α_i — некоторые положительные числа. В долгосрочном периоде для двух факторов производства имеем:

$$\begin{cases} \max(p_3 x_3 - p_1 x_1 - p_2 x_2), \\ x_3 = \alpha_1 x_1 + \alpha_2 x_2, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases}$$

Поскольку ни одна из частных производных целевой функции не обращается в ноль, то чем больше будет вовлечено в производственный процесс ресурсов, тем больше будет прибыль.

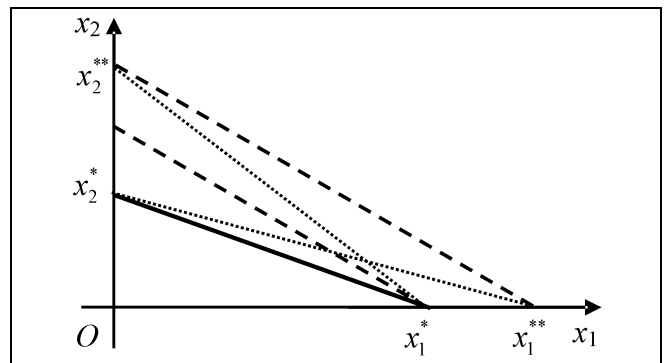


Рис. 1. Выбор производителя для линейной производственной функции

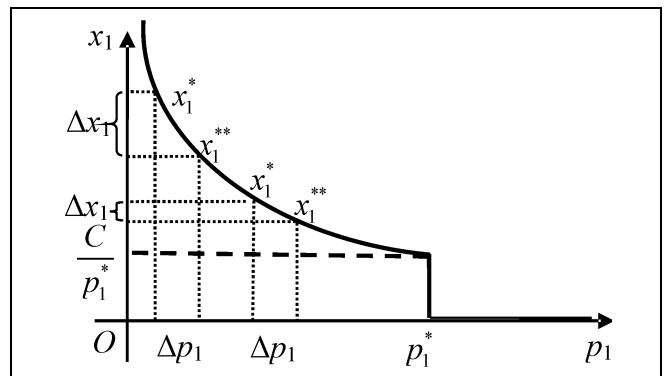


Рис. 2. Кривая спроса на ресурс для линейной производственной функции

В краткосрочном периоде задача (4) для линейной ПФ принимает вид

$$\begin{cases} \max(\alpha_1 x_1 + \alpha_2 x_2), \\ p_1 x_1 + p_2 x_2 \leq C, x_1 \geq 0, x_2 \geq 0, \end{cases}$$

а ее решение представлено на рис. 1. Изокванты (штриховые линии) представляют собой семейство параллельных прямых с угловым коэффициентом $-\alpha_1/\alpha_2$. Очевидно, что при разных наклонах изокосты (сплошная линия) и изокванты выбор производителя получается в одной из двух угловых точек (x_1^* или x_2^* на рис. 1), а функция спроса (рис. 2) на первый фактор производства принимает вид:

$$x_1^* = \begin{cases} C/p_1, & p_2 \alpha_1 > p_1 \alpha_2, \\ Ct/p_1, & p_2 \alpha_1 = p_1 \alpha_2, \quad t \in [0; 1], \\ 0, & p_2 \alpha_1 < p_1 \alpha_2. \end{cases}$$

Таким образом, для линейных производственных функций при увеличении информированности производителя о выбранном факторе производства (его цена при этом уменьшается) расширяет



исходное множество $x_2^* O x_1^*$ финансово доступных комбинаций производственных факторов до $x_2^{**} O x_1^{**}$. Соответственно, новая информация о ранее выбранном факторе влияет только на объем использования этого фактора (точка x_1^{**} на рис. 1), но не влияет на то, какой именно фактор производства будет применяться. Вместе с тем, повышение информированности производителя о ценах на другие ресурсы может изменить структуру его выбора, если конъюнктурная информация изменит исходное множество $x_2^* O x_1^*$ до состояния $x_2^{**} O x_1^{**}$ (см. рис. 1).

3. МОДИФИЦИРОВАННАЯ ЗАДАЧА ВЫБОРА ПРОИЗВОДИТЕЛЯ ДЛЯ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ ЛЕОНТЬЕВА

В случае использования фирмой ресурсов с постоянными пропорциями (совершенных дополнителей) производственный процесс можно описать функцией Леонтьева: $y(x_1, \dots, x_{n-1}) = \min\{\alpha_1 x_1, \dots, \alpha_n x_{n-1}\}$, где α_i — некоторые положительные числа.

Заметим, что задача (3) с ПФ Леонтьева принимает вид

$$\begin{cases} \max(p_3 x_3 - p_1 x_1 - p_2 x_2) \\ x_3 = \min\{\alpha_1 x_1; \alpha_2 x_2\}, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$$

и не имеет глобального экстремума. Следовательно, чем больше будет вовлечено в производственный процесс ресурсов, тем больше будет прибыль.

В краткосрочном периоде решение задачи

$$\begin{cases} \max(\min\{\alpha_1 x_1, \alpha_2 x_2\}), \\ p_1 x_1 + p_2 x_2 \leq C, x_1 \geq 0, x_2 \geq 0 \end{cases}$$

для ПФ Леонтьева с двумя факторами производства представлено на рис. 3. Изокванты (штриховые линии) представляют собой семейство углов, вершины которых находятся на прямой $x_2 = (\alpha_1/\alpha_2)x_1$. Поскольку производной ПФ Леонтьева в точке выбора производителя не существует, то, учитывая равенство $\alpha_1 x_1 = \alpha_2 x_2$ и ограничение на доступные комбинации производственных факторов, получаем следующие функции спроса: $x_1^* = \frac{\alpha_2 C}{\alpha_2 p_1 + \alpha_1 p_2}$,

$$x_2^* = \frac{\alpha_1 C}{\alpha_2 p_1 + \alpha_1 p_2}.$$

Оценим изменение спроса на ресурсы вследствие возможного изменения цен. Так как эlasticность спроса на определенный ресурс по цене на этот ресурс $E_{p_1}^{x_1} = \frac{\partial x_1^*}{\partial p_1} \frac{p_1}{x_1^*} = -\left(1 - \frac{\alpha_1 p_2}{\alpha_2 p_1 + \alpha_1 p_2}\right) < 0$,

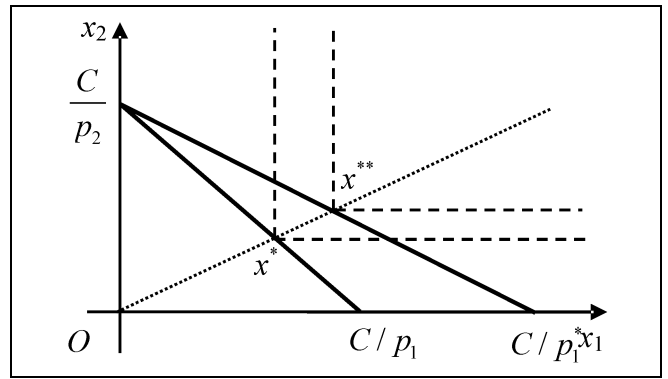


Рис. 3. Выбор производителя для производственной функции Леонтьева

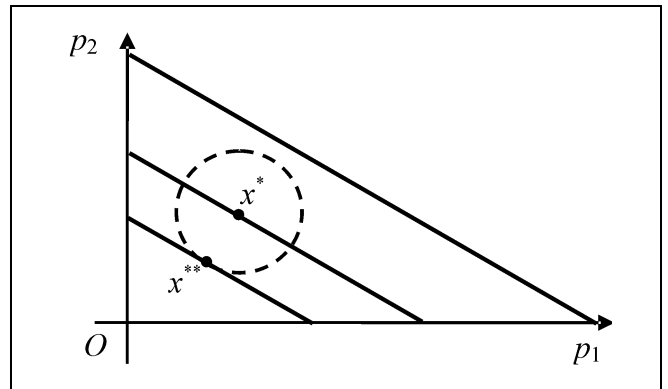


Рис. 4. Линии уровня функций спроса на ресурсы для производственной функции Леонтьева

а перекрестная эlasticность спроса по цене

$$E_{p_2}^{x_1} = \frac{\partial x_1^*}{\partial p_2} \frac{p_2}{x_1^*} = -\left(1 - \frac{\alpha_2 p_1}{\alpha_2 p_1 + \alpha_1 p_2}\right) < 0, \text{ то в общем}$$

случае имеем неэlasticный спрос.

Таким образом, для ПФ Леонтьева увеличение информированности производителя о ценах не будет влиять на увеличение использования ресурсов до тех пор, пока не появится возможность приобрести еще один комплект всех взаимодополняющих факторов производства.

Наглядно изменение выбора производителя с ростом его информированности можно представить, построив линии уровня поверхности спроса на ресурсы. Линии уровня $x_i^*(p_1, p_2, C) = \text{const}$ образуют семейство параллельных прямых с угловым коэффициентом $-\alpha_2/\alpha_1$, причем по мере увеличения объемов спроса на ресурсы расстояние между ними уменьшается (рис. 4). Таким образом, с ростом объемов использования факторов производства цена информации оказывает все большее влияние на изменение выбора производителя.

4. МОДИФИЦИРОВАННАЯ ЗАДАЧА ВЫБОРА ПРОИЗВОДИТЕЛЯ ДЛЯ МУЛЬТИПЛИКАТИВНЫХ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ

Среди мультипликативных ПФ одна из наиболее популярных функция Кобба—Дугласа, которая в общем случае имеет вид $y(x_1, \dots, x_{n-1}) = \mu \prod_{i=1}^{n-1} x_i^{\alpha_i}$, где коэффициенты $\alpha_i > 0$ отражают эластичность выпуска по отношению к объемам использования факторов производства, μ — коэффициент, отражающий влияние иных факторов (например, макроэкономических).

Для случая двух факторов производства в долгосрочном периоде получаем задачу

$$\begin{cases} \max(p_3x_3 - p_1x_1 - p_2x_2), \\ x_3 = \mu x_1^{\alpha_1} x_2^{\alpha_2}, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases}$$

Наличие максимума у функции прибыли фирмы, выпуск которой описывается ПФ Кобба—Дугласа, эквивалентно условиям:

$$\begin{aligned} \frac{\partial^2 \pi(x_1, x_2)}{\partial x_1^2} &= p_3 \mu \alpha_1 (\alpha_1 - 1) x_1^{\alpha_1 - 2} x_2^{\alpha_2} < 0, \\ \Delta &= \frac{\partial^2 \pi}{\partial x_1^2} \cdot \frac{\partial^2 \pi}{\partial x_2^2} - \left(\frac{\partial^2 \pi}{\partial x_1 \partial x_2} \right)^2 = \\ &= p_3^2 \mu^2 \alpha_1 \alpha_2 x_1^{2(\alpha_1 - 1)} x_2^{2(\alpha_2 - 1)} (1 - \alpha_1 - \alpha_2) > 0, \end{aligned} \quad (8)$$

где $\pi(x_1, x_2) = p_3 x_1^{\alpha_1} x_2^{\alpha_2} - p_1 x_1 - p_2 x_2$ — прибыль фирмы.

Поскольку знак выражения (8) определяется знаком множителя $(1 - \alpha_1 - \alpha_2)$, то при эластичности производства (определяемой выражением $\sum_{i=1}^{n-1} \alpha_i$) меньше единицы локальный максимум прибыли существует.

Если эластичность производства равна единице, то безубыточность деятельности фирмы, определяемая соотношением $p_3 \mu = \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1} \left(\frac{p_2}{\alpha_2}\right)^{\alpha_2}$ непосредственно зависит от конъюнктуры рынка. В случае строгого выполнения равенства получаем ситуацию с нулевой прибылью, описываемую в большинстве учебников *economics*. Если правая часть выражения будет больше левой, можно путем выбора оптимальной структуры используемых ресурсов постоянно увеличивать размер прибыли, иначе же фирма обречена на банкротство.

Если эластичность производства больше единицы, то технология производства и конъюнктура

рынка позволяют получить неограниченный рост прибыли с увеличением объема вовлекаемых ресурсов только после прохождения некоторой критической точки вовлечения ресурсов.

Таким образом, имеет смысл определять эффективную комбинацию производственных факторов только для случая $\sum_{i=1}^{n-1} \alpha_i < 1$. Получающиеся при этом функции спроса на ресурсы имеют вид:

$$\begin{aligned} x_1^* &= \\ &= (p_3 \mu)^{\frac{1}{1 - (\alpha_1 + \alpha_2)}} (\alpha_1 / p_1)^{\frac{1 - \alpha_2}{1 - (\alpha_1 + \alpha_2)}} (\alpha_2 / p_2)^{\frac{\alpha_2}{1 - (\alpha_1 + \alpha_2)}}, \\ x_2^* &= \\ &= (p_3 \mu)^{\frac{1}{1 - (\alpha_1 + \alpha_2)}} (\alpha_2 / p_2)^{\frac{1 - \alpha_1}{1 - (\alpha_1 + \alpha_2)}} (\alpha_1 / p_1)^{\frac{\alpha_1}{1 - (\alpha_1 + \alpha_2)}}. \end{aligned}$$

В случае использования оптимальной комбинации факторов производства прибыль фирмы выражается в виде $\pi(x_1^*, x_2^*) = \frac{p_2}{\alpha_2} (1 - \alpha_1 - \alpha_2) \times$
 $\times (p_3 \mu \alpha_1 / p_1)^{\frac{1}{1 - (\alpha_1 + \alpha_2)}} (\alpha_2 p_1 / \alpha_1 p_2)^{\frac{\alpha_2}{1 - (\alpha_1 + \alpha_2)}}$, откуда следует, что эффективная комбинация производственных факторов определяет строго положительное значение прибыли.

Прямая эластичность спроса по цене $E_{p_1}^{x_1} = \frac{\partial x_1}{\partial p_1} \cdot \frac{p_1}{x_1} = \frac{\alpha_2 - 1}{1 - (\alpha_1 + \alpha_2)}$, а поскольку справедлива оценка $0 > E_{p_1}^{x_1} = -1 - \frac{\alpha_1}{1 - (\alpha_1 + \alpha_2)} > -1$, то спрос эластичен. Перекрестная эластичность спроса по цене равна $E_{p_2}^{x_1} = \frac{\partial x_1}{\partial p_2} \cdot \frac{p_2}{x_1} = \frac{-\alpha_2}{1 - (\alpha_1 + \alpha_2)}$, а поскольку $E_{p_2}^{x_1} < 0$, то спрос может быть как эластичным, так и нет.

Решение задачи

$$\begin{cases} \max \mu x_1^{\alpha_1} x_2^{\alpha_2}, \\ p_1 x_1 + p_2 x_2 \leq C, x_1 \geq 0, x_2 \geq 0 \end{cases}$$

в краткосрочном периоде определяет функции спроса на ресурсы в виде $x_i^* = \frac{\alpha_i C}{p_i (\alpha_1 + \alpha_2)}$. При этом

величина $\sum_{i=1}^{n-1} \alpha_i$, в отличие от долгосрочного случая, не оказывает влияния на полученный результат.

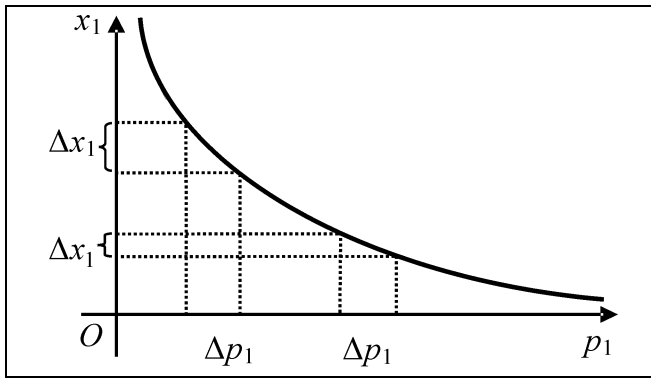


Рис. 5. Функция спроса на ресурс для мультипликативной производственной функции

Далее, очевидно, что повышение информированности только об одном факторе производства не изменит уровень использования других факторов производства, поскольку $x_i^{**} = \frac{\alpha_i C}{(p_1 - \Delta p_i)(\alpha_1 + \alpha_2)}$.

Иными словами, если деятельность фирмы описывается ПФ Кобба—Дугласа, то увеличение информированности производителя приводит к изменению спроса только на этот ресурс (рис. 5), не изменяя при этом объемы вовлечения в производство других факторов.

Одним из обобщений функции Кобба—Дугласа является функция Стоуна $y(x_1, \dots, x_{n-1}) = \mu \prod_{i=1}^{n-1} (x_i - x_i^0)^{\alpha_i}$, где α_i — все также эластичность выпуска по ресурсам, а x_i^0 — минимально необходимый уровень вовлечения ресурса i в производственный процесс.

