

А.В. Лотов, И.И. Пospelova

**Конспект лекций по теории и методам
многокритериальной оптимизации**

Учебное пособие

Москва
2014

Предисловие

Вниманию читателя предлагается конспект, составленный на основе курсов лекций “Основы многокритериальной оптимизации” и “Многокритериальные системы поддержки принятия решений”, читавшихся д.ф.-м.н. А.В. Лотовым на факультете Вычислительной математики и кибернетики МГУ им. М.В. Ломоносова в 1996-2013 годах. Конспект был переработан и дополнен доцентом ВМиК МГУ, к.ф.-м.н. И.И. Поспеловой.

Конспект предназначен в первую очередь для студентов старших курсов кафедры системного анализа и кафедры исследования операций факультета ВМиК, но может быть использован всеми, кто интересуется проблемами многокритериальной оптимизации и знаком с основами скалярной (одно-критериальной) оптимизации.

Введение

В практических задачах принятия решений часто встречается ситуация, когда нельзя ограничиться рассмотрением единственного критерия выбора решения. Попытка математической формализации таких задач привела к созданию теории многокритериальной оптимизации, которая используется при разработке методов, предназначенных для поддержки лиц, принимающих решения, при наличии нескольких критериев выбора. Данный конспект лекций посвящен основам теории многокритериальной оптимизации и методов поддержки принятия решений при нескольких критериях.

Надо отдавать себе отчет в том, что имеется принципиальное различие между теорией многокритериальной оптимизации (МКО) и практикой разработки методов поддержки принятия решений, основанных на ее использовании. *Теория многокритериальной оптимизации* — это математическая дисциплина, базирующаяся на аксиомах выбора решения и изучающая следствия этих аксиом. Она является развитием теории обычной, однокритериальной, оптимизации.

Напротив, методы поддержки принятия решений — это прикладная дисциплина, являющаяся составной частью *теории принятия решений*, которая не является строгой математической наукой, а базируется на знаниях о лице, принимающем решение (ЛПР). Под *лицом, принимающим решение*, принято понимать человека или группу лиц, ответственных за выбор действия, решения в какой-либо области. Отметим, что, в отличие от методов однокритериальной оптимизации, методы многокритериальной оптимизации требуют активного участия ЛПР в процессе поиска предпочтительного решения. Поэтому понимание сущности процесса принятия решения и возможностей человека в этом процессе является основой анализа методов, предназначенных для поддержки ЛПР. Таким образом, основой теории принятия решений является модель человека, в связи с чем эта теория принадлежит к социально-психологическим дисциплинам, хотя, конечно, использует различные математические методы.

Надо сразу сказать, что в настоящее время теория принятия решений является совокупностью концепций, зачастую не связанных между собой, а иногда и противоречащих друг другу (см. [9]). Такая ситуация объясняется тем, что теория принятия решений возникла, по крайней мере, из трех источников, принадлежащих различным отраслям человеческого знания, которые отличаются и языком, и методами исследования. Рассмотрим эти три источника.

1. Экономическая теория, которая объясняет и предсказывает экономическое поведение людей, т.е. поведение в области производства, обмена и потребления. Эта теория привнесла в теорию принятия решений такое базовое понятие, как функция полезности. Все же, к теории принятия решений экономическая теория имеет лишь косвенное отношение, так как ее задача — описание *характерного* поведения человека (а вовсе не его поддержка

в процессе принятия решений). Поэтому экономическая теория использует простую модель разумного человека — ограниченного, но логичного существа (по существу, автомата), действия которого направлены на максимизацию функции полезности. Эта модель показала свою эффективность при анализе экономических проблем. В то же время, при разработке методов поддержки принятия решений человеком такой подход не слишком эффективен — функция полезности обычно заранее не известна, и, более того, ее существование часто вызывает сомнение.

2. Психология, т.е. наука, изучающая психику человека, в том числе и его поведение. Одной из целей психологии является построение адекватных моделей принятия решений человеком в различных ситуациях, так что априорные модели человеческого поведения (типа максимизации полезности) хотя и встречаются в психологических исследованиях, не определяют их основные направления. В настоящее время психология включает совокупность различных конкурирующих моделей человека — от простейшей модели поведения в виде реакции на стимулы до сложнейших концепций ментальных (умственных) моделей, используемых человеком для предсказания реакции окружающего мира на его возможные действия. При этом часто учитывается социальная обусловленность предпочтений и привычек отдельного человека. Такие модели в той или иной степени являются обобщением результатов экспериментов с человеком и наблюдений над ним. Надо, однако, признать, что анализировать знания, добытые экспериментальной и теоретической психологией, значительно труднее, чем изучать простую математическую модель максимизации функции полезности, поэтому концепция функции полезности также иногда используется психологами, особенно в теоретических исследованиях.

3. Инженерное направление связано с практическим применением методов поддержки принятия решений. Систематическая разработка таких методов началось после Второй мировой войны, а с появлением компьютеров началось создание компьютерных систем поддержки принятия решений (СППР). Для инженерного подхода главным является не теоретический результат, а работающая система, которая может быть успешно использована людьми при принятии решений. Начав с применения методов однокритериальной оптимизации и простейших форм функции полезности, создатели СППР были вынуждены обратиться к более сложным постановкам, в том числе и к многокритериальным. В связи с широким распространением СППР являются источником важного (часто негативного) опыта, который требует своего осмысления. Это определяет усиление влияния инженерного направления на теорию поддержки принятия решений. Синтез различных концепций и практического опыта, особенно опыта создания и использования СППР, стал основой быстрого развития теории принятия решений, наблюдаемого в настоящее время. Разработано огромное число подходов и методов для анализа широкого класса проблем, в том числе и методы поддержки принятия решений при нескольких критериях, в условиях неопределенности и риска, при наличии вербальных (словесных) моделей и т.д.

Данная книга посвящена некоторым важнейшим аспектам многокритериальной оптимизации и теории и методам принятия решений при многих критериях в условиях определенности на основе математического моделирования множества возможных решений и количественных критериев выбора решения. В связи с этим многие вопросы теории принятия решений остались за пределами книги. В частности, в ней не рассматриваются методы поддержки принятия решений в условиях неопределенности и риска, а также методы, применяемые в экспертных СППР, которые базируются на предоставлении пользователям знаний экспертов в виде определенной структуры, удобной для тех, кто пользуется этими знаниями (см., например, гл. 9 и 10 в книге [9]).

Для того чтобы более точно описать круг задач, изучаемых в данной книге, рассмотрим широко распространенное представление о структуре процесса принятия решений, сформулированное лауреатом Нобелевской премии Гербертом Саймоном¹⁾ еще в 1960 году. Он выделял четыре основных этапа этого процесса:

- 1) идентификация проблемы и сбор информации,
- 2) выделение относительного небольшого числа альтернативных вариантов действия,
- 3) окончательный выбор наиболее предпочтительного варианта на основе детального анализа отобранных альтернатив,
- 4) применение выбранного варианта и анализ информации, полученной в этом процессе.

Собственно принятие решения осуществляется на втором и третьем этапах, которые сильно различаются. Выделение малого числа альтернатив из большого (часто бесконечного) числа возможных вариантов действия на втором этапе требует всестороннего (т.е. многокритериального) анализа этих вариантов. При этом анализ большого числа вариантов может быть довольно грубым, но должен по возможности включать все такие варианты. На третьем этапе, наоборот, требуется тщательно проанализировать и сравнить лишь малое число альтернатив, уже сформулированных в явном виде. Таким образом, должны существовать по крайней мере два типа многокритериальных методов — методы для поддержки выбора из большого (или бесконечного) числа возможных решений, предназначенные для второго этапа, и методы для поддержки выбора из малого числа альтернатив на третьем этапе.

Важно также, что методы, используемые на этих двух этапах, могут отличаться между собой ролью математического моделирования. Это отличие в значительной степени зависит от области практической деятельности, в которой применяются эти методы. Так, при поддержке принятия решений в политических или социальных проблемах, т.е. в тех областях, где методы

¹⁾Н. Simon The New Science of Management Decision. New York: Harper and Row, 1960.

математического моделирования только начинают развиваться, как правило, применяются многокритериальные подходы, не требующие использования каких-либо математических моделей. В таких случаях основное внимание уделяется третьему этапу, а выделение малого числа альтернативных вариантов осуществляется на основе опыта экспертов. Более того, даже оценка значений критериев для выделенных альтернатив часто осуществляется экспертами. Поскольку число альтернатив мало, методы поддержки принятия решений в этом случае направлены на классификацию альтернатив по предпочтительности (по-английски — outranking methods).

Наоборот, в тех областях человеческой деятельности, где математическое моделирование является традиционным средством анализа проблем (скажем, при проектировании технических систем), математическое моделирование используется в полной мере, как на втором, так и на третьем этапах. При этом эксперты продолжают играть свою важную роль, их опыт дополняет результаты использования моделей. Наконец, в таких областях, как анализ развития экономики страны или группы стран, корпоративное планирование или разработка и выбор системы мероприятий в области финансовой деятельности (финансовая инженерия), математические методы находят все более широкое применение.

Под *методами многокритериальной оптимизации* принято понимать методы поиска решения в задачах с большим (обычно бесконечным) числом возможных вариантов решения в том случае, когда условия задачи и количественные критерии выбора формулируются в математическом виде. Этим методам, наряду с математической теорией многокритериальной оптимизации, посвящена данная книга. Методы классификации малого числа альтернатив в данной книге рассматриваются лишь в §9 для того, чтобы дать общее представление об этом интересном направлении теории принятия решений.

Сразу отметим, что в большинстве методов многокритериальной оптимизации второй и третий этапы процесса принятия решения не разделяются, так что исследователи сразу стараются найти единственное решение, в наибольшей степени удовлетворяющее ЛПР. Это создает определенные проблемы при использовании таких методов в СППР, поэтому в последнее время все большее внимание уделяется методам многокритериальной оптимизации, позволяющим сформировать несколько альтернативных решений для дальнейшего детального анализа с помощью специально предназначенных методов.

В заключение рассмотрим новые тенденции, наблюдающиеся в процессе совершенствования методов многокритериальной оптимизации в последние годы. Новый импульс развитию методов был получен в связи с распространением компьютерной графики, позволившей использовать компьютерную визуализацию при разработке методов поддержки принятия решений, в том числе многокритериальных методов. Это связано с тем, что многокритериальные методы, как уже говорилось, требуют активного участия ЛПР в процессе поиска предпочтительного решения, что требует от исследователей обеспечить легкость взаимодействия человека с компьютером в процес-

се применения методов. Визуализация информации, под которой понимается трансформация символьных данных в геометрические фигуры, помогающие человеку при формировании ментального образа этих данных, значительно облегчает взаимодействие ЛПР с компьютером. Визуализация информации в процессе человеко-машинного взаимодействия является характерной чертой новых компьютерных технологий, позволяющей облегчить процесс усвоения информации человеком и повысить эффективность этого процесса. Известно, что человеческое восприятие в значительной степени базируется на зрении (по некоторым оценкам, около половины нейронов человеческого мозга так или иначе связано со зрением), поэтому компьютерная визуализация является прямым путем к усвоению информации, а применение визуализации в рамках многокритериальных методов в значительной степени облегчает их использование.

Отметим еще одну функцию визуализации, важную в рамках практики применения многокритериальных методов. Речь идет об интеллектуализации СППР, под которой понимается методика создания таких систем, которые могут использоваться лицом, принимающим решение, самостоятельно, без участия посредников. Дело в том, что сложность методов может потребовать в процессе их применения участия специалистов (часто авторов этих методов). Такое требование зачастую является неприемлемым для ЛПР в силу конфиденциальности данных, самого процесса принятия решений или его результатов. Использование визуализации в многокритериальных методах может способствовать их интеллектуализации, приводящей к исключению посторонних лиц из процесса принятия решений. Интеллектуальные компьютерные СППР должны обеспечивать простоту взаимодействия пользователя с ними и брать на себя функции посредников, чему, бесспорно, может способствовать компьютерная визуализация.

Наконец, совершенствование многокритериальных методов с середины 90-х годов XX века стимулируется развитием компьютерных сетей, потребовавших разработать такие методы, которые могли бы применяться в сети Интернет и других сетях. Этот факт также важно учитывать при анализе многокритериальных методов.

Предлагаемая книга состоит из двух глав. В первой главе дается введение в теорию многокритериальной оптимизации (более основательное изложение теории МКО можно найти в книге [12]), во второй главе рассматриваются методы многокритериальной оптимизации, а также один из методов классификации альтернатив.