Решая задачу

$$\begin{cases} \max(p_3 x_3 - p_1 x_1 - p_2 x_2), \\ x_3 = \mu (x_1 - x_1^0)^{\alpha_1} (x_2 - x_2^0)^{\alpha_2}, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$$

в долгосрочном периоде при условии $\alpha_1 + \alpha_2 < 1$, получаем следующие функции спроса:

$$\begin{aligned} x_1^* &= x_1^0 + \\ &+ (p_3 \mu)^{\frac{1}{1-(\alpha_1+\alpha_2)}} \left(\frac{\alpha_1}{p_1}\right)^{\frac{1-\alpha_2}{1-(\alpha_1+\alpha_2)}} \left(\frac{\alpha_2}{p_2}\right)^{\frac{\alpha_2}{1-(\alpha_1+\alpha_2)}}, \\ x_2^* &= x_2^0 + \\ &+ (p_3 \mu)^{\frac{1}{1-(\alpha_1+\alpha_2)}} \left(\frac{\alpha_2}{p_2}\right)^{\frac{1-\alpha_1}{1-(\alpha_1+\alpha_2)}} \left(\frac{\alpha_1}{p_1}\right)^{\frac{\alpha_1}{1-(\alpha_1+\alpha_2)}}. \end{aligned} \quad (9)$$

Составляя и решая задачу

$$\begin{cases} \max \mu (x_1 - x_1^0)^{\alpha_1} (x_2 - x_2^0)^{\alpha_2}, \\ p_1 x_1 + p_2 x_2 \leq C, x_1 \geq 0, x_2 \geq 0, \end{cases}$$

для краткосрочного периода получаем соответствующие функции спроса на ресурсы:

$$x_i^* = x_i^0 + \frac{\alpha_i}{(\alpha_1 + \alpha_2)} \cdot \frac{C - (p_1 x_1^0 + p_2 x_2^0)}{p_i}, \quad i = 1, 2. \quad (10)$$

Интерпретируя полученные решения (9) и (10), можно отметить, что как в краткосрочном, так и в долгосрочном периодах производитель сначала должен приобрести минимально необходимое количество каждого из ресурсов, а затем эти ресурсы вовлечь в производство пропорционально их «желательности» и обратно пропорционально цене.

Проанализируем полученные функции (10) как решение задачи (5). Прямая эластичность спроса по цене $E_{p_1}^{x_1} = \frac{\alpha_1 (C - p_2 x_2^0)}{\alpha_1 (C - p_2 x_2^0) + \alpha_2 p_1 x_1^0}$, а поскольку справедлива оценка $0 > E_{p_1}^{x_1} = -1 +$

$+\frac{\alpha_2 p_1 x_1^0}{\alpha_1 (C - p_2 x_2^0) + \alpha_2 p_1 x_1^0} > -1$, то спрос неэластичен. Перекрестная эластичность спроса по цене $E_{p_2}^{x_1} = -\frac{\alpha_1 p_2 x_2^0}{\alpha_1 (C - p_2 x_2^0) + \alpha_2 p_1 x_1^0} < 0$, поэтому спрос

может быть как эластичным, так и нет. Геометрически решение задачи (4) представлено на рис. 6. Расположение изоквант аналогично схеме функции Кобба—Дугласа, однако ветви кривых имеют асимптоты, не совпадающие с координатными осями.

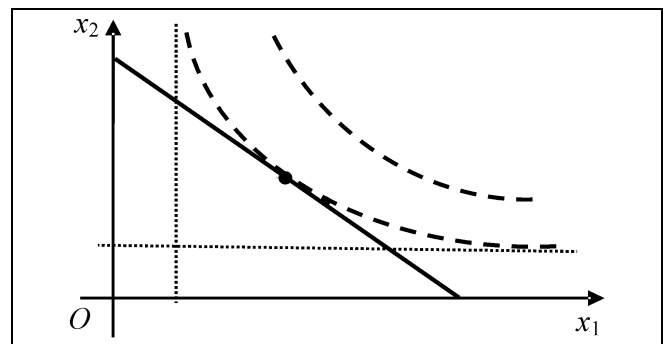


Рис. 6. Выбор производителя для мультипликативной производственной функции

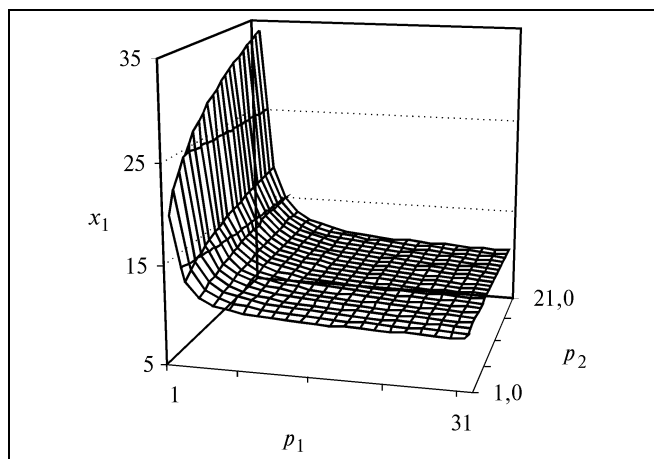


Рис. 7. Кривая спроса гипотетической фирмы на ресурс для производственной функции Стоуна

Определим границы влияния цены информации на объем вовлечения ресурса, исходя из того, что его использование изменится хотя бы на одну условную единицу. Последнее означает, что изменение в структуре потребления зависит не только от информированности производителя, но и от объема использования данного ресурса. Другими словами, при малом уровне использования фактора в производстве информация оказывается менее ценной, чем при большом, поскольку одна и та же единица конъюнктурной информации дает больший прирост его использования (рис. 7).

Таким образом, для производства, описываемого ПФ Стоуна, увеличение информированности производителя относительно одного ресурса приводит к изменению объема использования всех факторов, однако наибольшее влияние все же оказывается на объем вовлечения в производственный процесс именно данного ресурса.

5. МОДИФИЦИРОВАННАЯ ЗАДАЧА ВЫБОРА ПРОИЗВОДИТЕЛЯ ДЛЯ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ С ПОСТОЯННОЙ ЭЛАСТИЧНОСТЬЮ ЗАМЕЩЕНИЯ

Практически столь же часто, как и функция Кобба—Дугласа, для описания производственных процессов применяется функция с постоянной эластичностью замещения или CES-функция:

$$y(x_1, \dots, x_{n-1}) = \mu \left(\sum_{i=1}^{n-1} \frac{\beta_i}{x_i^\alpha} \right)^{-1/\alpha}, \text{ где } 0 < \beta_i < 1, \alpha > 0,$$

μ — так же, как и в мультипликативных функциях, коэффициент, отражающий влияние прочих факторов.

Поскольку функция прибыли для задачи (3), имеющей в данном случае вид

$$\begin{cases} \max \{ p_3 x_3 - p_1 x_1 - p_2 x_2 \}, \\ x_3 = \mu \left(\frac{\beta}{x_1^\alpha} + \frac{1-\beta}{x_2^\alpha} \right)^{-1/\alpha}, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, \end{cases}$$

не имеет локального максимума, то чем больше будет вовлечено в производственный процесс ресурсов, тем больше будет прибыль.

Геометрически решение задачи

$$\begin{cases} \max \mu \left(\frac{\beta}{x_1^\alpha} + \frac{1-\beta}{x_2^\alpha} \right)^{-1/\alpha}, \\ p_1 x_1 + p_2 x_2 \leq C, x_1 \geq 0, x_2 \geq 0 \end{cases} \text{ в краткосрочном}$$

периоде представлено на рис. 8. Расположение изоквант (штриховые линии) аналогично схеме функции Стоуна, однако у каждой изокванты свои асимптоты, определяемые параметрами CES-функции: $x_1 = \frac{C}{\mu} \beta^{1/\alpha}$, $x_2 = \frac{C}{\mu} (1-\beta)^{1/\alpha}$, где $C = x_3 =$

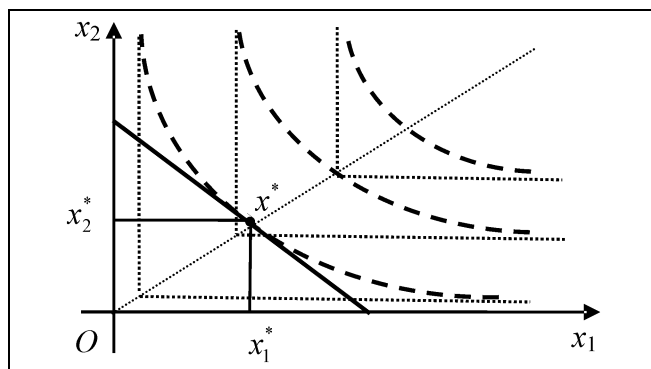


Рис. 8. Выбор производителя для CES-функции

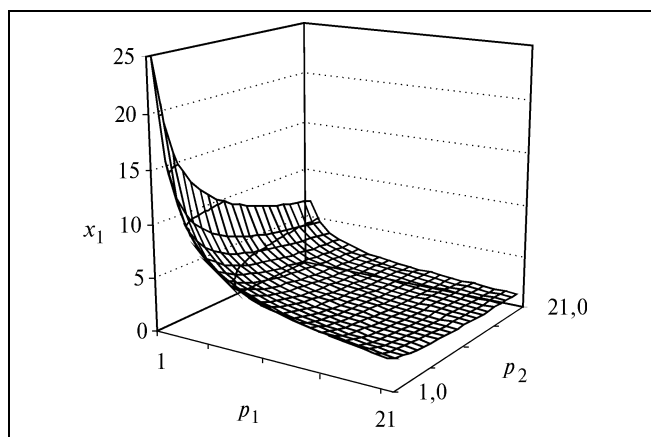


Рис. 9. Функция спроса гипотетической фирмы для CES-функции



= const. Аналитическое решение задачи (4) определяет функции спроса на факторы производства:

$$x_1^* = \frac{C\beta^\sigma}{p_1^\sigma(\beta^\sigma p_1^{1-\sigma} + (1-\beta)^\sigma p_2^{1-\sigma})};$$

$$x_2^* = \frac{C(1-\beta)^\sigma}{p_2^\sigma(\beta^\sigma p_1^{1-\sigma} + (1-\beta)^\sigma p_2^{1-\sigma})},$$

где $\sigma = (1 + \alpha)^{-1}$ — эластичность замещения факторов производства для данной ПФ [9]. График функции спроса представлен на рис. 9.

Таким образом, как и в случае ПФ Стоуна, производство, описываемое CES-функцией, реагирует на увеличение информированности производителя относительно одного ресурса изменением объема использования всех факторов.

6. МОДИФИЦИРОВАННАЯ ЗАДАЧА ВЫБОРА ПРОИЗВОДИТЕЛЯ ДЛЯ ПРОИЗВОДСТВЕННОЙ ФУНКЦИИ ФИЛИПОВА³

В некоторых случаях производственный процесс следует описывать с помощью s-образной производственной функции вида $y(x_1, \dots, x_{n-1}) =$

$$= \sum_{i=1}^{n-1} (\alpha_i \sin \beta_i x_i + \gamma_i x_i), \quad \alpha_i \geq 0, \beta_i > 0, \gamma_i > 0.$$

График объема выпуска для такой функции в случае двух факторов производства приведен на рис. 10. Допуская возможность такого описания производственного процесса, можно отказаться от принципов убывающей продуктивности факторов производства и убывающей предельной полезности [10].

Предположение ненасыщаемости позволяет определить важное соотношение для коэффициентов производственной функции данного вида: $\alpha_i \beta_i \leq \gamma_i$.

Для случая двух факторов производства получим следующую модель поведения фирмы в долгосрочном периоде:

$$\begin{cases} \max \{ p_3 x_3 - p_1 x_1 - p_2 x_2 \}, \\ x_3 = \alpha_1 \sin \beta_1 x_1 + \gamma_1 x_1 + \alpha_2 \sin \beta_2 x_2 + \gamma_2 x_2, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases}$$

Существование локального максимума прибыли фирмы зависит от одновременного выполнения условий $|p_i - \gamma_i| \leq p_3 \alpha_i \beta_i, i = 1, 2$. В противном слу-

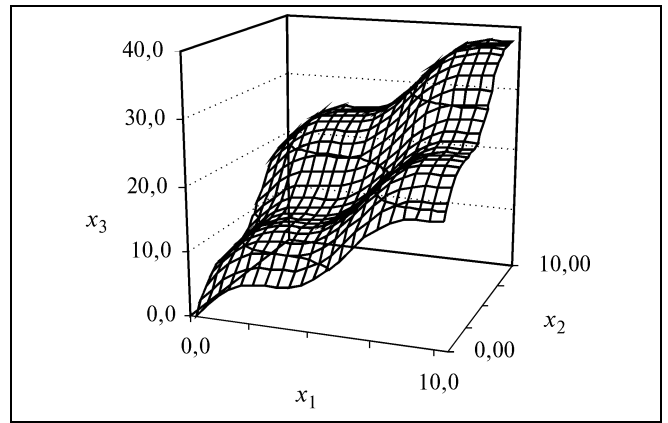


Рис. 10. Объем выпуска гипотетической фирмы для производственной функции Филиппова

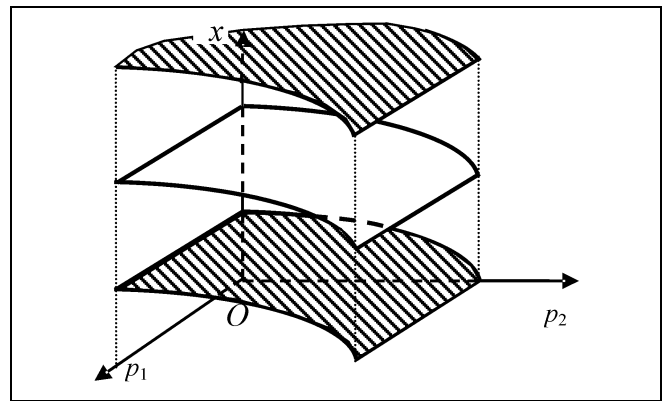


Рис. 11. Функция спроса для производственной функции Филиппова

чае имеем тривиальный случай, когда размер прибыли неограниченно увеличивается пропорционально объемам вовлекаемых ресурсов.

Таким образом, получаем функции спроса на ресурсы:

$$x_i^* = \frac{1}{\beta_i} \arccos \left(\frac{p_i - \gamma_i}{p_3 \alpha_i \beta_i} \right) + \frac{2\pi m}{\beta_i}, \quad i = 1, 2,$$

$$m \in \{0\} \cup N,$$

где N — множество натуральных чисел. График функции спроса в этом случае представлен на рис. 11.

Геометрическое решение задачи

$$\begin{cases} \max(\alpha_1 \sin \beta_1 x_1 + \gamma_1 x_1 + \alpha_2 \sin \beta_2 x_2 + \gamma_2 x_2), \\ p_1 x_1 + p_2 x_2 \leq C, \quad x_1 \geq 0, \quad x_2 \geq 0 \end{cases}$$

в краткосрочном периоде представлено на рис. 12. Изокванты (штриховые линии) в нескольких точках соприкасаются с границей множества финансово доступных комбинаций производственных

³ Данная функция названа в честь Леонида Андреевича Филиппова, известного алтайского экономиста-математика, разработавшего метод интервальных расчетов Вексичкого и предложившего один из более ранних вариантов записи s-образных функций подобного вида.

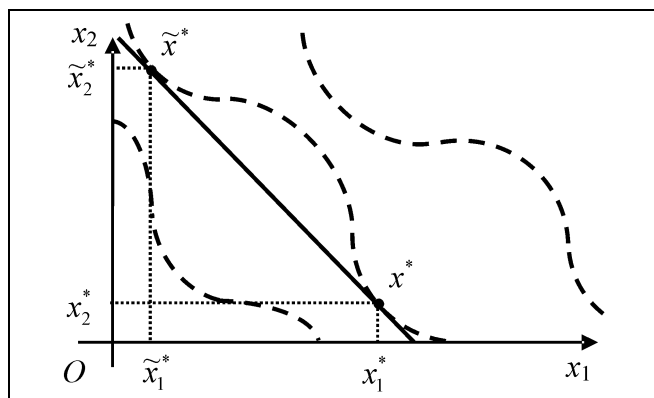


Рис. 12. Оптимальный выбор для производственной функции Филиппова

факторов, что означает наличие нескольких оптимальных решений (точки x^* , \tilde{x}^*), однако для каждой изокванты число точек различно. Более того, аналитически определить их расположение, а следовательно, и функции спроса для общего случая затруднительно.