ГЛАВА 1. ЭЛЕМЕНТЫ ТЕОРИИ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

Глава 1. Бинарные отношения как средство описания предпочтений

Предпочтение человека в самом простом, элементарном виде — это выражение того факта, что человек предпочел бы выбрать один из двух объектов, который, таким образом, является для него более предпочтительным. Отсюда следует, что предпочтения удобно выразить с помощью математического понятия, характеризующего пары объектов. Для этого можно предложить несколько различных подходов, наиболее естественный из которых основан на использовании так называемых *бинарных отношений*. Напомним, что бинарным называется отношение, заданное для пары объектов некоторого множества.

Вообще говоря, бинарные отношения применяются для описания попарных связей различного характера между объектами произвольной природы. На множестве действительных чисел определены известные из школьного курса алгебры бинарные отношения “больше”, “меньше”, “равно”, “не больше”, с помощью бинарных отношений описываются отношения предшествования, принадлежности и т.д. Хотя в дальнейшем нас будут интересовать только бинарные отношения предпочтения, т.е. отношения типа “лучше”, “не хуже” и “безразлично”, сначала рассмотрим общие понятия теории бинарных отношений и только затем перейдем к интересующим нас бинарным отношениям предпочтения.

1.1. Основные понятия

Пусть G — непустое множество произвольной природы. Рассмотрим прямое произведение $G^2 = G \times G = \{(x, y) \mid x, y \in G\}$ и выделим некоторое подмножество $\mathcal{R}$ множества G^2 .

Определение 1.1. *Говорят, что пара (x, y) находится в бинарном отношении $x\mathcal{R}y$, если $(x, y) \in \mathcal{R}$.*

Заметим, что использование одной и той же буквы (скажем, $\mathcal{R}$) для обозначения и множества $\mathcal{R} \subseteq G^2$, и бинарного отношения $(x, y) \in \mathcal{R}$ зачастую неудобно, поэтому для обозначения множества и бинарного отношения могут использоваться различные символы.

Поскольку бинарные отношения между элементами G задаются подмножествами G^2 , к бинарным отношениям применимы все теоретико-множественные операции, в частности, можно рассматривать объединение бинарных отношений $\mathcal{R}_1$ и $\mathcal{R}_2$, обозначаемое через $\mathcal{R}_1 \cup \mathcal{R}_2$, и их пересечение $\mathcal{R}_1 \cap \mathcal{R}_2$. Можно также рассмотреть дополнение к бинарному отношению

$\mathcal{R}$, которое определяется как $\overline{\mathcal{R}} = G^2 \setminus \mathcal{R}$. Помимо этого, определяются следующие операции:

- *сужение* бинарного отношения на множество $G_1 \subset G$:
 $\mathcal{R}_{G_1} = \{(x, y) \mid (x, y) \in (\mathcal{R} \cap G_1^2)\}$;
- бинарное отношение, *обратное* к данному:
 $\mathcal{R}^{-1} = \{(x, y) \in G^2 \mid (y, x) \in \mathcal{R}\}$.

Бинарные отношения могут обладать некоторыми свойствами, определяемыми свойствами тех попарных связей, которые моделируются с помощью бинарных отношений. Общеприняты следующие названия свойств бинарных отношений.

Определение 1.2. *Бинарное отношение $\mathcal{R}$ называется*

- *рефлексивным*, если $x\mathcal{R}x$ для любого $x \in G$;
- *иррефлексивным (антирефлексивным)*, если $x\overline{\mathcal{R}}x$ для любого $x \in G$;
- *симметричным*, если из $x\mathcal{R}y$ следует $y\mathcal{R}x$;
- *асимметричным*, если из $x\mathcal{R}y$ следует $y\overline{\mathcal{R}}x$;
- *антисимметричным*, если из $y\mathcal{R}x$, $x\mathcal{R}y$ следует $x = y$;
- *транзитивным*, если из $x\mathcal{R}y$, $y\mathcal{R}z$ следует $x\mathcal{R}z$;
- *негатранзитивным*, если $\overline{\mathcal{R}}$ транзитивно;
- *ацикличным*, если из $x\mathcal{R}z_1$, $z_1\mathcal{R}z_2, \dots, z_n\mathcal{R}y$ следует $x \neq y$;
- *связным (полным)*, если для любых $x, y \in G$ либо $x\mathcal{R}y$, либо $y\mathcal{R}x$, либо и то, и другое;
- *слабо связным (слабо полным)*, если для любых $x, y \in G$ таких, что $x \neq y$, либо $x\mathcal{R}y$, либо $y\mathcal{R}x$, либо и то, и другое.

Отметим, что неполное (несвязное) бинарное отношение иногда называют частичным.

Следующие утверждения показывают, что свойства бинарных отношений из определения 1.2 не являются независимыми.

Утверждение 1.1. *Если бинарное отношение $\mathcal{R}$ иррефлексивно и транзитивно, то оно асимметрично.*

Доказательство. Предположим противное, т.е. пусть $\mathcal{R}$ является иррефлексивным и транзитивным, но оно не асимметрично. По определению 1.2, отрицание асимметричности означает, что найдутся такие $x, y \in G$, что $x\mathcal{R}y$ и $y\mathcal{R}x$. Тогда в силу транзитивности получаем, что $x\mathcal{R}x$, но это противоречит иррефлексивности. $\square$

Утверждение 1.2. *Если бинарное отношение $\mathcal{R}$ иррефлексивно и транзитивно, то оно ациклично.*

Доказательство. Предположим противное, т.е. пусть $\mathcal{R}$ является иррефлексивным и транзитивным, но оно не ациклично. По определению 1.2, отрицание ацикличности означает, что найдутся такой $x \in G$ и такая последовательность $z_1, z_2, \dots, z_n \in G$, что $x\mathcal{R}z_1, z_1\mathcal{R}z_2, \dots, z_n\mathcal{R}x$. В силу транзитивности получаем, что $x\mathcal{R}x$, но это противоречит иррефлексивности. $\square$

Далее мы рассмотрим один из видов бинарных отношений, наиболее часто используемый на практике — отношение порядка.

1.2. Бинарные отношения порядка

Определение 1.3. *Бинарное отношение называется квазипорядком, если оно рефлексивно и транзитивно (обычно обозначается $\mathcal{R}$).*

Определение 1.4. *Бинарное отношение называется эквивалентностью, если оно рефлексивно, транзитивно, симметрично (обычно обозначается $\mathcal{I}$).*

Легко проверить, что если $\mathcal{R}$ — бинарное отношение квазипорядка, то $\mathcal{I} = \mathcal{R} \cap \mathcal{R}^{-1}$ является бинарным отношением эквивалентности (в этом случае говорят, что бинарное отношение эквивалентности $\mathcal{I}$ порождено бинарным отношением квазипорядка $\mathcal{R}$). Заметим, что симметричность означает, что одновременно выполнено $x\mathcal{R}y$ и $y\mathcal{R}x$, т.е. x и y эквивалентны в смысле бинарного отношения $\mathcal{R}$, откуда и происходит его название.

Определение 1.5. *Бинарное отношение называется строгим порядком¹⁾, если оно иррефлексивно и транзитивно (обычно обозначается $\mathcal{P}$).*

Из утверждений 1.1 и 1.2 следует, что отношение строгого порядка также является асимметричным и ациклическим.

Пусть $\mathcal{R}$ — квазипорядок, $\mathcal{I}$ — порожденное им бинарное отношение эквивалентности, тогда бинарное отношение $\mathcal{P}$, определенное как

$$\mathcal{P} = \mathcal{R} \setminus \mathcal{I} = \mathcal{R} \setminus (\mathcal{R} \cap \mathcal{R}^{-1}) = \mathcal{R} \setminus \mathcal{R}^{-1},$$

является отношением строгого порядка.

Примерам, иллюстрирующим введенные понятия, могут служить определяемые на множестве действительных чисел отношения квазипорядка “ $\leq$ ”, эквивалентности “ $=$ ” и строгого порядка “ $<$ ”.

Свойства отношения эквивалентности

Напомним следующее определение.

¹⁾Иногда для этого отношения используются и другие названия, например, “частичный строгий порядок”.

Определение 1.6. Совокупность множеств G_α , где $\alpha \in A$ (A — множество, не обязательно счетное), называется разбиением множества G , если

$$1) G = \bigcup_{\alpha \in A} G_\alpha,$$

$$2) G_\alpha \neq \emptyset, \forall \alpha \in A,$$

$$3) G_{\alpha_1} \cap G_{\alpha_2} = \emptyset \text{ для таких } \forall \alpha_1, \alpha_2 \in A, \text{ что } \alpha_1 \neq \alpha_2.$$

Имеет место следующее утверждение, доказательство которого приводить не будем.

Теорема 1.1. 1. Пусть $\mathcal{I}$ — бинарное отношение эквивалентности на G , $G \neq \emptyset$. Найдется такое разбиение $G_\alpha, \alpha \in A$ множества G , что $x\mathcal{I}y$ тогда и только тогда, когда найдется такое $\alpha \in A$, что $x, y \in G_\alpha$.

2. Пусть $\{G_\alpha, \alpha \in A\}$, — разбиение множества G . Бинарное отношение $\mathcal{R}$ такое, что $(x, y) \in \mathcal{R}$ тогда и только тогда, когда найдется такое $\alpha \in A$, что $x \in G_\alpha, y \in G_\alpha$, является бинарным отношением эквивалентности.

Таким образом, можно говорить, что отношение эквивалентности разбивает множество G на классы эквивалентности.

Свойства бинарного отношения строгого порядка $\mathcal{P}$

Теперь изучим свойства бинарного отношения строгого порядка $\mathcal{P}$ на непустом подмножестве X множества определения бинарного отношения G .

Определение 1.7. Элемент $x \in X$ называется $\mathcal{P}$ -максимальным (максимальным по $\mathcal{P}$) на X , если $y \bar{\mathcal{P}} x$ для $\forall y \in X$.

Существуют и другие, эквивалентные, определения элемента, максимального по $\mathcal{P}$ на X .

Определение 1.8. Элемент $x \in X$ называется $\mathcal{P}$ -максимальным на X , если $\{y \in X \mid y \mathcal{P} x\} = \emptyset$.

Определение 1.9. Элемент $x \in X$ называется $\mathcal{P}$ -максимальным на X , если из того, что $y \in G$ и $y \mathcal{P} x$ следует $y \in G \setminus X$.

В соответствии с этими определениями максимальные элементы часто называют *недоминируемыми*. Заметим, что в определениях 1.7–1.9 условие $y \neq x$ выполняется автоматически, так как отношение $\mathcal{P}$ не является рефлексивным.

Совокупность $\mathcal{P}$ -максимальных элементов на X будем обозначать через $\text{Max}_{\mathcal{P}} X$. Отметим одно важное свойство множеств максимальных элементов.

Утверждение 1.3. Пусть на $X \subseteq G$ заданы отношения строгого порядка $\mathcal{P}_1$ и $\mathcal{P}_2$, причем $\mathcal{P}_1 \subseteq \mathcal{P}_2$. Тогда $\text{Max}_{\mathcal{P}_2} X \subseteq \text{Max}_{\mathcal{P}_1} X$.

Доказательство. Пусть $x^0 \in \text{Max}_{\mathcal{P}_2} X$. Тогда $\{y \in X \mid y\mathcal{P}_2 x^0\} = \emptyset$ по определению 1.8. Поскольку $\mathcal{P}_1 \subseteq \mathcal{P}_2$, то, значит, и $\{y \in X \mid y\mathcal{P}_1 x^0\} = \emptyset$, а это, в свою очередь, означает, что $x^0 \in \text{Max}_{\mathcal{P}_1} X$. $\square$

Для бинарного отношения строгого порядка можно ввести понятие решения.

Определение 1.10. Говорят, что отношение $\mathcal{P}$ обладает свойством фон Неймана-Моргенштерна (НМ-свойством) на $X^* \subseteq X$, если для любого $x \in X$ либо $x \in X^*$, либо найдется $y \in X^*$ такой, что $y\mathcal{P}x$.

Определение 1.11. Если отношение $\mathcal{P}$ обладает НМ-свойством на $\text{Max}_{\mathcal{P}} X \subseteq X$, то $\text{Max}_{\mathcal{P}} X$ называется решением по фон Нейману-Моргенштерну (ядром по фон Нейману-Моргенштерну) для бинарного отношения $\mathcal{P}$ на X .

Возникает естественный вопрос о том, в каких случаях существует решение по фон Нейману-Моргенштерну. Рассмотрим одно простое утверждение.

Утверждение 1.4. Пусть X непусто и конечно. Тогда $\text{Max}_{\mathcal{P}} X$ непусто и является решением по фон Нейману-Моргенштерну для отношения $\mathcal{P}$ на X .