ЗАКЛЮЧЕНИЕ

Решение модифицированной с учетом цены конъюнктурной информации задачи выбора производителя в значительной мере зависит от вида производственной функции, подходящего для данной фирмы, а также от выбранного периода анализа — краткосрочного или долгосрочного. Для случаев ресурсов совершенных заменителей или дополнителей получаем, что конъюнктурная информация не оказывает принципиального влияния на выбор фирмы, по крайней мере, до момента изменения более сильных условий (для ресурсов-субститутов — цен обоих факторов, для ресурсов-дополнителей — цены всего набора факторов). Анализ функции Кобба—Дугласа также показал незначительное влияние, поскольку в случае большей информированности увеличивается объем использования только подешевевшего ресурса.

Принципиальное влияние конъюнктурной информации характерно для функций Стоуна и CES, где с повышением информированности производитель изменяет не только объем вовлечения данного ресурса, но и всю структуру ресурсного вектора, что может привести к необходимости смены ранее выбранной технологии производства.

Еще более интересные результаты демонстрирует логистическая производственная функция Филиппова, которая допускает не только существование нескольких оптимальных точек, но и разрывов в спросе на ресурс, что может быть вызвано, например, необходимостью накопления ценности данного ресурса для производственного процесса.

ЛИТЕРАТУРА

1. *Lesourne J.A.* A Theory of Individual for Economic Analysis. — Amsterdam, 1977. — P. 7—13.
2. *Stiglitz J.E.* Equilibrium in product markets with imperfect information // *American Economic Review*. — 1979. — Vol. 69, — N 2. — P. 339—345.
3. *Маленко Э.* Лекции по микроэкономическому анализу. — М.: Наука, 1984. — 390 с.
4. *Akerlof G.A.* The Market for «Lemons»: Qualitative Uncertainty and Market Mechanism // *Quarterly Journal of Economics*. — 1970. — August. — P. 488—500.
5. *Вальрас Л.* Элементы чистой политической экономии. — М.: Изограф, 2000. — С. 18.
6. *Терехов Л.Л.* Производственные функции. — М.: Статистика, 1974. — С. 10—18.
7. *Dreze J.* Some theory of labor management and participation // *Econometrica*. — 1976. — N 6. — Vol. 44. — P. 1125—1139.
8. *Производственные функции в управлении проектами* // Научные и учебно-методические разработки Ин-та инноватики СПбГУ. — [Электронный ресурс]. (http://www.iit.spb.ru/material/methodical_m/m_4/Production_functions_management_projects.pdf). — С. 1—3.
9. *Клейнер Г.Б.* Производственные функции. Теория, методы, применение. — М.: Финансы и статистика, 1986. — С. 42.
10. *Иванов Е.Ю., Филиппов Л.А.* Информация в экономике и бизнесе. — Барнаул: Изд-во АлтГУ, 2000. — С. 26—29.

☎ (3852) 24-65-58; e-mail: baykin@rol.ru, ieu@asu.ru, isaeva@econ.asu.ru

Статья представлена к публикации членом редколлегии Р.М. Нижегородцевым. □



ПОЗДРАВЛЯЕМ

главного редактора нашего журнала

Дмитрия Александровича **НОВИКОВА**

с избранием членом-корреспондентом Российской академии наук.

Желаем ему доброго здоровья и дальнейших творческих успехов!

Редколлегия и редакция
журнала «Проблемы управления»

МЕТОД ЭВОЛЮЦИИ ПАРАМЕТРОВ В НЕЛИНЕЙНЫХ СЕТЕВЫХ ЗАДАЧАХ НА ПРИМЕРЕ ОПТИМИЗАЦИИ РАСПРЕДЕЛЕНИЯ ВХОДЯЩИХ ПОТОКОВ НАГРУЗКИ

В.А. Жевнеров

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Метод эволюции параметров предложено применять для решения нелинейных задач оптимизации сетей. Особенности применения метода показаны на примере оптимального распределения входящих потоков нагрузки в сети передачи данных. Показано, что метод обеспечивает существенное сокращение времени решения задач по сравнению с известными алгоритмами при сравнимых затратах на реализацию.

ВВЕДЕНИЕ

Оптимальное распределение входящих потоков нагрузки представляет собой один из важных этапов проектирования сетевых систем различного назначения, например, таких как производственные предприятия, системы передачи данных и др. Особенность такой задачи заключается в большом числе переменных, описывающих возможные маршруты движения потоков, и ограничениях на допустимые значения интенсивностей потоков. Известные методы оптимизации [1–3] обеспечивают возможность численного решения задачи при числе узлов системы до нескольких десятков. Однако для многих крупных производственных объединений, а тем более систем связи, общее число узлов (филиалов или узлов коммутации) выше на порядок. Это определяет актуальность разработки более эффективных численных методов решения практических задач высокой размерности. Такие методы предлагается строить на основе метода эволюции параметров (ЭП) [4].

Предлагаемый метод ЭП может быть распространен для решения аналогичных классов сетевых задач, например, для транспортных задач [5], в которых, кроме нелинейной функции затрат, в зависимости от объема перевозок требуется определить транзитные маршруты и распределение нагрузки между ними. Задача о максимальном потоке [6] может ставиться в виде требования определения максимальной интенсивности потока, обеспечивающего заданное значение целевой функции.

Основные особенности применения метода ЭП достаточно хорошо демонстрируются на примере решения задачи оптимального распределения входящих потоков нагрузки для систем передачи данных. Суть задачи состоит в оптимальном распределении интенсивностей входящих потоков сообщений на множестве допустимых маршрутов передачи по критерию минимизации среднего времени доставки. В отличие от рассмотренной в работе [4] аналогичной задачи, здесь учитываются дополнительные нелинейные ограничения на максимально допустимое значение среднего времени передачи сообщений по отдельным ветвям сети.

1. МЕТОД ЭВОЛЮЦИИ ПАРАМЕТРОВ

Существо метода поясняется для задачи общего вида

$$\max_{\bar{x} \in \Omega_x} W(\bar{x}), \quad (1)$$

где Ω_x — область допустимых значений параметров $\bar{x} = \{x_i\}$, $i = \overline{1, n}$, определяемая системой M ограничений:

$$\varphi_j(\bar{x}) \leq c_j^0, \quad j = \overline{1, M}. \quad (2)$$

Для простоты изложения рассматривается одноэкстремальная задача вогнутого программирования: функции $\varphi_j(\bar{x})$ полагаются гладкими и выпуклыми, а функция $W(\bar{x})$ — гладкой и вогнутой, $W(\bar{x}) \geq 0$. Образованная системой ограничений (2) область Ω_x полагается односвязной.

Традиционный подход к численному решению задачи (1), (2) состоит во введении штрафных функций [1, 7–9] различного вида, учитывающих влияние ограничений. Наиболее распространено применение штрафных функций в соответствии с методом неопределенных множителей Лагранжа. Однако применение этого метода для исследования задач оптимизации структур сетевых систем весьма проблематично из-за большого числа ограничений и переменных. Например, частная задача распределения нагрузки по маршрутам полносвязной сети, состоящей из N узлов, требует $\sim N!$ переменных и ограничений по числу возможных маршрутов. Применимый для решения подобных задач численный метод, очевидно, должен обеспечивать возможность значительного снижения числа одновременно анализируемых переменных и ограничений при уточнении решения на каждой итерации.

Основная идея метода эволюции параметров состоит в отслеживании оптимального решения при изменяющихся условиях задачи. Изменения оптимального решения определяются математическим оператором, составленным на основе требования соблюдения необходимых и достаточных условий оптимальности. Начальный поиск экстремума производится для измененных условий задачи, при которых решение находится достаточно просто. Например, нагрузки принимаются близкими к нулевым, общие затраты на систему полагаются практически неограниченными и т. п. Затем условия задачи восстанавливаются (эволюционируют) до начального уровня при одновременном отслеживании изменений оптимального решения.

Применение метода ЭП для решения задачи распределения однородных потоков нагрузки с нелинейной целевой функцией $W(\bar{x})$ и взаимно независимыми линейными ограничениями показано в работе [4]. Далее рассматриваются особенности применения метода ЭП для решения задачи (1), (2) более общего вида.

2. ОПЕРАТОР ЭВОЛЮЦИИ ДЛЯ ОДНОЭКСТРЕМАЛЬНОЙ ЗАДАЧИ

Для задачи (1) в качестве параметров эволюции наиболее естественно выбрать значения правых частей ограничений (2), изменяющихся от начальных значений $c_j = c_j^H$ до c_j^0 с дискретным шагом Δc_j в общем случае переменной длины.

Оператор эволюции, определяющий взаимосвязь между решениями на каждом шаге эволюции, строится на основе соотношений, которым должно удовлетворять оптимальное решение. Такие уравнения, приводимые ниже, по виду совпа-

дают с необходимыми условиями оптимальности, сформулированными в теореме Куна—Таккера [2].

Пусть \bar{x}^0 — решение задачи (1). Тогда для достаточно малой вариации $\delta\bar{x}$ должно выполняться

$$\delta W(\bar{x}) = \sum_i \frac{\partial W(\bar{x}^0)}{\partial x_i} \delta x_i \leq 0 \quad (3)$$

при соблюдении условий

$$\sum_i \frac{\partial \varphi_j(\bar{x}^0)}{\partial x_i} \delta x_i = \delta c_j \leq 0, \quad j = \overline{1, M}. \quad (4)$$

Условия (4) составляются только для активных ограничений, т. е. для которых $\varphi_j^a(\bar{x}^0) = \varphi_j(\bar{x}^0) = 0$. Для неактивных ограничений знак δc_j не имеет значения. Рассмотрим случай, когда $M^a(\bar{x}^0)$ — число активных ограничений меньше n , т. е. $M^a(\bar{x}^0) < n$. Выражение (4), определяющее допустимые значения $\delta\bar{x}$ при $\delta c_j = 0$, принимает вид

$$\sum_i \frac{\partial \varphi_j^a(\bar{x}^0)}{\partial x_i} \delta x_i = 0, \quad j = \overline{1, M^a}, \quad (5)$$

и при этом должно выполняться условие (3). Так как в соответствии с выражением (5) для любой допустимой вариации $\delta\bar{x}$ вариация $-\delta\bar{x}$ также допустима, из формулы (3) следует

$$\sum_i \frac{\partial W(\bar{x}^0)}{\partial x_i} \delta x_i = 0. \quad (6)$$

Очевидно, соотношение (6) выполняется тогда и только тогда, когда оно является линейно зависимым относительно левых частей выражения (5). Иначе говоря, существуют коэффициенты μ_j , для которых

$$\frac{\partial W(\bar{x}^0)}{\partial x_i} \delta x_i = \sum_j \mu_j \frac{\partial \varphi_j^a(\bar{x}^0)}{\partial x_i} \delta x_j, \quad i = \overline{1, n}. \quad (7)$$

Пусть $\delta\bar{c} = \{\delta c_j\}$ — вариация значений c_j активных ограничений. Тогда вариация δW с учетом выражений (4) и (7) принимает следующий вид:

$$\delta W = \sum_i \sum_j \mu_j \frac{\partial \varphi_j^a(\bar{x}^0)}{\partial x_i} \delta x_i = \sum_j \mu_j \delta c_j. \quad (8)$$

Выражение (8) означает, что для обеспечения $\max W$ в точке \bar{x}^0 обязательно должно выполняться условие $\mu_j \geq 0, \forall j$, так как в противном случае для $\mu_j < 0$ допустимая вариация $\delta c_j < 0$ приведет в соответствии с соотношением (6) к увеличению значения $W(\bar{x}^0)$. Появление отрицательных значений



$\mu_j(\bar{x})$ в процессе эволюции означает необходимость перевода ограничений φ_j^a в разряд неактивных. При заданных значениях $\delta \bar{c}$ величины $\delta \bar{x} = \{\delta x_j\}$ и μ_j определяются из совместного решения систем уравнений (4) и (7). Случай $M^a(\bar{x}^0) < n$ означает, что в точке \bar{x}^0 ($n - M^a(\bar{x}^0)$) значений $\frac{\partial W(\bar{x}^0)}{\partial x_i}$ тождественно равны нулю, что является необходимым условием существования безусловного экстремума по x_i и приводит, соответственно, к $\delta x_i = 0$. В этом случае размерность систем уравнений (4) и (7) фактически понижается на $(n - M^a(\bar{x}^0))$.

Сказанное очевидным образом распространяется для случая $n = M^a(\bar{x}^0)$. При $n < M^a(\bar{x}^0)$ система ограничений либо вырожденная, либо противоречивая, не допускающая решения.

Эволюция значений c_j удобна тем, что для этого случая вариации δc_j для активных ограничений взаимно независимы, в то время как изменения $\delta \bar{x} = \{\delta x_j\}$ должны удовлетворять определенным взаимосвязям вида (4). Представление (8), кроме удобства реализации процесса эволюции значений c_j , дополнительно обеспечивает возможность проведения оценки на чувствительность решения к налагаемым ограничениям, что бывает часто необходимо при исследовании практических задач.

Активные ограничения на каждой итерации выявляются путем проверки нарушения строгого неравенства $\varphi_j(\bar{x}) < c_j^0$, а из списка активных ограничений они исключаются по отрицательному знаку μ_j .

Для задач распределения нагрузки в сетях наибольшее число ограничений имеет вид

$$\pm x_{i^*} \leq c_{i^*}^0, \quad (9)$$

определяющий границы допустимых значений x_{i^*} . Например, это требования положительных значений нагрузок системы, объемов вкладываемых ресурсов и т. п. Фактически все активные ограничения вида (9) должны быть включены в системы уравнений (4) и (7), но от этого весьма желательно избавиться, иначе размерность этих систем уравнений будет очень велика. Формально это достигается следующим образом;

— для активного ограничения (9) соответствующее уравнение системы (4) принимает вид $\delta x_{i^*} = 0$;

— с учетом, что $\delta x_{i^*} = 0$, значение x_{i^*} полагается равным константе на текущей итерации и величина δx_{i^*} исключается из числа неизвестных в системах (4) и (7);

— уравнение вида
$$\frac{\partial W(\bar{x}^0)}{\partial x_{i^*}} = \sum_{j \neq i^*} \mu_j \frac{\partial \varphi_j^a(\bar{x}^0)}{\partial x_{i^*}} + \mu_{i^*}$$

также исключается из системы (7) и служит только для определения знака μ_{i^*} (принадлежности ограничения к активному) на каждой итерации. Условие $\mu_{i^*} \geq 0$ активности ограничения в этом случае записывается в виде

$$\sum_{j \neq i^*} \left(\mu_j \frac{\partial \varphi_j^a(\bar{x}^0)}{\partial x_{i^*}} \right) - \frac{\partial W}{\partial x_{i^*}} \leq 0. \quad (10)$$

Проверка условия (10) для переменных x_{i^*} производится специализированными методами. Достаточно эффективны здесь методы динамического программирования [5, 6].

Таким образом, в соответствии с выражениями (4) и (7) решение задачи (1), (2) в процессе эволюции значений $\bar{c} = \{c_j\}$ должно удовлетворять следующей системе уравнений, составляемой для активных ограничений:

$$\begin{cases} \frac{\partial W}{\partial x_i} - \sum_j \left(\mu_j \frac{\partial \varphi_j^a(\bar{x}^0)}{\partial x_i} \right) = 0, & i = \overline{1, n}, \\ \varphi_j^a(\bar{x}^0) - c_j = 0, & j = \overline{1, M^a}. \end{cases} \quad (11)$$

Обозначим систему (11) для удобства оператором $F(\bar{x}, \bar{c}) = 0$. При изменении значений \bar{c} с дискретным шагом $\Delta \bar{c} = \{\Delta c_j\}$ соответствующее решение задачи в точке $\bar{x} + \Delta \bar{x}$ должно также удовлетворять условию

$$F(\bar{x} + \Delta \bar{x}, \bar{c} + \Delta \bar{c}) = 0. \quad (12)$$

Точное решение уравнений (12) обеспечивает определение изменения $\Delta \bar{x}$ оптимального решения задач. Однако даже для достаточно простого вида функции $W(\bar{x})$, например, в виде среднего времени или вероятности передачи потоков, уравнения (12) точно решить весьма затруднительно. Поэтому предлагается линеаризация уравнений (12) по малым параметрам Δx_i и $\Delta \mu_j$ — изменениям x_i и μ_j . Очевидно, что в этом случае значения $\partial W / \partial x_i$ и μ_j будут определяться с точностью $O((\Delta c)^4)$, как показано, например, в работе [4]. Ошибка вычисления $W(\bar{x})$ не будет накапливаться в процессе эволюции при достаточно малых $\Delta \bar{c}$, т. е. сохранится свойство релаксации ошибок, отмеченное ранее в работах [4, 10].

В процессе эволюции, в силу неточности выполнения условия (12) из-за линеаризации, получаемое значение $\varphi_j(\bar{x} + \Delta \bar{x}) = c_j(\bar{x} + \Delta \bar{x})$ может изменяться не точно на задаваемую величину Δc_j .