Доказательство. Отметим прежде всего, что из НМ-свойства $\mathcal{P}$ на $\text{Max}_{\mathcal{P}} X \subseteq X$ следует непустота $\text{Max}_{\mathcal{P}} X$, поэтому достаточно доказать НМ-свойство. Докажем его от противного, т.е. предположим, что отношение $\mathcal{P}$ не обладает НМ-свойством на $\text{Max}_{\mathcal{P}} X \subseteq X$. По определению 1.10 это означает, что существует такой $y \in X$, что $y \notin \text{Max}_{\mathcal{P}} X$ и не существует $x \in \text{Max}_{\mathcal{P}} X$, для которого $x\mathcal{P}y$. Из этого следует, что существует $z^{(1)} \in X$, $z^{(1)} \notin \text{Max}_{\mathcal{P}} X$, такой, что $z^{(1)}\mathcal{P}y$. Поскольку $z^{(1)}$ не является $\mathcal{P}$ -максимальным на X , то существует такой $z^{(2)} \in X$, что $z^{(2)}\mathcal{P}z^{(1)}$. Если бы $z^{(2)} \in \text{Max}_{\mathcal{P}} X$, то в силу транзитивности $\mathcal{P}$ имело бы место $z^{(2)}\mathcal{P}y$, $z^{(2)} \in \text{Max}_{\mathcal{P}} X$, что противоречит предположению. Таким образом, $z^{(2)} \in X \setminus \text{Max}_{\mathcal{P}} X$. Рассуждая дальше аналогичным образом, получим бесконечную последовательность $z^{(1)}, z^{(2)}, \dots, z^{(k)}, \dots$, для которой $z^{(1)}\mathcal{P}y$, $z^{(2)}\mathcal{P}z^{(1)}$, $\dots$, $z^{(k)}\mathcal{P}z^{(k-1)}$, $\dots$, причем $z^{(l)} \in X \setminus \text{Max}_{\mathcal{P}} X$, $l = 1, 2, \dots$.

В силу ацикличности $\mathcal{P}$ в этой последовательности все элементы различны, что противоречит конечности X . $\square$

Если множество X бесконечно, то условия того, что $\text{Max}_{\mathcal{P}} X$ непусто и является решением по фон Нейману-Моргенштерну, оказываются более сложными. Они рассмотрены в §3.

Бинарные отношения предпочтения

Бинарные отношения порядка могут быть использованы для формализации понятия предпочтения. Пусть требуется описать взаимную предпочтительность элементов множества G .

1. Если элемент $x' \in G$ не менее предпочтителен (*не хуже*), чем элемент $x'' \in G$, то будем писать $x' \succeq x''$.

2. Если безразлично, какой из двух элементов $x', x'' \in G$ предпочесть, то будем писать $x' \approx x''$.

3. Если элемент $x' \in G$ более предпочтителен (*лучше*), чем элемент $x'' \in G$, то будем писать $x' \succ x''$.

Таким образом, оказались введены три бинарных отношения, описывающих предпочтения. Если предположить, что эти бинарные отношения транзитивны, то в силу их смысла отношение $x' \succeq x''$ является квазипорядком, отношение $x' \approx x''$ — симметричной частью этого квазипорядка, а отношение $x' \succ x''$ — асимметричной частью. Таким образом, отношение $x' \approx x''$ является эквивалентностью, а $x' \succ x''$ — строгим порядком.

1.3. Отношения предпочтения в задаче скалярной оптимизации

Проиллюстрируем использование бинарных отношений на примере задачи скалярной оптимизации.

Пусть все мыслимые варианты решения образуют пространство W , называемое *пространством решений*. Пусть задана числовая функция — единственный критерий выбора решения

$$f : W \rightarrow \mathbb{R}^1,$$

где $\mathbb{R}^1$ — множество всех действительных чисел. Функцию f называют также *критерием оптимизации*. Для определенности будем считать, что предпочтительным является увеличение значения f .

Как правило, выбор решения осуществляется в рамках некоторых ограничений. Формально это значит, что выделяется некоторое подмножество $X \subseteq W$, называемое *множеством допустимых решений*. Тогда выбор решения трактуется как выбор некоторого элемента из X , доставляющего максимум критерию выбора решения. Таким образом, мы приходим к задаче скалярной (однокритериальной) оптимизации, которую принято записывать как

$$f(x) \rightarrow \max, \quad x \in X. \quad (1.1)$$

Описанная формулировка является самой простой с точки зрения формализации задачи выбора решения из заданного множества. В общем случае и значение критерия, и множество допустимых решений могут зависеть от случайных факторов, действий каких-либо других лиц, от изменения общей обстановки, в которой производится выбор решения. Задачи такого типа изучаются в классической теории исследования операций [3], и мы не будем рассматривать эти вопросы в рамках данного курса. Таким образом, множество X и критерий f будем считать заданными.

В случае единственного критерия выбора решения проблема поиска наилучшего решения сводится к чисто математической проблеме решения задачи (1.1). Мы не будем останавливаться на методах решения задачи скалярной оптимизации, см., например, [2], [13], а переформулируем ее в терминах бинарных отношений.

Поскольку на множестве действительных чисел $\mathbb{R}^1$ имеет место естественный квазипорядок, то в пространстве решений W с помощью единственного критерия f можно ввести

квазипорядок: $x' \succeq x'' \Leftrightarrow f(x') \geq f(x'')$;

эквивалентность: $x' \approx x'' \Leftrightarrow f(x') = f(x'')$;

строгий порядок: $x' \succ x'' \Leftrightarrow f(x') > f(x'')$.

Как обычно, эквивалентность разбивает пространство решений W на классы, которые в данном случае имеют вид:

$$f(x') = f(x'').$$

Понятие решения задачи оптимизации обычно вводят на основе использования квазипорядка.

Определение 1.12. *Элемент $x^* \in X$ называется решением задачи оптимизации (1.1), если $x^* \succeq x$ для всех $x \in X$.*

Отметим, что при некоторых простых предположениях такое решение задачи (1.1) существует (например, достаточно потребовать непрерывность функции $f(x)$ и компактность X). Если решение не единственно, т.е. несколько решений принадлежат к одному классу эквивалентности (совпадают по значению критерия с $f(x^*)$), то достаточно найти любое из них. Методы оптимизации обычно находят какое-то одно решение задачи, которое и предлагается использовать.

Определить понятие решения задачи оптимизации можно и на основе применения бинарного отношения строгого порядка и использования максимальных элементов множества допустимых решений X .

Определение 1.13. *Элемент $x^* \in X$ называется решением задачи оптимизации (1.1), если $x^* \in \text{Max}_\succ X$.*

Это определение, как легко показать, совпадает с определением, данным на основе квазипорядка. Такое совпадение объясняется тем, что и квазипорядок, и строгий порядок в однокритериальных задачах оптимизации основаны на сравнении действительных чисел, а любая пара действительных чисел связана квазипорядком “не меньше”, который, таким образом, является связным бинарным отношением.

1.4. Функция полезности

В связи с простотой формулировки задачи скалярной оптимизации внимание исследователей давно привлекал вопрос о том, когда бинарные отношения предпочтения $x' \succ x''$ и $x' \succeq x''$, заданные на парах элементов

множества G , можно описать с помощью числовой функции, заданной на элементах множества G .

Определение 1.14. Числовая функция $u(x)$, определенная на множестве G , называется функцией полезности, соответствующей бинарному отношению предпочтения $\succ$ (или $\succeq$), если $x' \succ x'' \Leftrightarrow u(x') > u(x'')$ ($x' \succeq x'' \Leftrightarrow u(x') \geq u(x'')$).

Условия существования функции полезности, соответствующей бинарному отношению $\succ$, заданному на конечном множестве G , дает следующая теорема, доказанная в конце XIX века.

Теорема 1.2. Для существования функции полезности $u(x)$, соответствующей бинарному отношению строгого порядка, заданному на конечном множестве G , необходимо и достаточно, чтобы строгий порядок был негатранзитивен.

Заметим, что достаточным условием негатранзитивности строгого порядка является его слабая связность. Таким образом, для слабо связанных отношений строгого порядка, заданных на конечных множествах, функция полезности, соответствующая этому отношению, существует. Немецкий математик Г. Кантор обобщил это утверждение на бесконечные множества элементов, используя при этом одно трудно проверяемое условие, на котором мы останавливаться не будем. Более удобное условие предложил французский экономист Ж. Дебре в середине XX века.

Пусть, для простоты, G принадлежит конечномерному линейному пространству $\mathbb{R}^m$ с некоторой метрикой. Пусть на множестве G задано бинарное отношение квазипорядка $\succeq$ и порожденное им отношение строгого порядка $\succ$. Дополнительно потребуем, чтобы отношение $\succeq$ обладало свойством непрерывности.

Определение 1.15. Отношение предпочтения $\succeq$ называется непрерывным на G , если множество $\{(x, y) | x \succ y\}$ является открытым подмножеством G^2 .

Содержательно это определение означает следующее: если элемент x^0 строго предпочтительнее элемента y^0 ($x^0 \succ y^0$), то при малом изменении каждого из этих элементов отношение строгого предпочтения сохраняется, т.е. если x и y близки соответственно к x^0 и y^0 , то $x \succ y$. Данное определение непрерывности можно заменить другим, эквивалентным.

Определение 1.16. Отношение предпочтения $\succeq$ называется непрерывным на G , если

- 1) множество $\{y | x \succ y\}$ открыто в G при любом $x \in G$,
- 2) множество $\{x | x \succ y\}$ открыто в G при любом $y \in G$.

Несмотря на то, что предположение о непрерывности отношения предпочтения кажется весьма естественным, нетрудно указать довольно простое и часто используемое бинарное отношение, не удовлетворяющее этому требованию. Оно называется *лексикографическим отношением предпочтения* и определяется следующим образом: $x \succeq y$, если $x_1 > y_1$, либо если существует $1 < p < t$ такой, что $x_1 = y_1, x_2 = y_2, \dots, x_p = y_p, x_{p+1} > y_{p+1}$. Отсутствие непрерывности у такого отношения предпочтения показывается следующим пример.

Пример. Пусть $G = \mathbb{R}^2$ и элементы $x', x'' \in \mathbb{R}^2$ таковы, что $x'_1 = x''_1$ и $x'_2 > x''_2$. Нетрудно видеть, что в таком случае множество $\{x | x \succ x'\}$ необязательно открыто в $\mathbb{R}^2$, поскольку в любой окрестности точки x' могут найтись такие точки x , что $x_1 < x'_1$. Рассмотрим, например, $x' = (0, 0.001)$, $x'' = (0, 0)$. Ясно, что $x' \succeq x''$, но для точки $x = (-0.0001, 0.001)$, находящейся в достаточно малой окрестности x' , это отношение уже не выполнено.

Отсутствие непрерывности лексикографического отношения предпочтения приводит к тому, что его рекомендуется использовать в дискретных задачах.

Существование функции полезности, представляющей отношение квазипорядка $\succeq$, утверждает теорема Дебре (см. [1]). В упрощенном виде для рассматриваемого нами случая $G \subseteq \mathbb{R}^m$ она имеет следующий вид.

Теорема 1.3. *Если множество $G \subseteq \mathbb{R}^m$ связно, а квазипорядок $\succeq$ является связным (полным) квазипорядком, непрерывным на G , то существует такая непрерывная функция полезности $u(x)$, что из $x' \succeq x''$ следует $u(x') \geq u(x'')$, а из $u(x') \geq u(x'')$ следует $x' \succeq x''$.*

Такую функцию полезности принято называть соответствующей бинарному отношению. Для частичного (неполного) отношения, для которого не существует функции полезности, соответствующей этому отношению, представляет интерес ответ на вопрос о существовании такой функции полезности $u(x)$, что из $x' \succeq x''$ следует $u(x') \geq u(x'')$. Известно, что всякое бинарное отношение может быть пополнено до связного (полного), поэтому поставленный вопрос сводится к свойствам связных отношений. В то же время, такое пополнение обычно неоднозначно, что приводит к сложности использования аппарата функций полезности при построении теории оптимизации при нескольких критериях. В то же время, построение функций полезности является одним из средств, используемых при разработке методов многокритериальной оптимизации (см. Главу 2).

Глава 2. Задача многокритериальной оптимизации

Итак, как уже говорилось в §1, при применении методов скалярной оптимизации предпочтительность решений описывается как увеличение значения некоторой единственной функции (единственного критерия выбора решения, например, функции полезности). Благодаря этому, проблема выбора наилучшего способа действия сводится к математической задаче поиска допустимого решения, доставляющего критерию максимальное значение. Если таких критериев несколько, то приходим к задаче *многокритериальной оптимизации*. Формулировка этой задачи требует подробного рассмотрения.

Предположим, что в задаче выбора решения каждому элементу пространства решений W ставится в соответствие m -мерный вектор критериев, т.е. считается заданным отображение

$$\varphi : W \rightarrow \mathbb{R}^m, \quad (2.1)$$

где W — пространство решений, в котором задано множество допустимых решений X . Таким образом, отображение φ является вектор-функцией. Сделанное предположение резко сужает класс задач принятия решения при нескольких критериях выбора, которые изучаются методами многокритериальной оптимизации (МКО). В частности, в МКО непосредственно не рассматриваются задачи с качественными и случайными переменными, а также многие другие задачи. Конечно, упомянутые задачи можно попытаться преобразовать к виду (2.1), но это уже другой вопрос.