При эволюции значения c_j , находящегося достаточно далеко от конечного значения c_j^0 , это может не иметь принципиального значения. Когда требуется обеспечить значение c_j с достаточной точностью, например, для активного не эволюционирующего ограничения, необходимо вводить в линеаризованную систему реальную невязку δc_j значения c_j . Тогда ошибка определения значения c_j будет иметь также порядок $O((\Delta c)^2)$, а в процессе коррекции решения для $\Delta \bar{c} = 0$ будет убывать квадратично. Верхние оценки ошибки вычисления W и \bar{c} получаются достаточно просто через коэффициенты линеаризованной системы (12) по аналогичным приведенным в работе [4] соотношениям.

Линеаризованная система (12) с учетом ограничений (11) принимает следующий вид:

$$\left\{ \begin{array}{l} \sum_k \left[\frac{\partial^2 W(\bar{x})}{\partial x_i \partial x_k} - \sum_j \left(\mu(\bar{x}) \frac{\partial^2 \varphi_j^a(\bar{x})}{\partial x_i \partial x_k} \right) \right] \Delta x_k - \\ - \sum_j \left(\Delta \mu_j \frac{\partial \varphi_j^a(\bar{x})}{\partial x_i} \right) = \delta_i, \quad i = \overline{1, n}, \\ \sum_i \frac{\partial \varphi_j^a(\bar{x})}{\partial x_i} \Delta x_i = c_j + \delta c_j, \quad j = \overline{1, M^a}, \end{array} \right. \quad (13)$$

где $\delta_i = \sum_j \left(\mu_j \frac{\partial \varphi_j^a}{\partial x_i} \right) - \frac{\partial W}{\partial x_i}$ — невязка соблюдения необходимых условий оптимальности; $\delta c_j = c_j - \varphi_j^a(\bar{x}^0)$ — ошибка обеспечения требуемого значения c_j на предыдущей итерации. Величины δ_i всегда должны входить в систему (13) как обеспечивающие релаксацию ошибок в процессе эволюции, а δc_j — при необходимости достаточно точно достижения требуемых значений c_j .

Если система (13) оказывается вырожденной, тогда для нахождения решения необходимо прибегать к разложениям по $\Delta \bar{x}$ более высокого порядка. При этом возможны бифуркация решения по параметру эволюции или вообще исчезновение решения в классе действительных чисел [11]. При неудачной параметризации задачи появление точек бифуркации может носить искусственный характер. Поэтому при построении алгоритмов решения конкретных классов задач необходимо исследовать возможность их появления в процессе эволюции. В то же время, наличие бифуркаций может быть определено многоэкстремальностью задачи. В этом случае построение оператора эволюции в точках бифуркации составляет самостоятельную проблему, общего решения которой не существует, так как в нелинейных задачах может

существовать бесконечное число бифуркаций разных видов. Это определяет необходимость изучения особенностей поведения решений конкретных задач и построения на этой основе конкретного оператора эволюции с бифуркациями.

Итак, алгоритм решения задачи (1), (2) в общем виде будет состоять из следующих основных этапов:

- 1) определение параметров эволюции и начальной задачи оптимизации;
- 2) задание шага эволюции;
- 3) определение множества активных ограничений;
- 4) составление и решение уравнений эволюции;
- 5) оценка точности полученного решения;
- 6) контроль окончания процесса эволюции (очередная итерация начинается со второго этапа).

Далее перечисляются основные особенности каждого из этапов.

В качестве параметров эволюции удобнее выбирать значения c_j для ограничений (2) задачи (1). В задачах распределения потоков в сетевых системах часто удобно искать начальное решение для нулевой нагрузки. Для такой начальной задачи определяются не только наилучшие характеристики процесса функционирования, но и проводится дополнительная проверка совместимости остальных ограничений. В ряде случаев удобно вводить фиктивные дополнительные ограничения, при которых начальное решение находится достаточно просто и значения правых частей которых изменяются до $+\infty$, что формально означает устранение дополнительных ограничений из исследуемой задачи.

Шаг эволюции $\Delta \bar{c} = \{\Delta c_j\}$ выбирается по оценкам точности линеаризации уравнений (12), откуда следует, что точность решения $\delta W \sim O(|\Delta \bar{c}|^4)$. Если требуется обеспечить заданную точность решения только для конечного значения $\bar{c} = \{c_j^0\}$, шаг эволюции выгодно взять большим, а дополнительную ошибку в конечной точке эволюции устранить на дополнительных итерациях при нулевом шаге эволюции (коррекция решения). Такой подход в большинстве случаев позволяет значительно уменьшить общее число итераций. В этом случае максимальный шаг $\Delta \bar{c}$ оценивается из условия сохранения свойства релаксации ошибок линеаризованным оператором $F(\bar{x}, \bar{c})$. Для задачи оптимизации распределения входящих потоков нагрузки (см. п. 3) свойство релаксации ошибок сохраняется при относительной ошибке определения W не более 10—15 %, а эволюция одновременно всех изменяемых параметров c_j требует меньшего числа итераций. Поэтому рекомендуется вводить единый параметр эволюции α , изменяемый от 0 до 1, представляя эволюцию параметров c_j , например, в виде $c_j = c_j^H + \alpha(c_j^0 - c_j^H)$.



Множество активных ограничений формируется на каждой итерации по следующим принципам:

- ограничение $\varphi_j^a(\bar{x})$ исключается из числа активных при $\mu_j(\bar{x}) < 0$;
- ограничение заносится в список активных при нарушении условия $\varphi_j(\bar{x}) < c_j^0$ на очередной итерации.

Для определения значений $\mu_j(\bar{x}) \geq 0$ дополнительно привлекаются специализированные методы поиска. Например, если производная функции $W(\bar{x})$ аддитивная, для поиска ограничений в соответствии с условием (10) рекомендуются методы динамического программирования. Эффективность применения метода ЭП в целом в большинстве случаев определяется эффективностью алгоритмов формирования множества активных ограничений.

Линеаризованные уравнения эволюции составляются в соответствии с выражениями (13), а для их решения применяется один из известных методов, выбираемый с учетом конкретных особенностей рассматриваемой задачи. Для таких методов в наихудшем случае объем вычислений будет пропорционален кубу числа уравнений [1].

Точность полученного решения оценивается так же, как и при выборе шага $\Delta\bar{c}$. Если ошибка неприемлема, то принимается решение о проведении итерации заново при меньшем шаге эволюции или коррекции полученного решения для $\Delta\bar{c} = 0$.

Процесс решения заканчивается при достижении всеми c_j значений c_j^0 , когда эволюционируемые условия задачи совпадут с исходными и будет достигнута требуемая точность решения. Если ограничения окажутся несовместимыми, в процессе эволюции возникнет ситуация невозможности формирования множества активных ограничений, удовлетворяющих условию $\mu_j(\bar{x}) \geq 0$.

2. ПРИМЕР ЗАДАЧИ ОПТИМИЗАЦИИ РАСПРЕДЕЛЕНИЯ ВХОДЯЩИХ ПОТОКОВ

В качестве примера рассматривается наиболее часто встречающаяся на практике задача оптимального распределения потоков сообщений в пуассоновской сети передачи данных по критерию минимума W — среднего времени передачи. Решение такой задачи подробно описано в работе [4] для случая взаимно независимых линейных ограничений, отражающих требование передачи всех потоков заданной интенсивности $\bar{\lambda}^0$, распределяемых одновременно по различным маршрутам. В качестве параметров эволюции выбирались значения $\bar{\lambda} = \alpha^* \bar{\lambda}^0$, изменяемые от $\bar{\lambda} = 0$ до $\bar{\lambda} = \bar{\lambda}^0$. Указанная задача обобщена введением дополни-

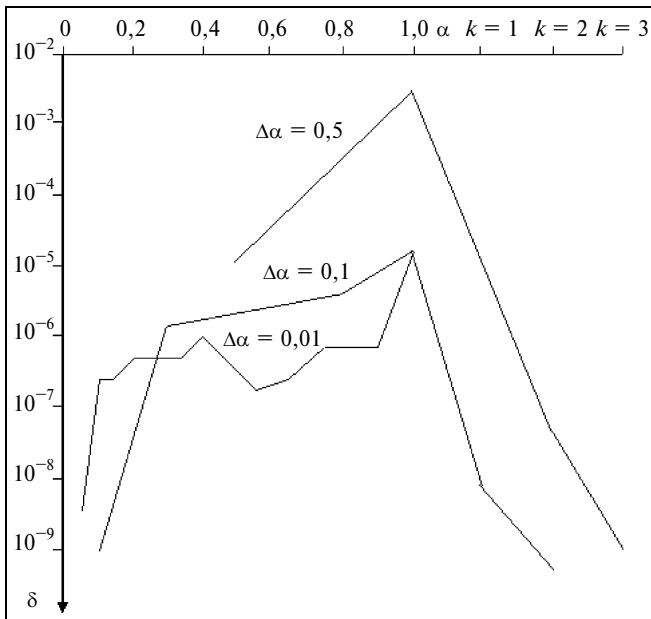
тельных ограничений на максимально допустимые значения среднего времени передачи сообщений по каждой ветви сети. Поскольку для пуассоновской сети функция W является дробно-линейной по параметрам $\bar{\lambda} = 0$ [3], такие дополнительные ограничения носят нелинейный характер. На каждый узел сети назначалось по одному входящему потоку для передачи на другой узел, назначаемому случайным образом.

Зависимости относительной ошибки $\delta = |\Delta W|/W$ от единого параметра эволюции $0 \leq \alpha \leq 1$ при различных шагах разбиения $\Delta\alpha$ и при проведении коррекции решения при $\Delta\alpha = 0$ практически не отличается от приведенных в работе [4]. Здесь также соблюдается зависимость $\delta \sim (\Delta\alpha)^4$. Рост трудоемкости вычислений в основном определяется увеличением размерности линеаризованных уравнений эволюции из-за появления дополнительных активных ограничений. В решаемой задаче для полностью связанной сети из 100 узлов, на каждый из которых поступает по одному внешнему потоку для передачи на определенный узел, введение дополнительных нелинейных 100 ограничений на максимально допустимое значение среднего времени передачи отдельных потоков по сети приводит к увеличению требуемого объема оперативной памяти примерно на 70 %, а времени решения — на 20—30 %. Такая постановка задачи имеет место при решении актуальной практической задачи управления приоритетными потоками с заданными нормативами на среднее время передачи данных. Меньший рост увеличения времени решения объясняется тем, что основные затраты времени приходится на этап поиска оптимальных маршрутов, на котором влияние дополнительных ограничений незначительно. Полное время решения на ЭВМ «Pentium-IV» составляет около 2 мин. Среднее число маршрутов передачи ~ 3 при средней длине маршрутов 3—4 ветви.

Экспериментально установлено, что наименьшее число итераций, требуемых для получения окончательного решения, достигается при организации процесса одновременной эволюции значений интенсивности входящих потоков нагрузки в виде $\Delta\bar{c} = \{\Delta c_j\}$, $\Delta c_j = \Delta\alpha c_j^0$, где α изменяется с шагом $\Delta\alpha$ от 0 до 1. В этом случае эволюция производится фактически по одному параметру.

Типичные зависимости относительной ошибки δ при росте параметра эволюции α от 0 до 1 для различных шагов $\Delta\alpha$ показаны на рисунке.

Как видно, значение δ пропорционально не 4-й, а 2-й степени $\Delta\alpha$. Это объясняется неточностью выполнения неравенства $\varphi(\bar{x} + \Delta\bar{x}) = c - \Delta c$ на каждой итерации. Здесь же приведены результаты для $k = \overline{1, 3}$ циклов коррекции окончательного решения при $\alpha = 1$ и $\Delta\alpha = 0$. Здесь ошибка



Зависимость относительной ошибки δ при росте параметра эволюции α

убывает наиболее быстро — пропорционально квадрату от начального значения.

Сравнение реализованного алгоритма с другим алгоритмом, реализующим метод девиации потоков [3], показывает, что затраты вычислительных ресурсов на выполнение одной итерации для обоих алгоритмов будут примерно одинаковыми. Типичные результаты сравнения алгоритмов по числу итераций при требуемой относительной точности 10^{-1} приведены в таблице, где показаны результаты численного решения двух примеров, взятых из работы [3]. В первом примере сеть имеет четыре узла, во втором — девять узлов.

В обоих примерах сети полагаются пуассоновскими, и в качестве критерия эффективности выбирается среднее время доставки сообщений.

Как следует из таблицы, алгоритмы, основанные на методе ЭП, не менее чем на порядок более эффективные. Кроме того, к существенному достоинству метода ЭП относится очевидная возможность учета нелинейных ограничений произвольного вида, определяющих, например, требования к качеству передачи отдельных потоков.

Результаты сравнения алгоритмов

Пример	Число итераций	
	Метод девиации потоков	Метод эволюции параметров
1	300	12
2	110	8

Фактически это приводит к увеличению размерности оператора эволюции на число новых активных ограничений. В то же время, учет нелинейных ограничений в методе девиации потоков напрямую невозможен и требует существенной доработки алгоритмов решения задачи.

ЗАКЛЮЧЕНИЕ

Применение метода эволюции параметров обеспечивает возможность составления эффективного алгоритма решения нелинейных задач оптимизации распределения потоков нагрузки в больших системах.

Предложенный вид линеаризации оператора эволюции позволяет производить текущую оценку точности определения оптимального решения, что приводит к существенному сокращению общего объема вычислений благодаря выбору максимально допустимого шага эволюции и обеспечивает возможность коррекции точности получаемых решений.

По сравнению с известными алгоритмами применение метода эволюции для нелинейных задач оптимальной маршрутизации потоков позволяет сократить общее время вычислений на 1—2 порядка при сравнимых затратах на реализацию.

ЛИТЕРАТУРА

1. Банди Б. Методы оптимизации. Вводный курс. — М.: Радио и связь, 1988.
2. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. — М.: Мир, 1985.
3. Клейнрок Л. Вычислительные системы с очередями. — М.: Мир, 1979.
4. Жевнеров В., Родионов И. Оптимальное проектирование информационных систем методом эволюции параметра // Журнал выч. матем. и матем. физики. — 1986. — Т. 26, № 3. — С. 64—67.
5. Филлипс Д., Гарсиа-Диас А. Методы анализа сетей. — М.: Мир, 1984.
6. Форд Л., Фалкерсон Д. Потоки в сетях. — М.: Мир, 1966.
7. Данилин Ю. Метод градиентного типа для минимизации гладких штрафных функций // Кибернетика — 1988. — № 4. — С. 25—29.
8. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных уравнений со многими неизвестными. — М.: Мир, 1975.
9. Емец О.А., Барболина Т.Н., Черненко О.А. Решение задач оптимизации с дробно-линейными целевыми функциями и дополнительными ограничениями на размещениях // Кибернетика и системный анализ — 2006. — Т. 42, № 5. — С. 42—47.
10. Кузьмина Л., Родионов И. Решение систем уравнений Хартри методом эволюции параметра // Там же. — 1985. — 25, № 8. — С. 57—82.
11. Литвинцева С., Родионов И. Бифуркация спектра несамосопряженных краевых спектральных задач / Препринт ИПМ АН СССР. — М., 1988. — № 40.

e-mail: jewn@mail.ru

Статья представлена к публикации членом редколлегии С.А. Редкозубовым. □

ИДЕНТИФИКАЦИЯ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ПРОИЗВОДСТВА ГУБЧАТОГО ТИТАНА

Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер, С.Л. Краев

Березниковский филиал Пермского государственного технического университета

Рассмотрены особенности построения моделей процессов восстановления и вакуумной сепарации губчатого титана. Предложено описывать динамику процессов дифференциальными уравнениями с переменными коэффициентами. Изложены способы идентификации параметров моделей в позиционных системах управления.

ВВЕДЕНИЕ

Основу промышленного производства губчатого титана составляют процессы восстановления тетраоксида титана магнием и последующей вакуумной сепарации, осуществляемые в аппаратах периодического действия, конструкцией которых предусмотрено разделение рабочего объема каждого аппарата на ряд зон нагрева (аппарат сепарации), зон нагрева и зону экзотермической реакции (аппарат восстановления). Характерная особенность таких систем как объектов управления состоит в необходимости поддержания позиционными регуляторами в каждой зоне нагрева (зоне реакции) определенной температуры с требуемой точностью. Температура зон нагрева аппаратов восстановления и сепарации регулируется притоком электрической энергии, поступающей от нагревателей, а температура зоны экзотермической реакции регулируется оттоком тепловой энергии посредством создаваемого вентилятором воздушного потока [1].

Системы позиционного регулирования температуры процессов восстановления и сепарации работают в режиме автоколебаний. Параметры автоколебаний существенно зависят от условий функционирования восстановления и сепарации и изменяются при вариациях нагрузок производства. Вопросы повышения качества позиционного управления температурным режимом данных процессов исследованы недостаточно, что существенно снижает технико-экономические показатели производства. Поэтому совершенствование систем управления температурным режимом процессов представляет собой одну из актуальных задач производства губчатого титана [2].

Такая задача решается в два этапа [3]:

- разработка математических моделей динамики процессов восстановления и вакуумной сепарации;
- синтез на основе моделей динамики оптимальных систем управления.

В настоящей статье рассматриваются вопросы построения моделей динамики технологических процессов (ТП) производства губчатого титана.

1. ВЫБОР МАТЕМАТИЧЕСКОЙ МОДЕЛИ И ОСОБЕННОСТИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ И СИСТЕМ УПРАВЛЕНИЯ ПРОИЗВОДСТВОМ ГУБЧАТОГО ТИТАНА

В общем случае построение математической модели для синтеза управления ТП предполагает решение следующих основных вопросов [4]:

- нахождение компромисса между точной передачей свойств ТП и относительной простотой модели;
- построение многомерных математических моделей при большом числе переменных состояния ТП;
- изучение возможности замены системы уравнений в частных производных, описывающих ТП с распределенными параметрами, системой обыкновенных дифференциальных уравнений;
- описание поведения ТП в условиях неопределенности и нестационарности, обусловленных неполной текущей и априорной информацией о возмущениях, изменением ТП и параметров управления;
- повышение точности описания ТП путем учета нелинейности уравнений модели.

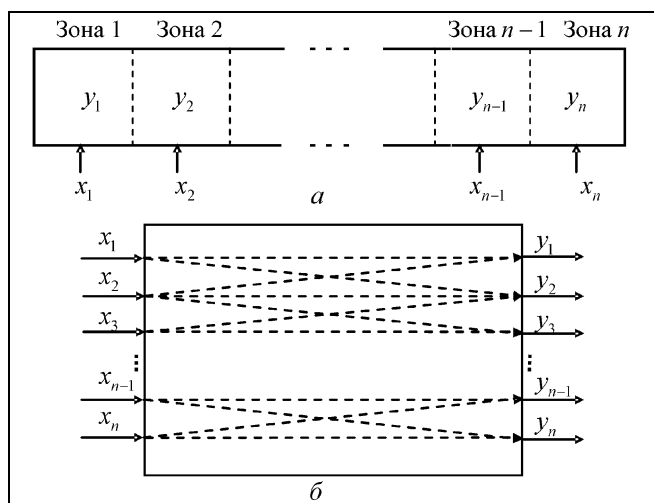


Рис. 1. Упрощенная схема n -зонного аппарата сепарации титана (а) и схема связей между входными и выходными переменными (б)

Установлено, что разработка математической модели занимает от 80 до 90 % усилий, необходимых для построения систем управления.

Рассмотрим в рамках сформулированных вопросов особенности производства губчатого титана, которые необходимо учитывать при математическом описании динамических свойств ТП [5–8].

Изменчивость параметров аппаратов восстановления и сепарации. В процессе эксплуатации реакторы удлиняются, уменьшается толщина их стенок, изнашивается футеровка печей, в условиях агрессивных сред снижается мощность нагревателей.

Нестационарность и нелинейность процессов как объектов управления. На объекты управления действуют неконтролируемые переменные возмущения, вызванные изменением теплофизических характеристик реакционной массы, ее структуры, состава и конфигурации, изменением выделяемого тепла экзотермической реакции восстановления титана при вакуумной очистке губчатого титана от примесей магния и хлорида магния. Статические и динамические характеристики зон нагрева и зоны реакции различны на разных стадиях процессов, характер их изменения не подчиняется какому-либо определенному закону.

Многомерность объектов управления, обусловленная многозонной системой нагрева аппарата сепарации и многозонной системой нагрева и охлаждения аппарата восстановления со значительным числом взаимосвязанных распределенных по временным и пространственным координатам входных и выходных величин. Упрощенную схему n -зонного аппарата сепарации можно представить в виде, показанном на рис. 1 [9, 10]. Здесь $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ — векторы

управления и выходных переменных соответственно, x_i — электрическая мощность nichромовых нагревателей, y_i — измеряемая температура в центральной части i -й зоны нагрева. Каждый вход влияет на несколько выходов и, соответственно, выход зависит от нескольких входов. Аналогично может быть представлена схема многозонного аппарата восстановления.

В динамике, в реальных условиях эксплуатации, *отсутствует априорная информация* о форме и степени связи выходных величин аппаратов с соответствующими входными величинами, которая одновременно учитывала бы влияние на выходные величины тепловых режимов соседних зон и неконтролируемых возмущений.

Нелинейность систем управления. При существующем аппаратном оформлении процессов восстановления и сепарации исполнительными механизмами служат релейные элементы (контакты, пускатели, клапаны, дозаторы и др.) с двумя рабочими положениями (включено/выключено). Это обстоятельство предопределило позиционный принцип управления ТП. Автоматизация осуществлена на базе нелинейных систем управления, в которых заданный температурный режим процессов поддерживают многоканальные двухпозиционные регуляторы. Низкое быстродействие регуляторов вносит дополнительное запаздывание, существенно снижающее качество регулирования температуры. Рабочие режимы регулирования температуры в нелинейных системах — сложные автоколебательные процессы, характеризующиеся асимметрией автоколебаний и остаточной неравномерностью.

В целом можно считать, что процессы восстановления и вакуумной сепарации представляют собой многомерные многосвязные нестационарные нелинейные объекты, а системы управления этими объектами — многомерные многосвязные нестационарные нелинейные системы управления.

На рис. 2 в качестве примера приведена структурная схема управления промышленным аппаратом сепарации с трехзонной системой нагрева.

Температура верхней $y_{в.3}(t)$, средней $y_{с.3}(t)$, нижней $y_{н.3}(t)$ зон аппарата сепарации АС поддерживается на соответствующих заданных уровнях $y_{в.3}^3$,

$y_{с.3}^3$ и $y_{н.3}^3$ многоканальным двухпозиционным регулятором (МДР) путем включения и выключения регулирующих воздействий — мощности нагрева верхней $x_{в.3}(t)$, средней $x_{с.3}(t)$ и нижней $x_{н.3}(t)$ зон. На выходные переменные $y_{в.3}(t)$, $y_{с.3}(t)$ и $y_{н.3}(t)$ оказывают влияние регулирующие воздействия соседних зон и неконтролируемые возмущения $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$.

Изложенные особенности производства губчатого титана свидетельствуют о сложности процес-

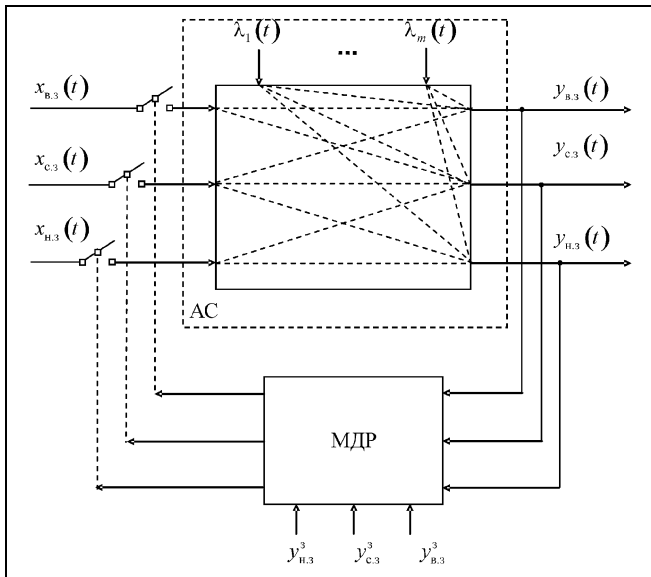


Рис. 2. Схема управления температурой аппарата сепарации титана

сов восстановления и сепарации как объектов управления и в целом характерны для управления и оценки состояния систем в условиях неопределенности и нестационарности [4, 11]. Это существенно затрудняет построение адекватного математического описания динамики процессов. В ранее проведенных исследованиях [5, 7] путем экспериментального определения динамических характеристик не удалось получить математическое описание, которое можно было бы использовать для разработки систем управления процессами производства губчатого титана. Известные математические модели [12, 13] предназначены для совершенствования технологии и аппаратуры.

2. ДЕКОМПОЗИЦИЯ СИСТЕМЫ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ

Предположим, что система управления процессом восстановления состоит из несвязанных друг с другом (автономных) нестационарных нелинейных систем регулирования температуры зон нагрева и зоны реакции. Управление процессом вакуумной сепарации состоит из несвязанных друг с другом нестационарных нелинейных систем регулирования температуры зон нагрева аппарата сепарации. Такие допущения вполне оправданы, поскольку температурные поля аппаратов достаточно хорошо изучены, и на практике взаимное влияние тепловых режимов устраняется рациональным размещением в рабочем объеме аппаратов восстановления и сепарации термопар и нагревателей зон, выбором оптимальной конструкции системы воздушного охлаждения зоны экзотермической реакции [14–17].

Предположим также, что объекты управления характеризуются сосредоточенными параметрами, динамика изменений которых может быть описана обыкновенными дифференциальными уравнениями. С учетом принятых допущений систему управления процессами восстановления и сепарации можно рассматривать состоящей из ряда типовых нестационарных нелинейных (позиционных) систем регулирования технологических параметров [3].

Каждая из таких систем регулирования состоит из обобщенного нестационарного объекта управления (ОНОУ), на вход которого воздействует неконтролируемое возмущение $z(t)$, и МДР — нелинейного регулятора [18] (рис. 3).

Под ОНОУ подразумевают зону нагрева, зону реакции аппарата восстановления или зону нагрева аппарата сепарации, под $z(t)$ — изменение тепловых потерь зоны нагрева, изменение выделяемого тепла зоны экзотермической реакции аппарата восстановления, изменение потребляемого тепла зоны нагрева аппарата сепарации. Подключением и отключением входной величины $x(t)$ (мощности нагрева зоны или мощности охлаждения зоны реакции) выходная величина $y(t)$ (температура) ОНОУ поддерживается на заданном уровне y_3 в соответствии с алгоритмом:

$$\phi[\bar{y}(t)] = \begin{cases} x(t) & \text{при } \bar{y}(t) < y_3 - \Delta y_0 \text{ и} \\ y_3 - \Delta y_0 \leq \bar{y}(t) \leq y_3 + \Delta y_0, & y'(t) > 0; \\ 0 & \text{при } \bar{y}(t) < y_3 + \Delta y_0 \text{ и} \\ y_3 - \Delta y_0 \leq \bar{y}(t) \leq y_3 + \Delta y_0, & y'(t) < 0; \end{cases}$$

$$\bar{y}(t) = y(\gamma T) \text{ при } \gamma T \leq t < (\gamma + 1)T, (\gamma = 1, 2, \dots),$$

где $\bar{y}(t)$ — квантованная по времени выходная величина ОНОУ, $\phi[\bar{y}(t)]$ — выходная величина МДР, $2\Delta y_0$ — зона нечувствительности МДР, $y'(t)$ — скорость изменения выходной величины, T — интервал квантования по времени.

Таким образом, при подключении $x(t)$ на входе ОНОУ действует разность $x(t) - z(t)$, при отключе-

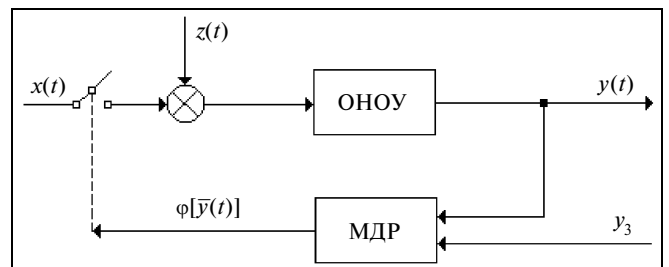


Рис. 3. Типовая структурная схема позиционного регулирования технологического параметра процесса восстановления (сепарации)

нии $x(t)$ поведение ОНОУ определяется возмущающим воздействием $z(t)$. Система регулирования при этом работает в режиме автоколебаний [8].

3. ПОСТАНОВКА ЗАДАЧИ ИДЕНТИФИКАЦИИ СТРУКТУРЫ МОДЕЛЕЙ

В общем случае для решения задач управления ТП необходимо математическое описание, устанавливающее связь между входными и выходными переменными. Назначение модели состоит в установлении таких количественных соотношений между входными и выходными переменными, на основании которых может быть выработано управление, обеспечивающее заданную цель функционирования объекта. Такая модель не вскрывает особенностей физических и химических процессов, протекающих в объекте управления. Однако получаемая с ее помощью связь образует совокупность тех сведений о динамических свойствах объекта, которая достаточна для разработки системы управления. Построение моделей динамики технологических объектов осуществляют экспериментальным методом. Результаты измерений служат основой построения математического описания динамики объекта. Для получения наиболее полной модели измерения проводят непосредственно в процессе управления с учетом действующих возмущений и реальных условий функционирования технологического объекта. Под идентификацией технологического объекта управления понимают определение его структуры и параметров по данным «вход-выход» [19]. В нашем случае задача идентификации заключается в определении структуры и параметров ОНОУ — уравнений связи входных и выходных величин и неизвестных коэффициентов этих уравнений [18].

Структуру модели динамики ОНОУ можно задать оператором A , устанавливающим соответствие между выходной $y(t)$ и входными функциями $x(t)$ и $z(t)$: $y(t) = A\{x(t), z(t)\}$. В теории управления ТП отсутствуют методы определения операторов объектов неизвестной структуры. Обычно, основываясь на реализациях входных и выходных величин, высказывают гипотезу о конкретном выражении операторов таких объектов, которую затем проверяют по результатам эксперимента [19].

Для оценки динамических свойств реального ОНОУ его можно с достаточной степенью точности представить эквивалентными нестационарными объектами без самовыравнивания и с самовыравниванием, описываемыми соответствующими дифференциальными уравнениями с переменными коэффициентами [20]:

$$\frac{d[y(t)]}{dt} = K_0(t)\{x[t - \tau(t)] - z(t)\}, \quad (1)$$

$$T_0(t) \frac{d[y(t)]}{dt} + y(t) = K_0(t)\{x[t - \tau(t)] - z(t)\}, \quad (2)$$

где $K_0(t)$ — коэффициент усиления, $T_0(t)$ — постоянная времени объекта с самовыравниванием, $\tau(t)$ — время запаздывания.

Таким образом, считаем, что априорно известны структуры операторов ОНОУ, представленные уравнениями (1) и (2). Нестационарность обусловлена тем, что коэффициенты этих уравнений — неизвестные функции времени в определении которых и состоит задача идентификации.