Введем понятие доминирования по Парето — одно из основных понятий теории МКО.

2.1. Доминирование по Парето

Доминированием по Парето называется бинарное отношение строгого порядка, которое определяется следующим образом.

Если рассмотреть на линейном (векторном) m -мерном пространстве $\mathbb{R}^m$ квазипорядок

$$y' \mathcal{R} y'' \iff y'_j \geq y''_j, \quad j \in \{1, \dots, m\},$$

то, как и всякий квазипорядок, он порождает на $\mathbb{R}^m$ отношение эквивалентности $\mathcal{I}$, которое в данном случае имеет вид

$$y' \mathcal{I} y'' \iff y' = y'',$$

и отношение строго порядка $\mathcal{P} = \mathcal{R} \setminus \mathcal{I}$, которое можно имеет вид

$$y' \mathcal{P} y'' \iff \{y'_i \geq y''_i, \quad y'_i \neq y''_i\}, \quad (2.2)$$

где $y' \geq y''$ означает $y'_i \geq y''_i, \quad i = 1, \dots, m$.

Так определенное бинарное отношение $\mathcal{P}$ имеет название *доминирования по Парето*. Для бинарного отношения Парето наряду с обозначением $\mathcal{P}$ часто используется обозначение $\succ_p$. Если $y' \succ_p y''$, то говорят, что вектор y' *доминирует* вектор y'' по Парето. Легко заметить, что бинарное отношение Парето транзитивно и иррефлексивно, т.е. является строгим порядком.

Отметим, что, вообще говоря, для введения бинарных отношений в линейных пространствах удобно использовать понятие конуса¹⁾. Пусть в $\mathbb{R}^m$ задан непустой выпуклый конус K , содержащий начало координат. Тогда квазипорядок в $\mathbb{R}^m$ можно ввести, например, так:

$$y' \mathcal{R} y \iff y' - y \in K,$$

а строгий порядок

$$y' \mathcal{P} y \iff y' - y \in K \setminus \{0\}.$$

Доминирование по Парето задается конусом K , совпадающим с неотрицательным ортантом $\mathbb{R}_+^m$. Поэтому

$$y' \succ_p y \iff y' - y \in \mathbb{R}_+^m \setminus \{0\},$$

а множество векторов, доминирующих по Парето точку $y \in \mathbb{R}^m$, задается конусом $y + \mathbb{R}_+^m \setminus \{0\}$. Этот факт важен при графической интерпретации понятия множества максимальных элементов доминирования по Парето.

2.2. Формулировка задачи МКО

Определение 2.1. *Задачей многокритериальной оптимизации называется задача выбора решения из некоторого множества $X \subseteq W$, имеющего бесконечное (или очень большое) число решений, причем критерии выбора задаются вектор-функцией (2.1), а предпочтения — доминированием Парето (2.2).*

В задаче МКО линейное (векторное) m -мерное пространство $\mathbb{R}^m$ принято называть (*критериальным пространством*), вектор $y \in \mathbb{R}^m$ — *критериальным вектором*, а каждую его компоненту — *частным критерием выбора*. По этой причине задачу МКО иногда называют задачей *векторной оптимизации*. Компоненты вектор-функции φ часто называют *целевыми функциями*.

Как легко увидеть, отличие задачи МКО от задачи скалярной оптимизации состоит в том, что критериальное пространство многомерно. В одномерном случае, как уже говорилось, для описания предпочтения на числах действительной оси может быть использован строгий порядок, который совпадает с доминированием по Парето.

¹⁾Множество $X \subseteq \mathbb{R}^m$ называется конусом, если $\lambda x \in X$ для любого $x \in X$ и любого числа $\lambda \geq 0$. Заметим, что по определению $0 \in X$. В дальнейшем под конусом в $\mathbb{R}^m$ мы будем понимать, наряду с обычным конусом с вершиной в точке 0, также “сдвинутый” конус с вершиной в любой точке $y \in \mathbb{R}^m$.

Отметим, что доминирование по Парето (2.2), введенное выше, подразумевает, что предпочтительно увеличивать значения каждого из критериев (при неизменных значениях остальных). В то же время в приложениях встречаются задачи, отличающиеся от задачи МКО, сформулированной выше, только тем, что значения всех критериев предпочтительно уменьшать. Наконец, часто встречаются задачи, в которых значения части критериев предпочтительно увеличивать, а остальных — уменьшать. Ясно, что с помощью умножения соответствующих целевых функций на минус единицу все эти задачи сводятся к задаче МКО, сформулированной выше. Хотя бинарное отношение Парето можно легко переформулировать и для случая минимизации всех критериев, и даже для случая разнонаправленных критериев, в дальнейшем ограничимся рассмотрением доминирования по Парето типа (2.2) и сформулированной выше задачи МКО, которую часто записывают в условном виде, аналогичном (1.1):

$$\varphi(x) \rightarrow \max, \quad x \in X \quad (2.3)$$

Надо отдавать себе отчет, что такая запись задачи МКО довольно условна (о доминировании по Парето в ней не говорится, хотя это и подразумевается), поэтому, в отличие от задачи скалярной оптимизации (1.1), она может не восприниматься людьми, мало знакомыми с МКО. К тому же, запись (1.1) сразу говорит о том, какая математическая задача должна быть решена, в то время как запись (2.3) об этом не говорит: она лишь дает информацию о множестве допустимых решений и о направлении улучшения значений критериев. Поэтому часто бывает необходимо не ограничиваться условной записью (2.3), а давать полную формулировку задачи МКО.

Кроме задачи МКО, существует много других задач выбора решения при нескольких критериях, не относящихся к области МКО. Это задачи с малым числом решений, задачи с качественными критериями, задачи с неопределенностями различных типов и т.д. Отметим, что в этих задачах набор частных критериев иногда также называют вектором, хотя свойств точки пространства $\mathbb{R}^m$ он уже не имеет. Такого распространения термина “вектор” на произвольные наборы критериев следует избегать (скажем, говорить просто о наборе частных критериев).

2.3. Понятие решения задачи МКО

В данном разделе вводится понятие решения задачи МКО. В сущности, проводимые рассуждения — это реализация общих идей решения фон Неймана-Моргенштерна для задачи МКО. Прежде всего, рассмотрим множество тех значений критериальных векторов, которые могут быть получены с использованием решений из множества X .

Определение 2.2. *Множество*

$$Y = \varphi(X) = \{y \in \mathbb{R}^m \mid y = \varphi(x), x \in X\}$$

называют множеством достижимых критериальных векторов (иногда — множеством достижимых значений критериев или множеством достижимых целей).

Множество Y в критериальном пространстве играет роль множества допустимых решений X в пространстве решений — оно содержит те и только те критериальные векторы, которые возможны (достижимы) согласно условиям задачи. Наличие бинарного отношения строгого порядка (доминирования по Парето) в пространстве $\mathbb{R}^m$ позволяет определить максимальные элементы Y .

Определение 2.3. Вектор $y^0 \in Y$, максимальный на Y по бинарному отношению Парето, называется оптимальным по Парето (недоминируемым по Парето, парето-оптимальным, парето-эффективным).

Заметим, что, поскольку при $m > 1$ бинарное отношение предпочтения Парето не является связным (т.е. в Y могут существовать векторы, не сравнимые с помощью этого отношения), то множество максимальных элементов, в общем случае, состоит более чем из одной точки множества Y .

Определение 2.4. Множество $\text{Max}_P Y$ максимальных векторов из Y по бинарному отношению Парето называют множеством парето-оптимальных критериальных векторов, множеством Парето, а также недоминируемой или эффективной по Парето границей (или просто границей Парето) множества Y и обозначают $P(Y)$.

Множество $P(Y)$ не всегда может претендовать на то, чтобы считаться решением задачи МКО. Так, можно легко построить примеры, когда множество $P(Y)$ не пусто, но его точки в совокупности не доминируют все точки множества достижимых критериальных векторов Y . В таком случае говорить о множестве $P(Y)$ как о решении задачи МКО было бы неверно.

Определение 2.5. Решение по фон Нейману-Моргенштерну для бинарного отношения Парето на множестве Y называется решением задачи МКО в критериальном пространстве.

В приведенном определении говорится о решении задачи МКО в критериальном пространстве. Существует и другое множество, претендующее на то, чтобы быть решением задачи МКО. Это множество пространства решений W . Прежде всего, отметим, что бинарное отношение Парето $\succ_P$, определенное в пространстве критериев, порождает бинарное отношение для $x', x'' \in W$, для которого будем использовать то же обозначение:

$$x' \succ_P x'' \Leftrightarrow \varphi(x') \succ_P \varphi(x'').$$

Это бинарное отношение, также называемое доминированием по Парето, является отношением строгого порядка в пространстве решений W . Поэтому можно рассмотреть множество максимальных элементов на множестве X согласно бинарному отношению Парето в пространстве решений.

Такое множество принято обозначать через $P(X)$. Множество $P(X)$ называют множеством решений, *эффективных по Парето* и *оптимальных по Парето*, а также множеством неуплучшаемых решений.

Определение 2.6. *Решение по фон Нейману-Моргенштерну для бинарного отношения Парето на множестве X называется решением задачи МКО в пространстве решений W .*

Таким образом, задача МКО имеет решение и в критериальном пространстве, и в пространстве решений. Эти решения связаны между собой: как легко заметить, $P(Y) = \varphi(P(X))$, причем $P(X)$ — полный прообраз $P(Y)$ на множестве X . Таким образом, связь между множествами $P(X)$ и $P(Y)$ в задаче МКО аналогична связи между множеством решений и оптимальным значением критерия в однокритериальной задаче оптимизации. В отличие от скалярного случая, в задачах МКО множество $P(Y)$ является сложным объектом исследования, свойства которого представляют интерес, а их знание, как мы увидим в этом курсе, помогает решить теоретические проблемы МКО.

Важно, что знание множества $P(Y)$ помогает решить практические вопросы выбора решения. Действительно, множество $P(Y)$ имеет простую геометрическую интерпретацию: оно является частью границы множества Y . Можно легко определить графически, принадлежит ли данная точка $y \in Y$ множеству $P(Y)$. Для этого достаточно выяснить, попадают ли другие точки множества Y в конус $y + \mathbb{R}_+^m \setminus \{0\}$ (см. рис. 2.1).

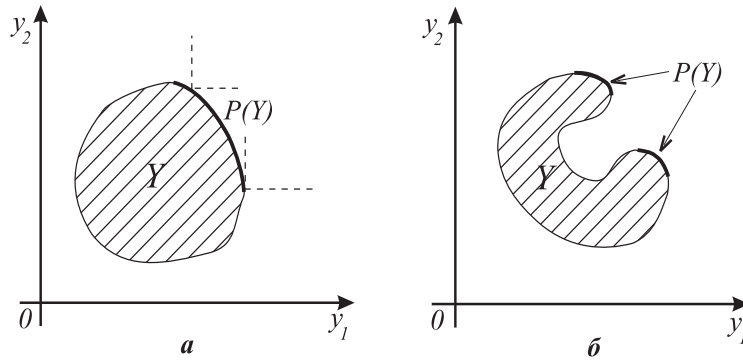


Рис. 2.1.

В то же время, в пространстве решений дать такую простую интерпретацию оптимального решения, как правило, невозможно из-за большой размерности пространства решений, что может осложниться нелинейностью отображения φ . В связи с этим, в методах МКО большое внимание уделяется множеству $P(Y)$. Надо, однако, отдавать себе отчет в том, что множество Y заранее не известно, поэтому описанный выше графический анализ границы Парето требует предварительного решения сложной задачи аппроксимации множества Y (или других, связанным с ним множеств).

Реализация такого подхода к решению задач МКО описано во второй главе нашей книги.

После того, как определены множества максимальных элементов, возникает вопрос об изучении их свойств. В данном курсе рассматриваются лишь те из них, которые играют существенную роль в дальнейшем изложении (о других свойствах, которые здесь не рассматриваются, см. [12]). Сразу отметим, что, например, даже у связных множеств Y граница Парето может быть несвязной (см. рис. 2.1 б).

Как всегда, стоит вопрос о существовании решения задачи МКО. Ответ на такой вопрос бывает непросто и в задачах однокритериальной оптимизации (скажем, в задачах, описываемых дифференциальными уравнениями), а в задачах МКО он осложняется их спецификой. Некоторые простые утверждения о существовании решения задачи МКО рассмотрены в следующем разделе.