4. ОПИСАНИЕ АВТОКОЛЕБАНИЙ В ПОЗИЦИОННОЙ СИСТЕМЕ УПРАВЛЕНИЯ

Будем считать ОНОУ на интервале идентификации (в течение периода автоколебаний) квазистационарным. Полагаем, что на интервале идентификации динамические свойства ОНОУ остаются постоянными, т. е. $K_0(t) = \text{const}$, $T_0(t) = \text{const}$, $z(t) = \text{const}$ и $\tau(t) = \text{const}$. В этом случае автоколебания температуры описываются системами конечных уравнений [20], устанавливающими взаимосвязь параметров автоколебаний с искомыми коэффициентами дифференциальных уравнений для ОНОУ без самовыравнивания:

$$\Delta y_{(+)} = \Delta y_0 + (\tau_{\text{off}} + \tau_{\text{d.off}})K_0(x - z); \quad (3)$$

$$\Delta y_{(-)} = \Delta y_0 + (\tau_{\text{on}} + \tau_{\text{d.on}})K_0z; \quad (4)$$

$$T_{\text{on}} = \tau_{\text{on}} + \tau_{\text{d.off}} + \frac{2\Delta y_0 + (\tau_{\text{on}} + \tau_{\text{d.on}})K_0z}{K_0(x - z)}; \quad (5)$$

$$T_{\text{off}} = \tau_{\text{off}} + \tau_{\text{d.on}} + \frac{2\Delta y_0 + (\tau_{\text{off}} + \tau_{\text{d.off}})K_0(x - z)}{K_0z} \quad (6)$$

и с самовыравниванием:

$$\Delta y_{(+)} = K_0(x - z) \left[1 - \exp\left(-\frac{\tau_{\text{off}} + \tau_{\text{d.off}}}{T_0}\right) \right] + \Delta y_0 \exp\left(-\frac{\tau_{\text{off}} + \tau_{\text{d.off}}}{T_0}\right); \quad (7)$$

$$\Delta y_{(-)} = K_0z \left[1 - \exp\left(-\frac{\tau_{\text{on}} + \tau_{\text{d.on}}}{T_0}\right) \right] + \Delta y_0 \exp\left(-\frac{\tau_{\text{on}} + \tau_{\text{d.on}}}{T_0}\right); \quad (8)$$

$$T_{\text{on}} = \tau_{\text{on}} + \tau_{\text{d.off}} + T_0 \ln \frac{K_0x - (K_0z - \Delta y_0) \exp\left(-\frac{\tau_{\text{on}} + \tau_{\text{d.on}}}{T_0}\right)}{K_0(x - z) - \Delta y_0}; \quad (9)$$

$$T_{\text{off}} = \tau_{\text{off}} + \tau_{\text{d.on}} + T_0 \ln \frac{K_0x - [K_0(x - z) - \Delta y_0] \exp\left(-\frac{\tau_{\text{off}} + \tau_{\text{d.off}}}{T_0}\right)}{K_0z - \Delta y_0}, \quad (10)$$

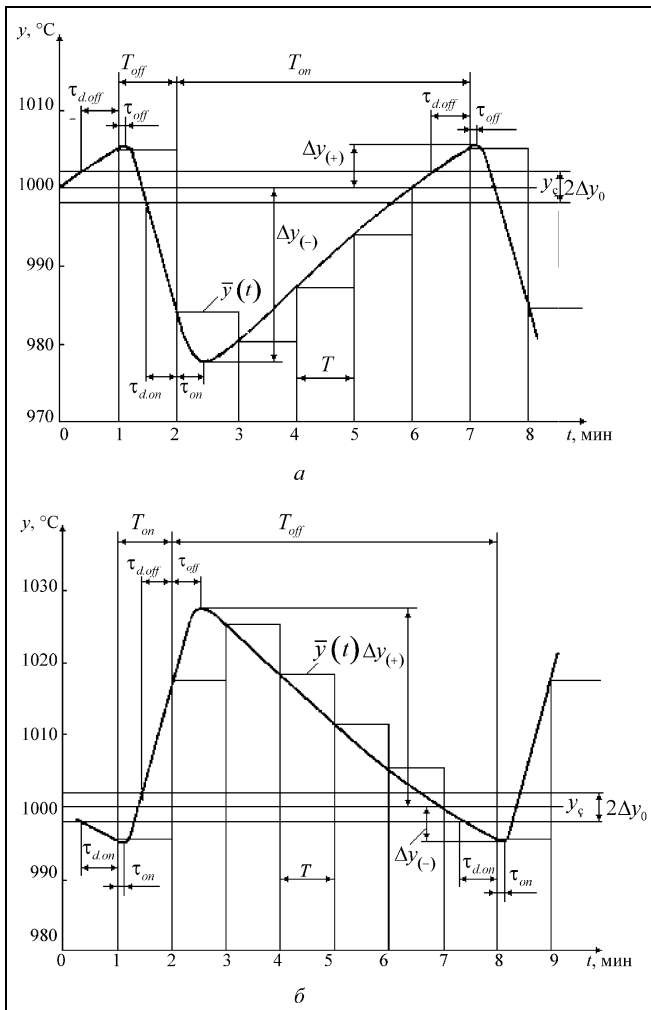


Рис. 4. Автоколебания температуры на интервале идентификации ОНОУ в начале (а) и в конце (б) процесса сепарации титана

где $\Delta y_{(+)}$ и $\Delta y_{(-)}$ — амплитуды положительного и отрицательного отклонений выходной величины от y_3 , T_{on} и T_{off} — время подключения и отключения входной величины, τ_{on} и τ_{off} — время запаздывания объектов при ее подключении и отключении, $\tau_{d.on}$ и $\tau_{d.off}$ — дополнительное время запаздывания МДР при подключении и отключении.

Считаем, что на интервале идентификации ОНОУ параметры автоколебаний могут быть измерены при известных значениях x и Δy_0 (рис. 4).

Тогда задача идентификации ОНОУ без самовыравнивания сводится к определению по соотношениям (3)—(6) и экспериментальным данным автоколебательного процесса неизвестных величин K_0 и z в уравнении (1), а задача идентификации ОНОУ с самовыравниванием — к определению по соотношениям (7)—(10) и экспериментальным данным неизвестных величин K_0 , T_0 и z в уравнении (2).

5. АЛГОРИТМЫ И СТРУКТУРНЫЕ СХЕМЫ ИДЕНТИФИКАЦИИ

Для решения задачи необходимо выбрать критерий качества идентификации — функцию потерь, характеризующую различие моделей и объекта, и разработать алгоритмы идентификации, минимизирующие эту функцию [21].

Введем для количественной оценки степени близости моделей (1), (2) и реального ОНОУ функцию потерь, которая зависит от расчетных и измеренных параметров автоколебаний. Наложим на эту функцию требование [18]:

$$\begin{aligned} \rho = & (\Delta y_{(+)}^p - \Delta y_{(+)}^3)^2 / \Delta y_{(+)}^3 + \\ & + (\Delta y_{(-)}^p - \Delta y_{(-)}^3)^2 / \Delta y_{(-)}^3 + (T_{on}^p - T_{on}^3)^2 / T_{on}^3 + \\ & + (T_{off}^p - T_{off}^3)^2 / T_{off}^3 \rightarrow \min, \end{aligned} \quad (11)$$

где ρ — функция потерь, индекс «р» относится к расчетным значениям параметров автоколебаний, индекс «э» — к экспериментальным.

Заметим, что системы (3)—(6) и (7)—(10) являются переопределенными (на четыре уравнения (3)—(6) — два неизвестных, на четыре уравнения (7)—(10) — три неизвестных) и для разработки алгоритмов решения систем уравнений (алгоритмов идентификации) целесообразно воспользоваться методами оптимизации [22]. Очевидно, что требование (11) будет выполняться, если значения K_0 и z ОНОУ без самовыравнивания и значения K_0 , T_0 и z ОНОУ с самовыравниванием будут найдены из систем уравнений (3)—(6), (7)—(10) и условия минимума функции потерь ρ , записанной для каждого для этих объектов.

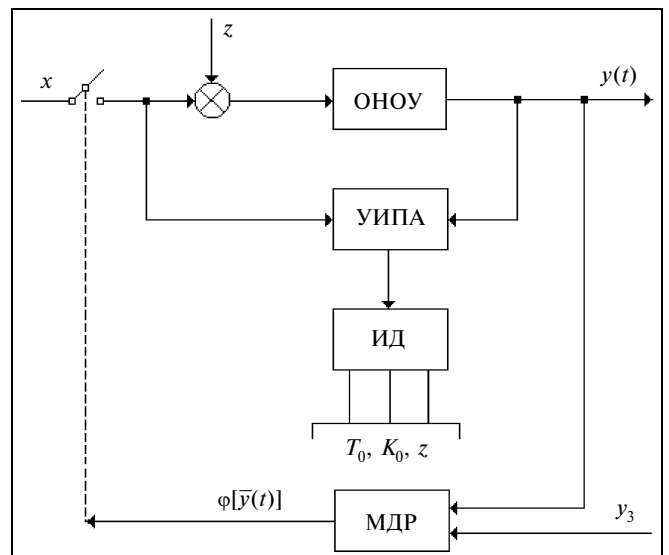


Рис. 5. Структурная схема идентификации ОНОУ в замкнутом контуре позиционного регулирования

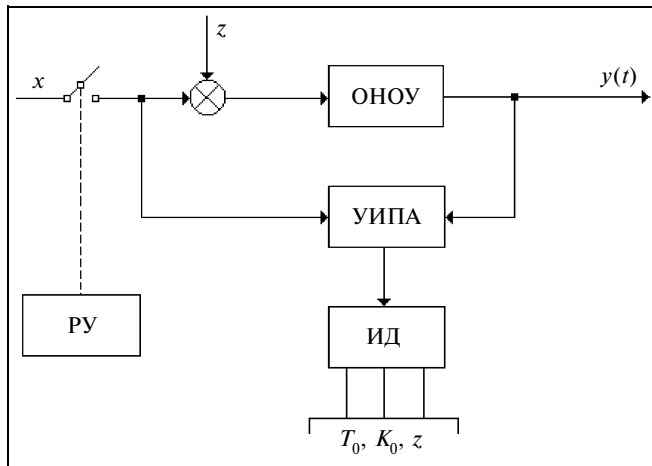


Рис. 6. Структурная схема идентификации ОНОУ в разомкнутом контуре позиционного регулирования

Соответствующие алгоритмы идентификации были реализованы с помощью метода пошаговой многомерной оптимизации с разной точностью [23]. Поиск минимума заканчивается при достижении заданной точности решения задачи, определяемой неравенствами:

$$\begin{aligned} |\Delta y_{(+)}^p - \Delta y_{(+)}^3| < \varepsilon_y, \quad |\Delta y_{(-)}^p - \Delta y_{(-)}^3| < \varepsilon_y, \\ |T_{on}^p - T_{on}^3| < \varepsilon_T \quad \text{и} \quad |T_{off}^p - T_{off}^3| < \varepsilon_T, \end{aligned} \quad (12)$$

где ε_y и ε_T — малые, наперед заданные величины (порядка $10^{-1} \dots 10^{-2}$).

Для решения задачи структурную схему (см. рис. 3) необходимо дополнить устройством измерения параметров автоколебаний УИПА и оценивающим устройством — идентификатором ИД с заложенными в него алгоритмами идентификации (рис. 5).

В процессе функционирования ОНОУ на его входе измеряются параметры автоколебаний T_{on} и T_{off} , а на выходе — параметры автоколебаний $\Delta y_{(+)}$, $\Delta y_{(-)}$ и времена запаздывания τ_{on} , τ_{off} , $\tau_{d.on}$ и $\tau_{d.off}$ (в быстродействующем МДР — микропроцессорном контроллере — можно принять $\tau_{d.on} = \tau_{d.off} = 0$). Идентификатор по результатам измерений вычисляет путем решения систем уравнений (3)—(6), (7)—(10) и минимизации соответствующей функции потерь ρ неизвестные величины K_0 и z ОНОУ без самовыравнивания и K_0 , T_0 и z ОНОУ с самовыравниванием. Задача идентификации ОНОУ в разомкнутом контуре позиционного регулирования (рис. 6) решается аналогично с той лишь разницей, что $y(t)$ поддерживается на уровне y_3 с помощью ручного управления РУ.

Предложенные подходы позволяют автоматизировать процесс идентификации моделей в составе АРМ операторов восстановления и сепарации [23]. Это дает возможность методом пассивного эксперимента непосредственно в рабочем режиме оценивать в реальном времени динамику и возмущения ТП, а также усредненные на интервале функционирования значения их динамических параметров. Существенно снижаются затраты времени на идентификацию и обработку результатов эксперимента.

6. ПРИМЕР

Решение задачи рассмотрим на примере идентификации ОНОУ в начале и в конце процесса сепарации (см. рис. 4).

Требуется определить такие значения K_0 и z ОНОУ без самовыравнивания и значения K_0 , T_0 и z ОНОУ с самовыравниванием, чтобы при из-

Таблица 1

Параметры автоколебаний температуры на интервале идентификации

Стадия процесса сепарации	Значения параметров							
	$\Delta y_{(+)}, ^\circ\text{C}$	$\Delta y_{(-)}, ^\circ\text{C}$	T_{on}, c	T_{off}, c	τ_{on}, c	τ_{off}, c	$\tau_{d.on}, \text{c}$	$\tau_{d.off}, \text{c}$
Начало процесса	5,9	21,1	300	60	25,9	7,5	29,6	34,4
Конец процесса	24,6	5,5	60	360	6,2	28,5	35,6	33,3

Таблица 2

Результаты идентификации

Стадия процесса сепарации	ОНОУ без самовыравнивания		ОНОУ с самовыравниванием		
	$K_0 \cdot 10^{-3}, ^\circ\text{C}/\text{c} \cdot \text{кВт}$	$z, \text{кВт}$	$K_0, ^\circ\text{C}/\text{кВт}$	T_0, c	$z, \text{кВт}$
Начало процесса	3,4	101,5	2,4	708,3	104,3
Конец процесса	3,5	25,7	2,5	700,2	21,1



вестных x , Δu_0 ($x = 130$ кВт, $\Delta u_0 = 2$ °С) и экспериментальным данным автоколебаний температуры (табл. 1) выполнялись условия (12). Результаты решения задачи приведены в табл. 2.

Как видно из табл. 2, параметры K_0 и T_0 ОНОУ изменяются в ходе процесса незначительно. Существенно изменяется (снижается) возмущающее воздействие z (тепло, потребляемое ОНОУ на испарение из губчатого титана примесей магния и хлорида магния).

ЗАКЛЮЧЕНИЕ

Проблема построения моделей сложных много-связных нестационарных технологических объектов — процессов восстановления и вакуумной сепарации губчатого титана — заменена решением в позиционных системах задач идентификации автономных нестационарных объектов управления. С помощью предложенных моделей синтезирован класс систем управления [24—27], обеспечивающих оптимизацию и интенсификацию процессов восстановления и вакуумной сепарации.

Описанный подход может быть применен для идентификации нестационарных технологических объектов в различных отраслях промышленности.

ЛИТЕРАТУРА

1. Тарасов А.В. Металлургия титана. — М.: ИКЦ «Академкнига», 2003. — 328 с.
2. Постановки задач оптимизации магнетермической технологии губчатого титана / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Проблемы и перспективы развития химической технологии на Западном Урале: Сб. науч. тр. — Пермь: ПГТУ, 2001. — С. 148—151.
3. Синтез оптимальных систем управления процессами магнетермического получения губчатого титана / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Математические методы в технике и технологиях — ММТТ-14: Сб. тр. 14-й Междунар. науч. конф. / СФ МЭИ. — Смоленск, 2001. — Т. 2. — С. 89—91.
4. Сергин М.Ю. Современное состояние и возможные пути решения проблем построения систем управления технологическими процессами // Приборы и системы. Управление, контроль, диагностика. — 2004. — № 1. — С. 2—8.
5. Тимченко Б.С. Экспериментально-статистическая оптимизация, контроль и автоматизация металлотермии. — М.: Цветметинформация, 1968. — 103 с.
6. Влияние различных факторов на срок службы реторт аппаратов магнетермического производства титана / О.А. Путина, А.А. Путин, А.И. Гулякин и др. // Цветные металлы. — 1979. — № 9. — С. 71—73.
7. Локшин Р.Г., Ревуцкий Е.Я. Автоматизация магнетермического процесса восстановления четырехлористого титана // Труды ВАМИ — Л., 1964. — № 52. — С. 128—137.
8. Минимизация длительности процессов сепарации титана в АСУТП / Ю.П. Кирин, А.И. Черепанов, Ю.А. Протасов и др. // Цветные металлы. — 1983. — № 1. — С. 51—54.
9. Муромцев Д.Ю., Тюрин И.В. Энергосберегающее управление сложными объектами // Промышленные АСУ и контроллеры. — 2005. — № 11. — С. 24—27.
10. Тюрин И.В., Яшин Е.Н. Оптимальное управление температурным профилем сложных объектов // Приборы и системы. Управление, контроль, диагностика. — 2006. — № 9. — С. 1—4.
11. Куржанский А.Б. Управление и наблюдение в условиях неопределенности. — М.: Наука, 1977. — 392 с.
12. Иодко Э.А. Термодинамика и теплофизика процесса магнетермического восстановления титана // Изв. АН СССР. Металлы. — 1977. — № 4. — С. 30—35.
13. Иодко Э.А. Испарение группы веществ из пористого тела // Инженерно-физический журнал. — 1979. — Т. 37, № 3. — С. 449—455.
14. Исследование температурных полей аппаратов сепарации различной конструкции / А.Н. Петрунько, А.Е. Андреев, Э.А. Арутюнов и др. // Цветные металлы. — 1969. — № 4. — С. 74—76.
15. Изучение температурных полей аппаратов для получения губчатого титана / А.И. Титаренко, Е.А. Черепанова, В.М. Мальшин и др. // Там же. — 1976. — № 4. — С. 50—51.
16. Температурные поля аппаратов восстановления / Е.А. Черепанова, В.М. Мальшин, А.П. Яценко и др. // Там же. — 1982. — № 3. — С. 51—52.
17. Опыт стабилизации температурного режима сепарации реакционной массы при производстве титана / Е.А. Черепанова, С.В. Мушков, А.И. Гулякин и др. // Там же. — 1983. — № 4. — С. 57—58.
18. Построение моделей динамики в системах управления процессами производства губчатого титана / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Вестник КГУ им. Н.А. Некрасова. — Кострома: КГУ, 2006. — Т. 12. — С. 43—47.
19. Основы управления технологическими процессами / Под ред. Н.С. Райбмана. — М.: Наука, 1978. — 440 с.
20. Идентификация математических моделей нестационарных объектов управления / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Математические методы в технике и технологиях — ММТТ-14: Сб. тр. 14-й Междунар. науч. конф. / СФ МЭИ. — Смоленск, 2001. — Т. 2. — С. 92—94.
21. Цыпкин Я.З. Основы информационной теории идентификации. — М.: Наука, 1984. — 320 с.
22. Амосов А.А., Дубинский Ю.А., Копченова Н.В. Вычислительные методы для инженеров. — М.: Высшая школа, 1994. — 544 с.
23. Применение методов оптимизации для идентификации объектов в позиционных системах регулирования / А.В. Затонский, Ю.П. Кирин, В.Ф. Беккер и др. // Математические методы в технике и технологиях — ММТТ-18: Сб. тр. 18-й Междунар. науч. конф. / КГТУ. — Казань, 2005. — Т. 10. — С. 67—70.
24. Управление вакуумной сепарацией губчатого титана с применением систем переменной структуры / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Математические методы в технике и технологиях — ММТТ-15: Сб. тр. 15-й Междунар. науч. конф. / ТГТУ. — Тамбов, 2002. — Т. 9. — С. 210—214.
25. Адаптивное позиционное регулирование с идентификатором / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Математические методы в технике и технологиях — ММТТ-16: Сб. тр. 16-й Междунар. науч. конф. / СПбГТИ (ТУ) — Санкт-Петербург, 2003. — Т. 2. — С. 85—88.
26. Синтез импульсной позиционной системы управления нестационарным объектом / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Математические методы в технике и технологиях — ММТТ-18: Сб. тр. 18-й Междунар. науч. конф. / КГТУ. — Казань, 2005. — Т. 10. — С. 44—47.
27. Совместная разработка технологии и системы управления вакуумной сепарацией губчатого титана / Ю.П. Кирин, А.В. Затонский, В.Ф. Беккер и др. // Приборы и системы. Управление, контроль, диагностика. — 2007. — № 9. — С. 7—10.

e-mail: zxenon2000@yandex.ru

Статья представлена к публикации членом редколлегии А.А. Амбарцумяном. □

СТРУКТУРНО-КЛАССИФИКАЦИОННЫЕ МЕТОДЫ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ В КРУПНОМАСШТАБНЫХ СИСТЕМАХ УПРАВЛЕНИЯ¹

Ю.А. Дорофеев

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Предложен метод решения задач анализа и прогнозирования в крупномасштабных системах управления. В качестве прогнозной модели для каждого объекта принята марковская цепь с r состояниями, где r — число структурных единиц (классов). Для эффективной реализации предложенного метода разработан комплексный алгоритм структуризации (классификации) объектов исследуемой системы.