2.4. Достаточные условия существования решения задачи МКО

В самом простом случае конечного множества X и, значит, конечного множества Y , из утверждения 1.4 для бинарного отношения Парето $\mathcal{P}$, являющегося отношением строгого порядка, получаем, что множество максимальных элементов $P(Y)$ непусто и является решением по фон Нейману-Моргенштерну для отношения $\mathcal{P}$ на Y . Таким образом, $P(Y)$ является решением задачи МКО в пространстве критериев. Совокупность точек множества X , порождающих точки множества $P(Y)$, т.е. множество $P(X)$, является решением задачи МКО в пространстве W .

В общем случае, однако, множество допустимых решений X и множество достижимых критериальных векторов Y бесконечны, поэтому для доказательства непустоты множеств максимальных элементов и выполнения свойства фон Неймана-Моргенштерна требуются определенные предположения.

Заранее обратим внимание на следующий факт. Напомним, что согласно определению свойства фон Неймана-Моргенштерна, для доказательства наличия этого свойства надо показать, что для любого $y^0 \in Y$ либо имеет место $y^0 \in P(Y)$, либо существует такой $y' \in P(Y)$, что $y' \succ_P y^0$. Поэтому для непустого множества Y из выполнения свойства фон Неймана-Моргенштерна сразу следует непустота $P(Y)$, так что последнее утверждение отдельно доказывать не требуется. Вспоминая, что $y' \succ_P y^0$ означает $y' \geq y^0$, $y' \neq y^0$, сразу получаем, что свойство фон Неймана-Моргенштерна для бинарного отношения Парето эквивалентно тому, что для любого $y^0 \in Y$ существует $y' \in P(Y)$, для которого $y' \geq y^0$.

Далее в этом разделе будем считать, что в $\mathbb{R}^m$ задано скалярное произведение $\langle y', y'' \rangle = \sum_{i=1}^n y'_i y''_i$ и соответствующие норма и расстояние.

Теорема 2.1. Пусть в задаче МКО (2.3) множество Y непусто, причем для любого $y \in Y$ множество $Y \cap \{y + \mathbb{R}_+^m\}$ компактно. Тогда множество

$P(Y)$ непусто и является решением задачи МКО.

Доказательство. Рассмотрим произвольную точку $y^0 \in Y$. Пусть $c \in \mathbb{R}^m$, $c > 0$. Рассмотрим задачу поиска $\max_{y \in Y \cap \{y^0 + \mathbb{R}_+^m\}} \langle c, y \rangle$. Поскольку функция $\langle c, y \rangle$ непрерывна, и по условию теоремы множество $Y \cap \{y + \mathbb{R}_+^m\}$ компактно, то максимум достигается. Пусть $\tilde{y}$ является точкой максимума, т.е.

$$\langle c, \tilde{y} \rangle = \max_{y \in Y \cap \{y^0 + \mathbb{R}_+^m\}} \langle c, y \rangle.$$

Покажем, что $\tilde{y} \in P(Y)$. Пусть это не так, т.е. найдется такой $\tilde{\tilde{y}} \in Y$, что $\tilde{\tilde{y}} \geq \tilde{y}$, $\tilde{\tilde{y}} \neq \tilde{y}$. Тогда в силу $c > 0$ имеем $\langle c, \tilde{\tilde{y}} \rangle > \langle c, \tilde{y} \rangle$, и при этом

$$\tilde{\tilde{y}} \in \{\tilde{y} + \mathbb{R}_+^m\} \subseteq \{y^0 + \mathbb{R}_+^m\}.$$

Получаем противоречие с тем, что $\tilde{y}$ является точкой максимума. Следовательно, $\tilde{y} \in P(Y)$. Таким образом, для произвольной точки $y^0 \in Y$ нашлась такая точка $\tilde{y} \in P(Y)$, что $\tilde{y} \geq y^0$. $\square$

Следствием теоремы 2.1 является нижеследующее простое утверждение.

Теорема 2.2. Пусть в задаче МКО (2.3) множество Y непусто и компактно. Тогда множество $P(Y)$ непусто и является решением задачи МКО.

Утверждение теоремы 2.1 не слишком конструктивно, поскольку возникает вопрос о том, как проверить компактность множества $\{y + \mathbb{R}_+^m\}$ для каждой точки $y \in Y$. Более конструктивной является следующая теорема.

Теорема 2.3. Пусть в задаче МКО (2.3) множество Y непусто и замкнуто, причем найдутся такие вектор $\mu > 0$ и число α , что для любого $y \in Y$ выполняется $\langle \mu, y \rangle \leq \alpha$, т.е. $Y \subseteq \{y \in \mathbb{R}^m \mid \langle \mu, y \rangle \leq \alpha\}$. Тогда множество $P(Y)$ непусто и является решением задачи МКО.

Утверждение этой теоремы может быть получено из теоремы 2.1 (подробнее см. в [12]).

Глава 3. Некоторые понятия МКО

Итак, в задаче многокритериальной оптимизации (2.3) под оптимальным решением понимаются связанные между собой множества в пространстве решений и пространстве критериев. Таким образом, рассуждения могут проводиться в одном из этих двух пространств. Теоремы о существовании решения задачи МКО, приведенные в предыдущем параграфе, показывают, что рассуждения, проводимые в критериальном пространстве, зачастую оказываются наглядными и простыми. Они позволяют отвлечься от особенностей, которые неизбежно проявляются при рассмотрении задач МКО в пространстве решений, и получить более универсальные утверждения. Понятия МКО, рассмотренные в данном параграфе, подтверждают наличие этой особенности задач МКО.

3.1. Применение квазипорядка в МКО

Рассмотрим в пространстве $\mathbb{R}^m$ квазипорядок $\mathcal{R}$, породивший бинарное отношение Парето, т.е.

$$y' \mathcal{R} y'' \iff y'_j \geq y''_j, \quad j \in \{1, \dots, m\},$$

Спрашивается, не может ли этот квазипорядок быть использован для формулировки понятия решения задачи МКО. Для ответа на этот вопрос введем некоторые понятия.

Определение 3.1. Элемент $x^* \in X$ называется абсолютно оптимальным решением задачи МКО (2.3), если $\varphi(x^*) \mathcal{R} \varphi(x)$ для всех $x \in X$, или, что то же самое, $\varphi(x^*) \geq \varphi(x)$ для всех $x \in X$.

Отметим, что даже в случае непрерывных вектор-функций $\varphi(x)$ и компактных множеств X такая точка $x^* \in X$ обычно не существует. Действительно, для ее существования требуется, чтобы максимум $\varphi_i(x)$ на X для всех $i = 1, \dots, m$ достигался в одной и той же точке x^* . Простейший пример отсутствия точки x^* : $X = \{(0, 1); (1, 0)\}$ и $\varphi_i(x) = x_i$, $i = 1, 2$. Эта пара допустимых точек не состоит в отношении “не хуже”, так что абсолютно оптимального решения не существует. Таким образом, непосредственное использование квазипорядка $\mathcal{R}$ непродуктивно. В то же время, никто не мешает нам рассмотреть максимальные значения критериев сами по себе (они существуют, в частности, для непрерывной $\varphi(x)$ и компактного множества X).

Определение 3.2. Под идеальной точкой задачи МКО (2.3) понимают такой вектор $y^* \in \mathbb{R}^m$, компоненты которого являются максимумами частных критериальных функций $\varphi_i(x)$ по отдельности, т.е.

$$y_i^* = \max_{x \in X} \varphi_i(x), \quad i = 1, \dots, m.$$

Вектор y^* (идеальная точка) является достижимым только в том случае, когда существует абсолютно оптимальное решение x^* . Поскольку наличие такого $x^* \in X$ является исключением, точку y^* часто называют утопической. На рис. 3.1 для $m = 2$ приведены примеры множеств Y , в которых вектор y^* является достижимым (рис. 3.1 а) и недостижимым (рис. 3.1 б).

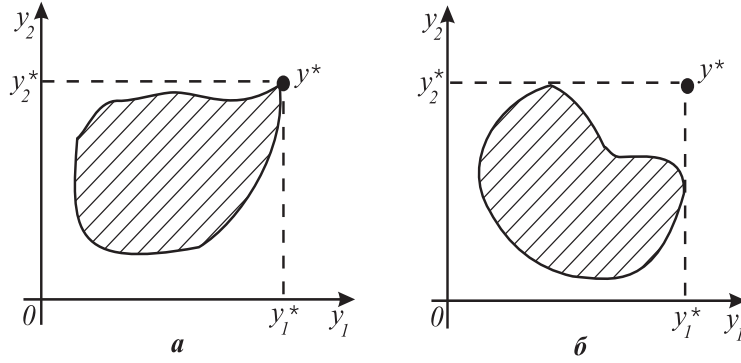


Рис. 3.1.

Таким образом, непосредственное использование квазипорядка $\mathcal{R}$ обычно не дает возможности ввести понятие решения задачи МКО, поэтому используется подход, основанный на использовании строгого порядка по Парето.

3.2. Оболочка Эджворта–Парето

В данном параграфе в дальнейшем без дополнительного упоминания рассматривается задача МКО (2.3).

Определение 3.3. Оболочкой Эджворта-Парето (ОЭП) множества Y называют множество

$$Y_P = Y + (-\mathbb{R}_+^m) = \{y \in \mathbb{R}^m | y \leq \varphi(x), x \in X\}.$$

Из определения 3.3 сразу следует

Лемма 3.1. Пусть задан произвольный вектор $y \in \mathbb{R}^m$. Тогда для того, чтобы $y \in Y_P$, необходимо и достаточно, чтобы существовал такой вектор $y' \in Y$, что $y' \geq y$.

Таким образом, оболочка Эджворта-Парето содержит все те точки $\mathbb{R}^m$, которые доминируются по Парето точками множества Y . Имеет место еще одно очевидное утверждение.

Лемма 3.2. Если в задаче МКО (2.3) бинарное отношение Парето обладает НМ-свойством на $P(Y) \subset Y$, то $Y_P = P(Y) + (-\mathbb{R}_+^m)$.

Термин *Оболочка Эджворта-Парето* является переводом с английского термина *Edgeworth-Pareto Hull*. Зачастую вместо Edgeworth-Pareto Hull в англоязычной литературе используется термин *Free Disposal Hull*, взятый из экономических исследований. Надо отметить, что хотя математическая формулировка обоих понятий близка, такая подмена неудачна, поскольку ОЭП имеет иной смысл.

Роль ОЭП в задачах МКО поясняет следующее утверждение.

Теорема 3.1. $P(Y_P) = P(Y)$.

Доказательство. 1. Докажем, что $P(Y) \subseteq P(Y_P)$. Пусть это не так, т.е. существует $y^0 \in P(Y)$ такой, что $y^0 \notin P(Y_P)$. Так как $P(Y) \subseteq Y \subseteq Y_P$, то $y^0 \in Y_P$. Поскольку $y^0 \notin P(Y_P)$, то найдется такой $y' \in Y_P$, что $y' \geq y^0$, $y' \neq y^0$. По лемме 3.1 это значит, что существует $y'' \in Y$ такой, что $y'' \geq y'$ и, следовательно, $y'' \geq y^0$, $y'' \neq y^0$. Поскольку $y'' \in Y$ доминирует по Парето вектор y^0 , то $y^0 \notin P(Y)$. Получили противоречие.

2. Докажем теперь обратное включение, т.е. $P(Y_P) \subseteq P(Y)$. Пусть $y^0 \in P(Y_P)$, но $y^0 \notin P(Y)$. Рассмотрим два случая.

2а. Если $y^0 \in Y$, то в силу $y^0 \notin P(Y)$, существует $y' \in Y$ такой, что $y' \geq y^0$, $y' \neq y^0$. Так как $Y \subseteq Y_P$, то $y' \in Y_P$, откуда $y^0 \notin P(Y_P)$. Таким образом, мы пришли к противоречию.

2б. Остается рассмотреть случай $y^0 \notin Y$. Так как $y^0 \in P(Y_P)$, то $y^0 \in Y_P$, значит, по лемме 3.1 найдется такой $y'' \in Y$, что $y'' \geq y^0$. Кроме того, поскольку $y^0 \notin Y$, то $y'' \neq y^0$, так что $y'' \in Y \subset Y_P$ доминирует $y^0 \in Y_P$. Отсюда $y^0 \notin P(Y_P)$. Опять получено противоречие. $\square$

Можно получить и другие свойства ОЭП, например, следующее.

Теорема 3.2. Если Y компактно, то Y_P замкнуто.

Доказательство. Поскольку Y замкнуто и ограничено, а множество $\mathbb{R}_+^m$ замкнуто, то утверждение теоремы сразу следует из замкнутости суммы двух замкнутых множеств, одно из которых ограничено [4], [7]. $\square$

Имеются и другие свойства ОЭП, важные с точки зрения методов МКО. Понятие ОЭП полезно и в теоретических исследованиях: использование этого понятия позволяет удобно сформулировать многие утверждения, в частности, рассмотренные далее в данном параграфе.