ВВЕДЕНИЕ

Многие крупномасштабные системы управления, прежде всего, организационно-административные, функционируют в условиях большой информационной размытости и неопределенности. Именно поэтому в последнее время для исследования таких систем стали широко применяться методы структурного анализа данных, базирующиеся на алгоритмах классификационного анализа данных [1].

В настоящей статье рассматривается задача анализа и прогнозирования в крупномасштабных системах управления, причем считается, что такая система состоит из достаточно большого числа объектов, каждый из которых характеризуется многочисленным набором разнородных параметров. Основная идея предлагаемого метода решения такой задачи состоит в том, что исследуются не точные значения параметров, описывающих состояние каждого объекта (например, траектории состояний), а лишь класс, к которому принадлежит каждый объект в рамках некоторой структуры (классификации) множества объектов, входящих в исследуемую крупномасштабную систему [2]. Такое интегральное описание объектов, входящих в

крупномасштабную систему, позволяет существенно повысить эффективность анализа поведения системы, а также устойчивость и робастность процедур принятия управленческих решений и прогнозов. Для формализации задачи используется методология классификационного анализа данных [1].

1. СТРУКТУРНО-КЛАССИФИКАЦИОННЫЙ АНАЛИЗ В КРУПНОМАСШТАБНЫХ СИСТЕМАХ УПРАВЛЕНИЯ

1.1. Постановка задачи

Пусть исследуемая система состоит из n объектов, каждый из которых характеризуется набором из k параметров. Изучается поведение этого множества объектов в дискретные моменты времени. Вводится в рассмотрение k -мерное пространство параметров X , в котором j -й объект в момент времени t представляется точкой $x_j(t) = (x_j^1(t), x_j^2(t), \dots, x_j^k(t))$. Упорядоченная совокупность точек $x_j(t_1), \dots, x_j(t_m)$ представляет собой известную часть траектории, характеризующей динамику j -го объекта.

В большинстве приложений для принятия управленческого решения в момент времени t_m используется совокупная информация об известных траекториях каждого объекта и прогноз значений

¹ Работа выполнена при частичной финансовой поддержке РФФИ, проект 08-07-00349-а.



$x_j(t_m + 1), j = 1, \dots, n$. Как правило, информация по каждому объекту рассматривается независимо от остальных [3]. Однако для многих прикладных задач требуется знать не точные значения параметров-характеристик в моменты времени t_1, t_2, \dots, t_m и прогнозировать значения в момент t_{m+1} , а знать (и прогнозировать) лишь класс, к которому принадлежит (будет принадлежать) этот объект в соответствующие моменты времени в рамках некоторой структуры (классификации) множества объектов изучаемой системы. Так, например, в процессе исследования социально-экономического развития субъектов РФ (в данном случае крупномасштабная система — народное хозяйство РФ) вовсе необязательно знать (и прогнозировать) значения социально-экономических параметров для каждого региона, достаточно лишь знать, в какой класс этот регион попадает в данный и прогнозируемый моменты времени (условно, в классы «хорошо», «средне» и «плохо» развивающихся объектов).

Основу предлагаемого подхода составляет процедура выявления структуры объектов, входящих в исследуемую систему. Предполагается, что вектор значений параметров $x_j(t)$ достаточно полно характеризует состояние j -го объекта в момент времени t . А это, в свою очередь, означает, что взаиморасположение точек $x_1(t), \dots, x_n(t)$ в пространстве X отражает реальную структуру (типологию) исследуемого множества объектов. Для выявления такой структуры в работе применяется комплексный алгоритм автоматической классификации, специально разработанный для решения таких задач. Он включает в себя алгоритмы: m -локальной оптимизации заданного критерия J , выбора информативных параметров, выбора начального разбиения, выбора числа классов, заполнения пропущенных наблюдений. Рассмотрим каждый из этих алгоритмов в отдельности.

1.2. Алгоритм m -локальной оптимизации

Вначале опишем работу алгоритма 1-локальной оптимизации. Для простоты изложения рассматривается случай двух классов $r = 2$. Пусть задано начальное разбиение R_0 всех точек классифицируемой выборки x_1, \dots, x_n . Обозначим через $x_j \in A_1$ точки, относящиеся к первому классу, а через $x_j \in A_2$ — ко второму. Алгоритм итерационный — на каждом шаге рассматривается одна точка из последовательности $x_1, \dots, x_n, x_1, \dots, x_n, x_1, \dots$ («зацикленная» исходная последовательность). Отнесение точки к одному из двух классов обозначается с помощью индекса

$$\rho(x_j) = \begin{cases} 1, & \text{если } x_j \in A_1, \\ -1, & \text{если } x_j \in A_2. \end{cases}$$

Тогда алгоритм 1-локальной оптимизации определяется следующим образом: $\rho(x_j) = \text{sign}[J(x_j \in A_1) - J(x_j \in A_2)]$.

В итоге точка x_j относится к тому классу, при отнесении к которому значение критерия J будет больше (если эти значения равны, то для определенности точка относится к классу с меньшим номером). Алгоритм заканчивается, если на некотором цикле среди точек x_1, \dots, x_n не будет сделано ни одной «переброски» точки из класса в класс.

Алгоритм m -локальной оптимизации — это поэтапное применение к выборке алгоритмов s -локальной оптимизации, $s = 1 \dots m$. На s -м этапе алгоритм работает по той же схеме, только на каждом его шаге происходит пробная «переброска» из класса в класс не одной, а s точек. Подсчитывается значение критерия J до и после «переброски», Принадлежность каждой из s точек к классу либо остается неизменной (значение J до «переброски» больше, чем после), либо меняется на другой класс — в противном случае. В данном случае цикл — это число шагов, равное числу различных s точек в выборке. Доказана сходимость алгоритма за конечное число шагов к локальному максимуму критерия J . Разработан эвристический алгоритм сокращенного перебора, в котором на каждом шаге для пробной «переброски» используются s точек, в определенном смысле ближайших к границе между классами.

При моделировании и в приложениях в качестве критерия J принимался функционал J_1 средней близости точек в классах, определяемый через потенциальную функцию близости точек x и y :

$$K(x, y) = 1/\{1 + \alpha R^p(x, y)\}, \quad (1)$$

где $R(x, y)$ — расстояние между точками x и y , α и p — настраиваемые параметры алгоритма. Средняя близость точек в классе определяется как

$$K(A_i, A_j) = \frac{2}{n_i(n_i - 1)} \sum_{i=1}^{n_i} \sum_{j>i} K(x_i, x_j),$$

где $K(x_i, x_j)$ определяется формулой (1), n_i — число точек в классе A_i . Тогда критерий J_1 определяется как

$$J_1 = \sum_{i=1}^r \frac{n_i}{n} K(A_i, A_i). \quad (2)$$

Специально отметим частный случай алгоритма m -локальной оптимизации для $k = 1$ (одномерный случай). Дело в том, что одномерный случай имеет уникальное свойство, существенно упрощающее процедуры целенаправленного перебора, применяемые для автоматической классифика-

ции, а именно: ввиду одномерной упорядоченности классов границей между двумя классами (в детерминированном случае) служит только одна точка, и таких границ может быть не более двух (для крайне правого и крайне левого классов — только одна). Работа детерминированного (в отличие от общего — размытого) варианта модификации этого алгоритма для одномерного случая описана в работе [4].

1.3. Алгоритм выбора информативных параметров

Этот алгоритм базируется на одном из алгоритмов экстремальной группировки параметров, а именно, на алгоритме «квадрат» [5]. В результате его применения получают разбиение исходных k параметров на небольшое (заданное) число групп, а также значения факторов для этих групп. В приложениях используются либо новые интегральные параметры — факторы групп, либо набор параметров, каждый из которых является ближайшим к фактору в соответствующей группе.

В большинстве приложений исходные или выделенные информативные параметры имеют неравнозначную важность для определения структуры объектов. Для выявления таких показателей важности в работе предлагается применять процедуру экспертного оценивания. Наиболее хорошие результаты дает процедура многовариантной экспертизы [6], когда к получению параметра важности для каждого оцениваемого параметра привлекаются несколько групп экспертов — специалистов в различных аспектах исследуемой проблемы. В результате процедуры экспертизы каждый параметр получает определенный вес (показатель его «важности») для формирования структуры объектов.

1.4. Алгоритм построения начального разбиения

На первом шаге из всех точек выборки x_1, \dots, x_n находится пара наиболее удаленных друг от друга точек, x_l и x_p , одна из которых x_l относится к первому классу, а другая x_p — ко второму. Если n достаточно велико, то используется усеченный вариант первого шага, а именно: x_l выбирается случайно, а x_p ищется как точка, наиболее от нее удаленная.

На втором шаге ищутся точки x_{l+1} и x_{p+1} — ближайšie, соответственно, к точкам x_l и x_p ; точка x_{l+1} относится к первому классу, а x_{p+1} — ко второму.

На $(s+1)$ -ом шаге ищутся точки x_{l+s} и x_{p+s} , ближайšie в среднем к уже найденным точкам,

соответственно, первого и второго классов. Точка x_{l+s} определяется следующим образом:

$$x_{l+s} = x_j / \min_{x_j} \frac{1}{s} \sum_{m=0}^s K(x_j, x_{l+m}). \quad (3)$$

Точка x_{p+s} определяется аналогично. Если возникает «конфликт», т. е. одна и та же точка является ближайшей к первому и ко второму классам одновременно, то она относится к первому классу. Процедура (3) повторяется до тех пор, пока не будут исчерпаны все точки выборки. Полученное разбиение принимается в качестве начального разбиения R_0 .

1.5. Алгоритм выбора числа классов

Для выбора числа классов применяется специальная экспертно-компьютерная процедура, которая заключается в следующем. Сначала эксперт-пользователь оценивает диапазон (r_{\min}, r_{\max}) , в пределах которого заведомо находится искомое число классов. Далее, с помощью любого алгоритма автоматической классификации (в настоящей работе применялся алгоритм m -локальной оптимизации), анализируемое множество объектов разбивается на $r_{\min}, r_{\min} + 1, \dots, r_{\max}$ классов. Качество каждой из полученных классификаций оценивалось с помощью критерия $J_3 = J_1 - qJ_2$, где критерий J_1 вычисляется по формуле (2), величина J_2 , а также некоторые вспомогательные величины вычисляются по формулам:

$$J_2 = \frac{1}{r-1} \sum_{i=1}^r \sum_{j>i}^r \frac{n_i + n_j}{n} (A_i, A_j);$$

$$K(A_i, A_j) = \frac{1}{n_i n_j} \sum_{x_l \in A_i} \sum_{x_p \in A_j} K(x_l, x_p) — \text{мера близости}$$

классов A_i и A_j , где потенциальная функция $K(x_i, x_j)$ определяется формулой (1); q , α и p из формулы (1) — настраиваемые параметры алгоритма. Фактически, параметр q является масштабирующим параметром, приводящим к соизмеримым средним значениям функционалов J_1 и J_2 ; на практике значение q примерно 2...7 (обычно во столько раз отличается средняя близость внутри классов от средней близости между самими классами).

Формально, в качестве «оптимального» можно выбрать такое число классов r_{opt} , которое соответствует максимальному значению $J_3(r_j)$, т. е. $r_{\text{opt}} = r_j$, для которого $\max J_3(r_j)$, $r_j = r_{\min}, \dots, r_{\max}$. Однако наличие существенной, но неиспользованной при классификации информации, например, ввиду отсутствия данных, может привести к тому, что по-



лученное таким способом r_{opt} не будет «истинно оптимальным».

Для компенсации этого недостатка предлагается следующая экспертная процедура. Экспертам-специалистам в соответствующей предметной области представляются значения $J_3(r_j)$, $r_j = r_{\text{min}}, \dots, r_{\text{max}}$, представленные для удобства в виде графика, на котором отмечается значение r_{opt} (оно соответствует максимальной точке на графике $J_3(r_j)$). Используя эту информацию, эксперты могут корректировать выбираемое число классов. В подавляющем большинстве случаев экспертное число классов либо совпадает со значением r_{opt} , либо незначительно (± 1) отличается от него.

При классификации многомерных объектов во время такой экспертизы анализируется также классификация каждого объекта. Для этой цели экспертам сообщается информация о мере близости $K(x_i, c_j)$ каждой точки x_i до центров классов c_j , $j = 1, \dots, r_{\text{opt}}$ в оптимальной классификации, т. е. матрица близости $\|K(x_i, c_j)\|$, $i = 1, \dots, n, j = 1, \dots, r_{\text{opt}}$. Перенесение точки (объекта) x_i из j -го класса в l -й считается допустимым, если величины $K(x_i, c_j)$ и $K(x_i, c_l)$ отличаются незначительно. Другими словами, содержательно обоснованное перенесение допустимо для точек, расположенных вблизи границы между соответствующими классами.

1.6. Алгоритм заполнения пропущенных наблюдений

Во многих приложениях имеются пропуски в данных. В этой ситуации нужно либо применять специальные процедуры подсчета расстояний между объектами, в параметрах которых имеются пропуски, либо разрабатывать специальные процедуры заполнения таких пропусков. В подавляющем большинстве случаев пропуски по каждому параметру заполняются средним известных значений соответствующего параметра (для исходной выборки). В настоящей работе была разработана специальная процедура заполнения пропусков в исходных данных с помощью алгоритмов автоматической классификации. Ее основная идея состоит в следующем. Если множество изучаемых объектов структурировано (т. е. их можно разделить на классы, достаточно компактно расположенные в пространстве параметров X), то дисперсия (диапазон) изменения каждого параметра в пределах каждой группы, как правило, будет существенно меньше, чем этот показатель для значения этого параметра по всей выборке. Таким образом, если по данным с пропусками удастся определить реальную структуру взаиморасположения точек (т. е. провести классификацию, адекватную этой струк-

туре), то заполнять пропущенное значение l -го параметра для объекта из i -го класса можно средним этого параметра по его известным значениям для всех объектов, попавших в i -й класс. Исходя из сделанного предположения, отклонение полученного значения от «истинного» должно быть существенно меньше (в среднем), чем обычная схема заполнения по общему среднему.