3.3. Доминирование по Слейтеру

В МКО используется еще одно бинарное отношение строгого порядка, играющее вспомогательную, но важную роль. Рассмотрим в пространстве $\mathbb{R}^m$ бинарное отношение $\mathcal{S}$:

$$y' \mathcal{S} y'' \Leftrightarrow y' > y'' \quad (y'_i > y''_i, \quad i = 1, \dots, m).$$

Легко заметить, что это бинарное отношение транзитивно и иррефлексивно, т.е. является бинарным отношением строгого порядка.

Для строгого порядка $\mathcal{S}$ также используется обозначение $\succ_S$. Это бинарное отношение принято называть доминированием по Слейтеру. Если $y' \succ_S y''$, то говорят, что вектор y' доминирует вектор y'' по Слейтеру. Строгий порядок по Слейтеру задается конусом $\text{int } \mathbb{R}^m = \mathbb{R}_+^m \setminus \text{Fr } \mathbb{R}_+^{m-1}$ ¹⁾. Множество векторов, доминирующих точку y по Слейтеру, можно представить в виде $y + \mathbb{R}_+^m \setminus \text{Fr } \mathbb{R}_+^m$.

Определение 3.4. Вектор $y^0 \in Y$, максимальный по отношению $\mathcal{S}$ на Y , называется оптимальным по Слейтеру, недоминируемым по Слейтеру или слабо эффективным.

Поскольку при $m > 1$ отношение предпочтения $\mathcal{S}$ не является связным (т.е. в Y могут существовать векторы, не сравнимые с помощью этого отношения), то множество максимальных элементов может, в общем случае, состоять более чем из одной точки множества Y .

Определение 3.5. Множество $S(Y) = \text{Max}_S Y$ называют множеством Слейтера в пространстве критериев или слабо эффективной границей²⁾ множества Y .

В последнее время для $S(Y)$ также появились такие термины как слабая граница Парето и слабое парето-оптимальное множество.

В силу того, что $S \subseteq \mathcal{P}$, из утверждения 1.3 следует, что $P(Y) \subseteq S(Y)$.

При рассмотрении задачи МКО бинарное отношение $\succ_S$ в пространстве критериев порождает соответствующее отношение на множестве допустимых решений X :

$$x' \succ_S x'' \Leftrightarrow \varphi(x') \succ_S \varphi(x'').$$

Через $S(X)$ обозначается множество максимальных элементов на множестве X . Множество $S(X)$ называют множеством решений, оптимальных по Слейтеру, а также множеством слабо эффективных решений³⁾.

Замечание 3.1.1. Поскольку $P(Y) \subseteq S(Y)$, теоремы 2.1 и 2.2 дают достаточное условие непустоты $S(Y)$ и выполнения НМ-свойства для бинарного отношения Слейтера на $S(Y) \subseteq Y$.

3.4. Свойства множества Слейтера

Напомним, что через $\text{Fr } Y$ обозначена граница множества $Y \subset \mathbb{R}^m$. Заметим, что утверждение $S(Y_P) = \text{Fr } Y_P$ выполняется далеко не всегда, например, оно не выполняется для незамкнутых множеств Y .

Теорема 3.3. Справедливы следующие соотношения:

$$1. S(Y) = S(Y_P) \cap Y,$$

¹⁾Напомним, что через $\text{int } X$ обозначается внутренность множества X , а через $\text{Fr } X$ — граница множества X .

²⁾Иногда под слабо эффективной границей понимают только $S(Y) \setminus P(Y)$.

³⁾Иногда под множеством слабо эффективных решений понимают только $S(X) \setminus P(X)$.

$$2. S(Y) = \text{Fr } Y_P \cap Y.$$

Доказательство. 1. Покажем, что $S(Y) \subseteq S(Y_P) \cap Y$. Предположим, что это не так, т.е. существует такой вектор $y^0 \in S(Y)$, что $y^0 \notin S(Y_P) \cap Y$. Поскольку $y^0 \in S(Y) \subseteq Y$, то из предположения следует, что $y^0 \notin S(Y_P)$. В силу $Y \subseteq Y_P$ имеем $y^0 \in Y_P$. Поскольку $y^0 \notin S(Y_P)$, то существует $y' \in Y_P$ такой, что $y' > y^0$. Применяя лемму 3.1 для y' , получаем, что существует $y'' \in Y$ такой, что $y'' \geq y'$, откуда, $y'' > y^0$, т.е. $y^0 \notin S(Y)$. Получили противоречие.

Покажем теперь, что $S(Y_P) \cap Y \subseteq S(Y)$. Пусть $y^0 \in S(Y_P) \cap Y$, но $y^0 \notin S(Y)$. Тогда найдется $y' \in Y$ такой, что $y' > y^0$. Но так как $Y \subseteq Y_P$, то $y' \in Y_P$, следовательно, $y^0 \notin S(Y_P)$. Получили противоречие.

2. Покажем, что $S(Y) \subseteq \text{Fr } Y_P \cap Y$. Предположим, что это неверно. Тогда найдется такой $y^0 \in S(Y)$, что $y^0 \notin \text{Fr } Y_P \cap Y$. Поскольку $y^0 \in S(Y) \subseteq Y$, то из предположения следует, что $y^0 \notin \text{Fr } Y_P$. Так как $y^0 \notin \text{Fr } Y_P$ и $y^0 \in Y_P$, то $y^0 \in \text{int } Y_P$, т.е. для некоторого $\varepsilon > 0$ имеем $U_\varepsilon(y^0) \subset Y_P$, где $U_\varepsilon(y^0)$ — ε -окрестность точки y^0 . Следовательно, существует $y' \in Y_P$ такой, что $y' > y^0$. Согласно лемме 3.1 найдется $y'' \in Y$ такой, что $y'' \geq y' > y^0$ и, следовательно, $y^0 \notin S(Y)$. Получили противоречие.

Покажем теперь, что $\text{Fr } Y_P \cap Y \subseteq S(Y)$. Пусть $y^0 \in Y \cap \text{Fr } Y_P$, но $y^0 \notin S(Y)$. Последнее означает, что существует $y' \in Y$ такой, что $y' > y^0$. Следовательно, найдется $\varepsilon > 0$ такое, что для всех $y \in U_\varepsilon(y^0)$ также имеем $y' > y$, т.е. $U_\varepsilon(y^0) \subset Y_P$, откуда $y^0 \in \text{int } Y_P$. Это противоречит тому, что $y^0 \in \text{Fr } Y_P$. $\square$

3.5. Оптимальность по Джоффриону

Рассмотрим простой пример задачи МКО. Пусть множество достижимых критериальных векторов имеет вид $Y = \{y \in \mathbb{R}^2 \mid (1 + y_1)^2 + y_2^2 \leq 1\}$ (см. рис. 3.2). Рассмотрим задачу многокритериальной максимизации (2.3).

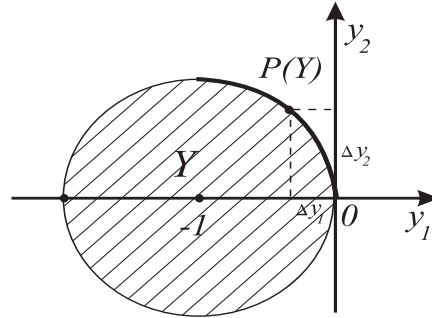


Рис. 3.2.

Множество критериальных точек, оптимальных по Парето, имеет вид $P(Y) = \{y \in \mathbb{R}^2 \mid -1 \leq y_1 \leq 0, y_2 = \sqrt{-y_1}\}$. Как видно, точка $(0,0)$ принадлежит паретовой границе множества Y , в то же время, она качественно

отличается от остальных точек. Производная функции $y_2 = \sqrt{-y_1}$ во внутренних точках области задания $-1 \leq y_1 \leq 0$ равна $-1/(2\sqrt{-y_1})$, и при стремлении y_1 к нулю величина производной стремится к $-\infty$. Это означает, что, переходя из точки $(0, 0)$ в некоторую близкую к ней точку паретовой границы, можно увеличить значение второго критерия на некоторую величину Δy_2 первого порядка малости, в то время как уменьшение значения первого критерия Δy_1 будет иметь второй порядок малости. Ясно, что точка $(0, 0)$ менее предпочтительна, чем такая точка $(\Delta y_1, \Delta y_2)$ (исключая патологический случай, когда критерий y_1 несравнимо более важен, нежели критерий y_2). Поэтому точка $(0, 0)$ может быть исключена из рассмотрения. Для того, чтобы при решении задачи МКО не рассматривать точки такого типа, вводится понятие оптимальности по Джоффриону.

Определение 3.6. Точка $y^0 \in P(Y)$ называется *собственно эффективной* (оптимальной по Джоффриону), если существует такое положительное число $\theta(y^0)$, что для всех точек $y \in Y$ справедливо следующее: если $y_i > y_i^0$, то найдется такой номер $j \in \{1, \dots, t\}$, что $y_j < y_j^0$, причем

$$\frac{y_i - y_i^0}{y_j^0 - y_j} \leq \theta(y^0).$$

В приведенном выше примере для точки $y^0 = (0, 0)$ такого $\theta(y^0) > 0$ не существует, поэтому она не является собственно эффективной.

Совокупность точек, оптимальных по Джоффриону, принято обозначать через $G(Y)$. По определению, $G(Y) \subseteq P(Y)$. В то же время можно показать, что в условиях теоремы 2.3 выполнено включение $P(Y) \subseteq \overline{G(Y)}$ (доказательство см. на стр. 145 книги [12]). Таким образом, в задачах, удовлетворяющих условию теоремы 2.3, множество $G(Y)$ мало отличается от $P(Y)$.

Глава 4. Свертки критериев

Ранее было определено понятие решения задачи многокритериальной оптимизации (2.3), в качестве которого были предложены множество $P(X)$ и его образ в критериальном пространстве $P(Y)$, и были рассмотрены простейшие свойства этих множеств. Далее продолжим изучение задачи МКО (2.3). Важнейшим теоретическим и практическим вопросом является вопрос о характеристике этих множеств на основе решения задачи скалярной оптимизации, т.е. их описания в виде совокупности решений некоторых задач скалярной оптимизации. Для характеристики этих множеств, а также для нахождения их отдельных точек был предложен *метод сверток критериев*, состоящий в построении скалярных функций вектора критериев и решении соответствующих задач скалярной оптимизации. Изначально метод сверток критериев предназначался для формирования единого критерия выбора решения, отражающего интересы пользователя. В настоящее время свертки критериев в теории МКО не связываются с интересами: они используются для того, чтобы сформулировать условия оптимальности, а также применяются при разработке методов решения задач МКО.

4.1. Скалярные функции, не убывающие по бинарному отношению

Под сверткой критериев в МКО принято понимать любую скалярную функцию критериев. Для установления свойств сверток критериев отвлечемся от задач МКО и рассмотрим свойства скалярных функций, не убывающих по бинарному отношению строгого порядка.

Пусть G — некоторое множество, на котором задано бинарное отношение строгого порядка $\mathcal{P}$ и пусть $\psi : G \rightarrow \mathbb{R}^1$. Пусть $a, b \in G$.

Определение 4.1. Функция ψ называется *неубывающей по $\mathcal{P}$* , если из $a\mathcal{P}b$ следует $\psi(a) \geq \psi(b)$. Функция ψ называется *возрастающей по $\mathcal{P}$* , если из $a\mathcal{P}b$ следует $\psi(a) > \psi(b)$.

Лемма 4.1 (Основная лемма). Пусть $A \subseteq G$, $\psi : G \rightarrow \mathbb{R}^1$, ψ не убывает по $\mathcal{P}$ на A и для некоторого $a^0 \in A$ имеем $\psi(a^0) = \max_{a \in A} \psi(a)$. Для того чтобы $a^0 \in \text{Max}_{\mathcal{P}} A$, достаточно выполнения одного из следующих условий:

- 1) ψ — возрастающая по $\mathcal{P}$ на A ,
- 2) a^0 — единственный элемент множества A , на котором достигается $\max_{a \in A} \psi(a)$.

Доказательство. 1. Предположим, что $a^0 \notin \text{Max}_{\mathcal{P}} A$. Это значит, что существует $a' \in A$ такой, что $a'\mathcal{P}a^0$, т.е. по определению 4.1 имеем $\psi(a') > \psi(a^0)$, но это противоречит тому, что $\psi(a^0) = \max_{a \in A} \psi(a)$.

2. Предположим, что $a^0 \notin \text{Мах}_P A$. Это значит, что найдется такой $a \in A$, что $a' \mathcal{P} a^0$ для некоторого $a' \in A$. Тогда $\psi(a') \geq \psi(a^0)$. Поскольку a^0 — единственный элемент множества A , на котором достигается $\max_{a \in A} \psi(a)$, то получили противоречие. $\square$

4.2. Общая теория сверток

Вернемся к задаче МКО (2.3).

Следствия из основной леммы:

- 1) если свертка критериев $\psi(y)$ является возрастающей по бинарному отношению Парето (Слейтера) на множестве Y , $y^0 \in Y$ и $\psi(y^0) = \max_{y \in Y} \psi(y)$, то $y^0 \in P(Y)$ ($y^0 \in S(Y)$);
- 2) если свертка критериев $\psi(y)$ является неубывающей по бинарному отношению Парето на множестве Y и существует единственный $y^0 \in Y$ такой, что $\psi(y^0) = \max_{y \in Y} \psi(y)$, то $y^0 \in P(Y)$.

Таким образом, для того, чтобы доказать, что некоторый вектор критериального пространства является недоминируемым, достаточно построить подходящую функцию свертки, для которой он является точкой максимума на множестве Y . Далеко не всегда, однако, ясно, как это сделать.

Если поступить наоборот, т.е. найти максимум некоторой функции свертки, возрастающей по $\mathcal{P}$ (или по $\mathcal{S}$) на множестве Y , то можно найти одну недоминируемую точку. Для того чтобы таким образом можно было найти все множество недоминируемых точек, функцию свертки можно задать в параметрическом виде. Тогда, меняя параметры свертки, при выполнении определенных условий можно получить все множество Парето. Отметим сразу, что это удастся далеко не для всех сверток.

Заметим, что в исходной формулировке задачи МКО множество Y не задано в явном виде и его построение требует существенных усилий (см. об этом §14). Поэтому при оптимизации сверток критериев на самом деле решаются несколько другие задачи оптимизации: вместо нахождения такого $y^0 \in Y$, что $\psi(y^0) = \max_{y \in Y} \psi(y)$, находится такой $x^0 \in X$, что $\psi(\varphi(x^0)) = \max_{x \in X} \psi(\varphi(x))$, а уже по x^0 находится $y^0 = \varphi(x^0)$. Это, однако, не меняет суть проблемы, так что в дальнейшем мы будем говорить о максимизации сверток в пространстве критериев.

Наиболее важными свертками критериев являются

- 1) *линейная свертка* $\psi_c(y) = \sum_{i=1}^m c_i y_i = \langle c, y \rangle$, где $c_i \geq 0$ — параметры свертки, $i = 1, \dots, m$;
- 2) *свертка Гермейера* $\psi_G(y) = \min_i (y_i / a_i)$, где $a_i > 0$ — параметры свертки, $i = 1, \dots, m$;
- 3) *свертки на основе отклонения от идеальной точки* $\psi_\rho(y) = -\rho(y^*, y)$, где y^* — идеальная точка, $\rho(\cdot, \cdot)$ — расстояние в $\mathbb{R}^m$.

4.3. Линейная свертка

Легко проверяются следующие утверждения.

Лемма 4.2. *Линейная свертка $\psi_c(y) = \langle c, y \rangle$ при $c > 0$ является возрастающей по бинарным отношениям Парето и Слейтера.*

Лемма 4.3. *Линейная свертка $\psi_c(y) = \langle c, y \rangle$ при $c \geq 0, c \neq 0$, является неубывающей по бинарному отношению Парето и возрастающей по бинарному отношению Слейтера.*

Из следствия основной леммы и лемм 4.2 и 4.3 следуют утверждения о недоминируемости по Парето (Слейтеру) точки оптимума линейной свертки.

Следствие 4.1 (из лемм (4.1), (4.2)). *Пусть $c \in \mathbb{R}^m, c > 0$. Если найдется такой вектор $y^0 \in Y$, что $\langle c, y^0 \rangle = \max_{y \in Y} \langle c, y \rangle$, то $y^0 \in P(Y) \subseteq S(Y)$.*

Следствие 4.2 (из лемм (4.1), (4.3)). *Пусть $c \in \mathbb{R}^m, c \geq 0, c \neq 0$. Если найдется такой вектор $y^0 \in Y$, что $\langle c, y^0 \rangle = \max_{y \in Y} \langle c, y \rangle$,*

1. *то $y^0 \in S(Y)$;*
2. *если, кроме того, вектор y^0 — единственная точка максимума $\langle c, y \rangle$ на Y , то $y^0 \in P(Y)$.*

Следствие 4.3 (из лемм (4.1), (4.2)). *Пусть $y^0 \in Y$. Если найдется такой $c \in \mathbb{R}^m, c > 0$, что $\langle c, y^0 \rangle = \max_{y \in Y} \langle c, y \rangle$, то $y^0 \in P(Y) \subseteq S(Y)$.*

Таким образом, линейная свертка порождает точки множеств Парето и Слейтера.

Следствие 4.4 (из лемм (4.1), (4.3)). *Пусть $y^0 \in Y$. Если найдется такой $c \in \mathbb{R}^m$, что $\langle c, y^0 \rangle = \max_{y \in Y} \langle c, y \rangle$ и $c \geq 0, c \neq 0$, то*

1. *$y^0 \in S(Y)$;*
2. *если, кроме того, вектор y^0 — единственная точка максимума $\langle c, y \rangle$ на Y , то $y^0 \in P(Y)$.*

Сформулированные следствия являются достаточными условиями принадлежности точек $y \in Y$ множеств Парето или Слейтера. Подчеркнем, однако, что эти условия не являются необходимыми — не все точки множеств Парето и Слейтера могут быть получены с помощью линейных свертки критериев. Для понимания этого факта достаточно рассмотреть границу Парето невыпуклого множества Y (см. рис. 4.2 б).

4.4. Свертка Гермейера

Свертку, предложенную Ю.Б. Гермейером [3], записывают в одной из двух форм:

$$\psi_G(y) = \min_i (y_i/a_i), \text{ где } a_i > 0 \text{ — параметры свертки, } i = 1, \dots, m,$$

$$\psi_G(y) = \min_i (\mu_i y_i), \text{ где } \mu_i > 0 \text{ — параметры свертки, } i = 1, \dots, m.$$

Очевидно, что переход от первой формы ко второй осуществляется заменой $\mu_i = 1/a_i$.

Рассмотрим свертку Гермейера в случае двух критериев (y_1, y_2) . Пусть в точке A с координатами $y^A = (y_1^A, y_2^A)$ (см. рис. 4.1) при некоторых значениях параметров выполняются равенства $\mu_1 y_1^A = \mu_2 y_2^A = 1$, т.е. $y_1^A = 1/\mu_1$, $y_2^A = 1/\mu_2$. Поэтому $\psi_G(y^A) = 1$. Тогда на лучах, исходящих из точки A параллельно осям координат в направлении возрастания критерия, имеем

1. $\mu_1 y_1 \geq 1$, $\mu_2 y_2 = 1$ на горизонтальном луче,
2. $\mu_1 y_1 = 1$, $\mu_2 y_2 \geq 1$ на вертикальном луче.

Отсюда получаем, что на этих лучах $\psi_G(y) = 1$. Таким образом, линии уровня свертки Гермейера совпадают с границами конуса $\{y^A + \mathbb{R}_+^m\}$. Аналогичным образом, линии уровня $\psi_G = \gamma$ задаются вертикальными и горизонтальными лучами, выходящими из точки $\hat{y} = (\hat{y}_1, \hat{y}_2)$, где $\mu_1 \hat{y}_1 = \gamma$, $\mu_2 \hat{y}_2 = \gamma$. Таким образом, системы линий уровня свертки Гермейера $\psi_G(y) = \gamma$ — это границы конуса $\{\hat{y} + \mathbb{R}_+^m\}$, где $\hat{y} = \gamma y^A$.

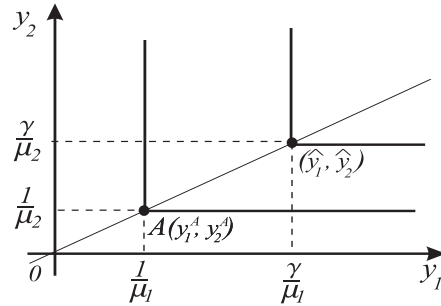


Рис. 4.1.

В общем случае m критериев поверхности уровня свертки Гермейера задаются границами конусов $\{\hat{y} + \mathbb{R}_+^m\}$, где $\hat{y}$ — любая точка, удовлетворяющая соотношению $\mu_i \hat{y}_i = \gamma$, $i = 1, \dots, m$, $\gamma \geq 0$. Таким образом, поверхности уровня свертки Гермейера — границы конуса точек, доминирующих вершину конуса. Благодаря этому свертка Гермейера обладает замечательными свойствами, которые доказаны ниже.

Лемма 4.4. *Свертка Гермейера является возрастающей по бинарному отношению Слейтера и неубывающей по бинарному отношению Парето.*

Доказательство. 1. Докажем, что свертка Гермейера является возрастающей по бинарному отношению Слейтера. Рассмотрим векторы $y'', y' \in \mathbb{R}^m$. Пусть $y'' > y'$, тогда справедливо представление $y'' = y' + e$, где $e = (e_1, \dots, e_m)$, $e_i > 0$, $i = 1, \dots, m$. В силу этого выполнены следующие равенства и неравенство: $\psi_G(y'') = \min_i (\mu_i y''_i) = \min_i (\mu_i (y'_i + e_i)) \geq$

$\min_i(\mu_i y') + \min_i(\mu_i e_i)$. Поскольку $\mu_i e_i > 0$, $i = 1, \dots, m$, то $\min_i(\mu_i e_i) > 0$, поэтому $\psi_G(y'') > \min_i(\mu_i y'_i) = \psi_G(y')$.

2. Докажем, что свертка Гермейера является неубывающей по бинарному отношению Парето. Рассмотрим векторы $y'', y' \in \mathbb{R}^m$. Пусть $y'' \geq y'$, $y'' \neq y'$. Тогда имеем $y'' = y' + e$, где $e = (e_1, \dots, e_m)$, $e_i \geq 0$, $i = 1, \dots, m$, $e \neq 0$. Отсюда $\psi_G(y'') = \min_i(\mu_i y''_i) = \min_i(\mu_i(y'_i + e_i)) \geq \min_i(\mu_i y'_i) + \min_i(\mu_i e_i)$. Поскольку $\mu_i > 0$, $i = 1, \dots, m$, то $\min_i(\mu_i e_i) \geq 0$, значит $\psi_G(y'') \geq \min_i(\mu_i y'_i) = \psi_G(y')$. $\square$

Из этой леммы сразу следует, что если y^0 — точка максимума $\psi_G(y)$ на Y , то $y^0 \in S(Y)$; если же она к тому же — единственная точка максимума, то $y^0 \in P(Y)$.

Лемма 4.5. Пусть $y^0 \in Y$, $y^0 > 0$. Тогда из того, что $y^0 \in S(Y)$ следует, что найдется такой набор параметров μ_i^0 , $\mu_i^0 > 0$, $i = 1, \dots, m$, что для функции $\psi_G^0(y) = \min_i \mu_i^0 y_i$ имеем $\psi_G^0(y^0) = \max_{y \in Y} \psi_G^0(y)$.

Заметим, что условие $y^0 > 0$ может быть удовлетворено с помощью переноса начала координат, так что это требование не ограничивает возможности применения леммы.

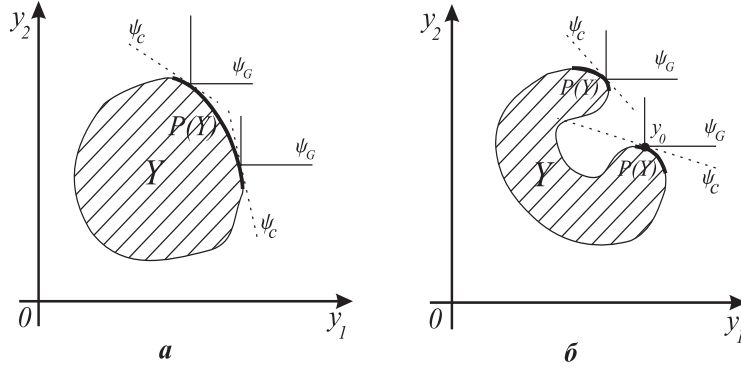
Доказательство. Из того, что $y^0 \in S(Y)$, следует, что не существует такого $y \in Y$, что $y > y^0$. Другими словами, для любого $y \in Y$ найдется координата i_0 такая, что $y_{i_0} \leq y_{i_0}^0$. Тогда $\min_i \frac{y_i}{y_{i_0}^0} \leq \frac{y_{i_0}}{y_{i_0}^0} \leq 1$. Выберем набор параметров μ_i^0 так, что $\mu_i^0 = \frac{1}{y_{i_0}^0}$, $i = 1, \dots, m$. Следовательно, для функции $\psi_G^0(y)$ с параметрами μ_i^0 имеем $\psi_G^0(y) = \min_i \frac{y_i}{y_{i_0}^0} \leq 1$. В то же время, $\psi_G^0(y_0) = 1$. Отсюда следует, что y_0 является точкой максимума свертки $\psi_G^0(y)$ с параметрами μ_i^0 , $i = 1, \dots, m$. $\square$

Следствием лемм 4.4 и 4.5 является следующая теорема.