2. МЕТОДИКА СТРУКТУРНО-КЛАССИФИКАЦИОННОГО ПРОГНОЗИРОВАНИЯ В КРУПНОМАСШТАБНЫХ СИСТЕМАХ УПРАВЛЕНИЯ

2.1. Динамическая структуризация исследуемых объектов

Вначале (в момент времени t_1) с помощью комплексного алгоритма автоматической классификации, описанного в § 1, производится структуризация n точек в пространстве X на r классов, каждый из которых и характеризует определенный тип объекта. Число классов r выбирается с помощью человеко-машинной процедуры, входящей в комплексный алгоритм автоматической классификации. Вводится понятие модели (эталона) класса $a_i(t)$, $i = 1, \dots, r$ (чаще всего — это центр класса) [1]. Для каждого объекта, кроме принадлежности к классу, вычисляются расстояния до эталонов всех классов $R_{ij}(t)$, $i = 1, \dots, r; j = 1, \dots, n$.

Заметим, что на практике структуризация объектов чрезвычайно редко проводится в пространстве исходных признаков, обычно сначала выделяется набор информативных параметров. В настоящей работе для этой цели применялась специальная процедура, также входящая в комплексный алгоритм, она описана в п. 1.3.

В момент времени t_2 каждая точка $x_j(t_2)$ с помощью одного из алгоритмов распознавания образов с учителем относится к тому или иному классу в рамках классификации, полученной на первом шаге. Для этого применяется алгоритм метода потенциальных функций, который в спрямляющем пространстве эквивалентен алгоритму ближайшего среднего [7]. А именно, каждая точка $x_j(t_2)$ относится к классу A_l для которого заданная мера близости $K(x_j(t_2), A_l)$ точки $x_j(t_2)$ к этому классу максимальна, т. е. $K(x_j(t_2), A_l) = \max_i K(x_j(t_2), A_i)$, $i = 1, \dots, r$, $i = 1, \dots, n$. В качестве такой меры близости служит величина $K(x_j, A_l) = \frac{1}{n_l} \sum_{x_i \in A_l} K(x_j, x_i)$, где n_l — число точек в классе A_l , $K(x, y)$ — потенциальная функция (1).

После определения принадлежности всех точек к тому или иному классу, производится пересчет эталонов $a_l(t_2)$, $i = 1, \dots, r$. Для каждой точки с

предыдущего шага пересчитываются, а для каждой новой точки вычисляются расстояния до новых эталонов $R(x_j(t_2), a_i(t_2))$, $i = 1, \dots, r, j = 1, \dots, n$. Такая процедура выполняется для всех m моментов времени. В итоге для каждого объекта получается последовательность (траектория) из m позиций. В каждой позиции находится $r + 1$ число, первое из которых — это номер класса, к которому относился этот объект в соответствующий момент времени, а последующие числа — это значения расстояний до центров классов в тот же момент времени. Требуется спрогнозировать номер класса (тип объекта), к которому будет относиться каждый объект в момент времени t_{m+1} .

2.2. Алгоритм прогнозирования

В качестве прогнозной модели для каждого объекта используется марковская цепь с r состояниями, т. е. на каждом шаге рассчитываются элементы матрицы переходных вероятностей $P = \|p_{ij}\|$, $j = 1, \dots, n; i = 1, \dots, r$. Разработан специальный алгоритм пересчета на каждом шаге соответствующих переходных вероятностей p_{ij} с использованием информации о значениях расстояний до центров классов и условий нормировки $\sum_{i=1}^r p_{ij} = 1$ для всех $j = 1, \dots, n$.

Алгоритм работает следующим образом. Пусть после первого шага, для точек $x_j(t_1)$ подсчитаны расстояния до эталонов $R_{ji}^{(1)} = R(x_j(t_1), a_i(t_1))$, $i = 1, \dots, r, j = 1, \dots, n$. Тогда элементы матрицы переходных вероятностей $p_{ji}^{(1)} = p_{ji}(t_1)$ рассчитываются следующим образом:

$$p_{ji}^{(1)} = \alpha_j^{(1)} / R_{ji}^{(1)}, \quad (4)$$

где $\alpha_j^{(1)} = \frac{\prod_{i=1}^r R_{ji}^{(1)}}{\sum_{l=1}^r \frac{1}{R_{jl}^{(1)}} \prod_{i=1}^r R_{ji}^{(1)}}$ — нормирующий множитель.

На s -м шаге элементы матрицы переходных вероятностей (4) модифицируются с помощью следующей процедуры. Введем обозначения $\Delta R_{ji}^{(s)} = R_{ji}^{(s-1)} - R_{ji}^{(s)}$, $\Delta \hat{R}_{ji}^{(s)} = \frac{R_{ji}^{(s-1)} - R_{ji}^{(s)}}{R_{ji}^{(s-1)} + R_{ji}^{(s)}}$. Если j -я точка

совпадает с эталоном i_0 -го класса ($x_j(t_s) = a_{i_0}(t_s)$), т. е. $R_{ji_0}^{(s)} = 0$, то

$$p_{ji}^{(s)} = \begin{cases} 1, & \text{если } i = i_0, \\ 0, & i = 1, \dots, r, i \neq i_0. \end{cases}$$

Другими словами, если точка совпадает с эталоном некоторого класса, то вероятность для этой точки остается в этом классе равна 1, а вероятность перехода в другой класс равна 0.

Для случая, когда $R_{ji_0}^{(s)} \neq 0$, все переходные вероятности модифицируются по следующей схеме:

$$p_{ji}^{(s)} = \gamma \left[p_{ji}^{(s-1)} + \left(\frac{1 + \text{sign}(\Delta R_{ji}^{(s)})}{2} - p_{ji}^{(s-1)} \text{sign}(\Delta R_{ji}^{(s)}) \right) \Delta \hat{R}_{ji}^{(s)} \right], \quad (5)$$

где, как обычно, $\text{sign}(z) = \begin{cases} 1, & \text{если } z \geq 0, \\ -1, & \text{если } z < 0, \end{cases}$ а γ — нормирующий множитель, определяемый условием нормировки переходных вероятностей $\sum_{i=1}^r p_{ji}^{(s)} = 1$:

$$\gamma = \frac{1}{1 + ((1 + \text{sign}(\Delta R_{ji}^{(s)}))/2 - p_{ji}^{(s-1)} \text{sign}(\Delta R_{ji}^{(s)})) \Delta \hat{R}_{ji}^{(s)}}.$$

Введение в формулу (5) величины $\text{sign}(\Delta R_{ji}^{(s)})$ вызвано необходимостью модификации различными способами переходных вероятностей для случаев увеличения и уменьшения расстояния от точки $x_j(t_s)$ до эталонов классов $a_i(t_s)$ на s -м шаге. А именно, в случае уменьшения величины $R_{ji}^{(s)}$ по отношению к $R_{ji}^{(s-1)}$ (т. е. $\Delta R_{ji}^{(s)} < 0$), соответствующая переходная вероятность изменяется путем ее увеличения на некоторую долю от $(1 - p_{ji}^{(s-1)})$; а в случае увеличения величины $R_{ji}^{(s)}$ по отношению к $R_{ji}^{(s-1)}$ (т. е. $\Delta R_{ji}^{(s)} > 0$), соответствующая переходная вероятность изменяется путем ее уменьшения на некоторую долю от $p_{ji}^{(s-1)}$. Это необходимо для выполнения условий нормировки для переходных вероятностей $0 < p_{ji}^{(s)} < 1, i = 1, \dots, r$.

Построенная с помощью описанного алгоритма матрица переходных вероятностей P используется для прогнозирования принадлежности объекта к тому или иному классу. На практике обычно применяется не рандомизированная, а байесовс-



кая схема, когда объект относится к тому классу i_0 , для которого $p_{ji_0} = \max_{i=1, \dots, r} p_{ji}$. В случае равенства переходных вероятностей p_{ji} для прогнозируемого объекта для двух или нескольких классов, он относится к классу с наименьшим номером.

2.3. Модификации

Разработана модификация процедуры прогнозирования, когда классификация объектов задается заранее (например, экспертным путем) и в последующем остается неизменной.

Разработан также вариант алгоритма «с памятью», когда используются данные только об s прошлых состояниях множества объектов (s — глубина памяти алгоритма).

Оказалось, что для некоторых приложений (с достаточно высоким уровнем помех при измерении параметров) существенно более эффективными оказываются алгоритмы размытой классификации, в том числе с фоновым классом [1].

ЗАКЛЮЧЕНИЕ

Разработанная методика применялась для анализа и совершенствования процедур принятия решений для нескольких больших систем управления, в основном регионального уровня, в том числе — региональная система управления здравоохранением, пассажирскими автоперевозками, система анализа, управления и прогнозирования социально-экономического развития субъектов

РФ и др. Во всех приложениях, а также при машинном моделировании была подтверждена высокая эффективность разработанного комплекса алгоритмов структурно-классификационного анализа и прогнозирования.

ЛИТЕРАТУРА

1. Бауман Е.В., Дорофеюк А.А. Классификационный анализ данных / Тр. Междунар. конф. по проблемам управления. — М.: СИНТЕГ, 1999. — Т. 1. — С. 62–67.
2. Дорофеюк А.А., Дорофеюк Ю.А. Методы структурно-классификационного прогнозирования многомерных динамических объектов // Искусственный интеллект. — 2006. — № 2. — С. 138–141.
3. Статистическое моделирование и прогнозирование / Под ред. А.Г. Гранберга. — М.: Финансы и статистика, 1990. — 382 с.
4. Процедуры классификационного анализа в задаче формирования информативных признаков при исследовании ритмической структуры биосигнала / А.А. Десова, А.А. Дорофеюк, В.В. Гучук и др. // Автоматика и телемеханика. — 2008. — № 6.
5. Браверман Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. — М.: Наука, 1983.
6. Дорофеюк А.А., Покровская И.В., Чернявский А.Л. Экспертные методы анализа и совершенствования систем управления // Автоматика и телемеханика. — 2004. — № 10. — С. 172–188.
7. Айзерман М.А., Браверман Э.М., Розоноэр Л.И. Метод потенциальных функций в теории обучения машин. — М.: Наука, 1970.

☎ (495) 334-90-70, e-mail: tigress86@bk.ru

Статья представлена к публикации членом редколлегии А.С. Манделем. □



Мультиконференция по информационным технологиям и управлению в промышленности

Институт проблем управления им. В.А. Трапезникова РАН,
3—7 июня 2009 г., г. Москва

В рамках мультиконференции пройдут следующие мероприятия:

- 13-й симпозиум «Information Control Problems in Manufacturing» (INCOM'09) — информационные технологии и управление в промышленности под эгидой Международной федерации по автоматическому управлению (International Federation of Automatic Control — IFAC)
- Международная научно-практическая конференция-выставка «Автоматизация в промышленности»
- Семинар «Информационные технологии в промышленности» для руководящего звена промышленности и бизнеса

Дополнительную информацию можно получить по тел./факсу (495) 334-89-90,
на сайтах <http://incom09.org>, <http://multi.sicpro.org>.
E-mail: incom09@ipu.ru, income@ipu.ru.

**CONTENTS & ABSTRACTS**

A CONCEPT OF A MATHEMATICAL KNOWLEDGE BANK FOR SCIENTIFIC RESEARCH. P. 1. A METAPHOR. 2

Kleschev A.S.

Basis on an extendable internal model of mathematical practice and of a model for the analogy between proofs, a concept of a computer support system for scientific research in the field of mathematics and its mechanisms is offered. A system metaphor and the structure of an external model of mathematical practice are presented.

NETWORK PROGRAMMING TECHNIQUE IN THE SYMMETRIC TRAVELING SALESMAN PROBLEM. 7

Burkova I.V.

A dual problem is formulated where the constraints are split into 2 groups with corresponding division of arcs into 2 parts and solution of the resulting 2 evaluation problems. The sum of the objective functions of optimal solutions to the evaluation problems gives the lower bound for the original problem. The solution of the evaluation task is reduced to the design of the shortest i -trees. A new method for building the lower bounds for evaluation problems underlain by the shortest-distance tree design is proposed. The paper shows that the design of i -trees and the shortest-distance tree for the original distance matrix would not deliver an optimal solution to the dual problem.

ON THE SPECTRUM OF A RING-STRUCTURED DIRECTED GRAPH'S ADJACENCY MATRIX AND ITS APPLICATION 11

Agaev R.P., Nikiforov S.V., Andryushina N.A.

In the paper, it was investigated that if 2 arcs were removed from ring-structured directed graph with 2 Hamilton circuits, its spectrum would remain real if and only if the number of nodes were even, and the «distance between the arcs» removed from one cyclic path were maximal. The applicability of the results to fault-tolerance estimation of ring topology networks is examined.

NON-DIFFERENTIABLE 2-D RAVINE TEST FUNCTIONS 16

Rykov A.S., Matvienko M.Yu.

The paper discusses the design of two-dimensional ravine functions for testing optimization techniques. These demonstrate a new type of non-differentiable test-functions featuring the ravines with acute piecewise-linear ravine bottom and piecewise-linear slopes. The algorithm for designing ravine test functions with desirable properties is offered.

CONTROLLABLE SYSTEMS ROBUST WITH RESPECT TO INITIAL DATA 20

Yakovenko G.N.

A concept of robustness with respect to initial data is proposed: several control impacts transfer a system from a fixed initial state to a same finite state; the same controls transfer the system from any other initial state to a same finite state. A theorem is proved: a controllable system is robust w. r. t. initial conditions if and only if it allows the maximal group of state symmetries. Case studies of various controllable systems robustness investigation are included.

GENERALIZED PROBABILISTIC CRITERIA IN THE NONLINEAR PARAMETRIC IDENTIFICATION PROBLEM 24

Sokolov S.V., Kucherenko P.A.

The topicality of investigating new and enhancing the existing methods of nonlinear stochastic parametric identification is shown. A solution to an identification problem based on applying the generalized probability criteria explicitly dependent on a posteriori density function is proposed. An identification algorithm is synthesized using the criterion of minimum estimation error probability. A numerical example illustrating the effectiveness of the approach proposed is included. The method proposed can be effectively applied in various fields such as communication, control, measurement, etc.

ANALYSIS OF COLLECTIVISM AND EGOISM IN THE CONTEXT OF SOCIAL WELFARE 30

Chebotaev P.Yu., Loginov A.K., Tsodikova Ya.Yu., Lezina Z.M., Borzenko V.I.

Comparative utility of major behavioral patterns including collectivism and egoism is investigated in the context of group decision making.

APPLICATION OF GENERALIZED MEDIAN SCHEMES FOR THE DESIGN OF STRATEGY-PROOF MULTICRITERION ACTIVE EXPERTISE MECHANISMS 38

Burkov V.N., Iskakov M.B., Korghin N.A.

Multicriterion active expertise mechanisms are represented as generalized median voter schemes in terms of right/left coalitions systems. This allows to apply the results of the public choice theory to the design of strategy-proof mechanisms.

MULTICRITERION COMPETITIVENESS EVALUATION AND THE RANKING OF ECONOMIC OBJECTS OPERATING UNDER VARIOUS UNCERTAINTY CONDITIONS 48

Kuvshinov B.M., Chelyadin A.S., Shiryaev V.I.

An approach to multiple criteria estimation of competitiveness of economical objects and to their ranking is proposed. The approach is based on estimating their economic potential. Pattern recognition techniques is used to estimate the competitiveness. The peculiarities of the economic potential estimation problem as well as those of the interval estimates comparison under uncertain source information are examined.

VENDOR SELECTION MANAGEMENT SUBJECT TO MARKET INFORMATION INFLUENCE 55

Baikin A.A., Ivanov E.Yu., Isaeva O.V.

This paper presents the investigation results of a vendor selection model modified subject to information cost. The investigation was undertaken for several types of production functions, such as linear, Leontieff, Cobb-Douglas, Stone, CES, and Philippov. The estimates of the nature and the influence of supplemental market information on the structure and volumes of the resources used by the company are presented.

PARAMETERS EVOLUTION TECHNIQUES IN NONLINEAR NETWORK PROBLEMS WITH A CASE STUDY OF THE OPTIMIZATION OF INCOMING LOAD FLOWS DISTRIBUTION 65

Zhevnerov V.A.

Parameters evolution techniques is proposed to apply in nonlinear network optimization tasks. The application features are demonstrated in a case study of optimal distribution of incoming load flows in a data communication network. The paper shows that the method ensures significant problem time reduction as against the known algorithms under comparable implementation costs.

PROCESS IDENTIFICATION PROBLEMS IN SPONGY TITANIUM PRODUCTION 71

Kirin Yu.P., Zatonsky A.V., Bekker V.F., Kraev S.L.

The features of spongy titanium recovery and vacuum separation process modeling are discussed. The paper suggests to describe processes dynamics with differential equations with variable coefficients. The methods of model dynamics identification in positional control systems are included.

STRUCTURALLY-CLASSIFICATION METHODS FOR ANALYSIS AND FORECASTING IN LARGE-SCALE CONTROL SYSTEMS 78

Dorofeyuk J.A.

The method for solving analysis and forecasting problems in large-scale control systems is presented. A markovian chain with r states, where r — the number of structural units (classes) was used as a forecasting model. For the effective realization of presented method the complex classification algorithm was developed.