Теорема 4.1 (Гермейер). Пусть $y^0 \in Y$, $y^0 > 0$. Включение $y^0 \in S(Y)$ верно тогда и только тогда, когда существует набор параметров $\mu_i^0 > 0$, $i = 1, \dots, m$, при котором для $\psi_G^0(y) = \min_i \mu_i^0 y_i$ имеем $\psi_G^0(y^0) = \max_{y \in Y} \psi_G^0(y)$.

Теорема Гермейера утверждает, что, в отличие от линейной свертки, свертка $\psi_G(y)$ позволяет сформулировать не только достаточное, но и необходимое условие принадлежности критериальной точки множеству Слейтера: каждая точка множества Слейтера (в том числе и точка множества Парето) может быть найдена как максимум $\psi_G(y)$ при соответствующем наборе параметров $\mu_i > 0$. Для лучшего понимания аппарата свертки изобразим линии уровня линейной свертки и свертки Гермейера в точках паретовой границы двух множеств (см. рис. 4.2). Видно, что если множество

является выпуклым (рис. 4.2 а), то в любой точке его паретовой границы с помощью любой из рассматриваемых сверток можно провести линию уровня, которая будет соответствовать максимальному допустимому значению свертки. Если же множество невыпукло (рис. 4.2 б), то это можно будет сделать только с помощью свертки Гермейера и нельзя с помощью линейной — линия уровня может оказаться только локально максимальной, как, например, в точке y_0 .



Об. 4.2.

4.5. Свертки на основе идеальной точки

Свертка на основе использования идеальной точки y^* имеет вид $\psi_\rho(y) = -\rho(y^*, y)$, где $\rho(\cdot, \cdot)$ — некоторая метрика в $\mathbb{R}^m$. При этом, если, как обычно, идеальная точка не достижима, то $\psi_\rho(y) < 0$ для всех векторов $y \in Y$. При построении этой свертки могут быть использованы различные метрики. Наиболее часто используется чебышевская метрика $\rho(y^*, y) = \max_i |y_i^* - y_i|$ или взвешенная чебышевская метрика $\rho(y^*, y) = \max_i \{\lambda_i |y_i^* - y_i|\}$, $\lambda_i > 0$. Может быть использована и любая другая метрика. При использовании чебышевской метрики свертка называется *чебышевской* и обозначается через $\psi_T(y)$. Ясно, что она является аналогом свертки Гермейера, только в качестве выделенной точки берется не нулевая, а идеальная точка.

Заметим, что поскольку $y_i \leq y_i^*$ для $y \in Y$, то при максимизации свертки по $y \in Y$ взвешенную чебышевскую метрику можно записать в более простом виде: $\psi_T(y) = -\max_i [\lambda_i (y_i^* - y_i)]$. Эта свертка, как легко доказать аналогично лемме 4.4, является возрастающей по бинарному отношению Слейтера и неубывающей по бинарному отношению Парето.

Для описания $S(Y)$ свертку $\psi_T(y)$ надо несколько преобразовать, используя вместо y^* точку y^{**} , $y_i^{**} = y_i^* + \delta$, $i = 1, \dots, m$, где δ — малое положительное число. Тогда для того, чтобы $y^0 \in S(Y)$, требуется, чтобы нашелся такой набор $\lambda_i^0 > 0$, $i = 1, \dots, m$, при котором y^0 — решение задачи

$$\min_{y \in Y} \psi_T(y),$$

где

$$\psi_T(y) = \max_i [\lambda_i^0 (y_i^{**} - y_i)].$$

4.6. Коррекция свертки критериев

На практике, когда ЛПР интересуют только точки, оптимальные по Парето, можно рассмотреть скорректированную свертку $\psi_\varepsilon(y) = \psi(y) + \varepsilon \langle c, y \rangle$, где $\psi(y)$ — некоторая свертка, не убывающая по бинарному отношению Парето, $c = (1, 1, \dots, 1)$, ε — малый положительный параметр. Благодаря этой небольшой коррекции, свертка $\psi_\varepsilon(y)$ становится возрастающей по бинарному отношению Парето. Например, взяв свертку $\psi_G^\varepsilon(y) = \psi_G(y) + \varepsilon \langle c, y \rangle$, получим свертку, максимум которой может достигаться только в точках паретовой границы. Недостатком такого подхода является то, что не все точки $P(Y)$ могут быть получены из-за возмущающего действия коррекции $\varepsilon \langle c, y \rangle$.

Глава 5. Условия оптимальности в МКО

В теории МКО большую роль играет линейная свертка. Причиной этого являются наличие экономической интерпретации параметров линейной свертки как цен, а также возможность построения теории двойственности на основе использования линейных сверток критериев. Часть утверждений о линейной свертке сразу следует из общей теории сверток, другие утверждения приведены в этом параграфе. Итак, рассматривается задача МКО (2.3).

5.1. Эффективно выпуклые задачи

Определение 5.1. *Задача многокритериальной оптимизации (2.3) называется эффективно выпуклой, если оболочка Эджворта-Парето выпукла.*

Множество Y в этом случае также иногда (не очень удачно) называют *эффективно выпуклым*, поскольку выпуклым оно может не являться. Рассмотрим достаточные условия эффективной выпуклости.

Определение 5.2. *Вектор-функция φ называется вогнутой на выпуклом множестве X , если для любых $x', x'' \in X$ и $\lambda \in [0, 1]$ справедливо $\varphi(\lambda x' + (1 - \lambda)x'') \geq \lambda\varphi(x') + (1 - \lambda)\varphi(x'')$.*

Лемма 5.1 (Карлин). *Пусть множество X выпукло и вектор-функция φ вогнута. Тогда множество Y_P является выпуклым.*

Доказательство. Пусть $y', y'' \in Y_P$. Требуется доказать, что $y^\lambda = \lambda y' + (1 - \lambda)y'' \in Y_P$ при любом $\lambda \in [0, 1]$. Из леммы 3.1 следует, что существуют такие $x', x'' \in X$, что $\varphi(x') \geq y'$, $\varphi(x'') \geq y''$, откуда $y^\lambda \leq \lambda\varphi(x') + (1 - \lambda)\varphi(x'')$. Выпуклость множества X означает, что $x^\lambda = \lambda x' + (1 - \lambda)x'' \in X$ для любого $\lambda \in [0, 1]$, поэтому $\varphi(x^\lambda) \in Y$. В силу вогнутости φ имеем $\varphi(x^\lambda) \geq \lambda\varphi(x') + (1 - \lambda)\varphi(x'') \geq y^\lambda$, откуда по лемме 3.1 получаем, что $y^\lambda \in Y_P$. $\square$

Это достаточное свойство эффективной выпуклости иногда используется в качестве ее определения, что, конечно, сужает область применимости рассматриваемой теории. Поэтому в дальнейшем будем придерживаться определения, сформулированного выше.

В дальнейшем нам потребуется известная из курса теории оптимизации (см., например, гл. 4 книги [2]) теорема об отделимости множеств (теорема об опорной гиперплоскости).

Теорема 5.1. *Пусть множество $Z \subset R^m$ выпукло. Тогда любая граничная точка y^0 множества Z отделима от Z , т.е. найдется такой вектор c , $\|c\| \neq 0$, что $\langle c, y^0 \rangle \geq \langle c, y \rangle$ для $y \in Z$.*

Сформулированная теорема является центральным пунктом доказательства основного утверждения теории эффективно-выпуклых задач.

Теорема 5.2 (По-лунг Ю). Пусть задача МКО эффективно выпукла. Пусть $y^0 \in Y$. Тогда для того, чтобы $y^0 \in S(Y)$, необходимо и достаточно существование такого $c \geq 0$, $c \neq 0$, что $\langle c, y^0 \rangle = \max_{y \in Y} \langle c, y \rangle$.

В данной книге теорема доказывается для случая, когда Y компактно (в общем случае доказательство приведено в [12]). Предварительно докажем одно полезное свойство множества Y_P .

Лемма 5.2. Пусть множество Y компактно. Тогда для любого $c \geq 0$ выполнено равенство $\max_{y \in Y_P} \langle c, y \rangle = \max_{y \in Y} \langle c, y \rangle$.

Доказательство. Согласно лемме 3.1 из того, что $y \in Y_P$ следует, что существует $y' \in Y$ такой, что $y' \geq y$. Тогда для любого $c \geq 0$ имеет место неравенство $\max_{y \in Y} \langle c, y \rangle \geq \max_{y \in Y_P} \langle c, y \rangle$. С другой стороны, из того, что $Y \subseteq Y_P$ следует неравенство в обратную сторону: $\max_{y \in Y} \langle c, y \rangle \leq \max_{y \in Y_P} \langle c, y \rangle$. Таким образом, утверждение леммы доказано. $\square$

Перейдем к доказательству теоремы.

Доказательство. Достаточность была доказана нами при рассмотрении свойств линейных сверток. Докажем необходимость. Пусть $y^0 \in S(Y)$. Согласно теореме 3.3 имеем $S(Y) = \text{Fr } Y_P \cap Y$, поэтому $y^0 \in \text{Fr } Y_P$. Из выпуклости и замкнутости Y_P , а также того, что $y^0 \in \text{Fr } Y_P$, с помощью теоремы об отделимости получаем, что найдется вектор c , $\|c\| \neq 0$, такой, что $\langle c, y^0 \rangle = \max_{y \in Y_P} \langle c, y \rangle$. Докажем, что все координаты вектора c неотрицательны.

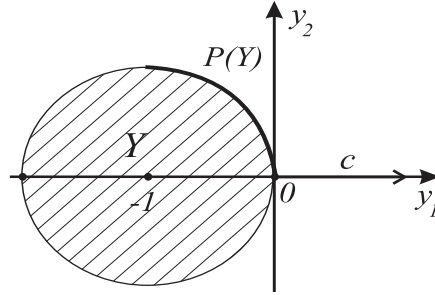
Предположим противное, т.е. пусть $c_{i_0} < 0$ для некоторого i_0 . По определению Y_P , при $c_{i_0} < 0$ это множество содержит все точки вида $\tilde{y} = y^0 + t(0, \dots, c_{i_0}, \dots, 0)$, где $t \geq 0$. Но тогда величина $\langle c, \tilde{y} \rangle = \langle c, y^0 \rangle + t(c_{i_0})^2$ не ограничена сверху на Y_P . Получили противоречие, поэтому $c_i \geq 0$ для всех i .

Поскольку по лемме 5.2 имеем $\max_{y \in Y_P} \langle c, y \rangle = \max_{y \in Y} \langle c, y \rangle$, то $\langle c, y^0 \rangle = \max_{y \in Y} \langle c, y \rangle$, где $c \geq 0$, $c \neq 0$. Теорема доказана. $\square$

Из теоремы По-лунг Ю и леммы Карлина следует свойство слейтеровских решений.

Теорема 5.3. Пусть X выпукло, φ вогнута, $x^0 \in X$. Тогда для того, чтобы $x^0 \in S(X)$ необходимо и достаточно существование такого $c \geq 0$, $c \neq 0$, что $\langle c, \varphi(x^0) \rangle = \max_{x \in X} \langle c, \varphi(x) \rangle$.

Отметим, что если в теореме 5.2 использовать условие $c > 0$ вместо $c \geq 0$, то заменить множество Слейтера на множество Парето нельзя. Приведем пример для случая $m = 2$ (см. рис. 5.1). Здесь точка $(0, 0)$ является оптимальной по Парето, а единственное опорное направление в этой точке определяет вектор $c = (1, 0)$.



Өйб. 5.1.

Однако использование строгого неравенства $c > 0$ отнюдь не бессмысленно — оно дает возможность описать множество Джоффриона.

Теорема 5.4 (Джоффрион). *Пусть X — выпуклое множество, вектор-функция φ вогнута, $x^0 \in X$. Тогда для того, чтобы $\varphi(x^0) \in G(Y)$, необходимо и достаточно существование такого $c > 0$, что $\langle c, \varphi(x^0) \rangle = \max_{x \in X} \langle c, \varphi(x) \rangle$.*

Доказательство теоремы можно найти в [12], и здесь мы его приводить не будем.

Укажем еще одно важное следствие леммы Карлина и теоремы По-лунг Ю. Рассмотрим задачу (2.3), в которой $D \subseteq \mathbb{R}^n$,

$$X = \{x \in D | g_j(x) \geq 0, j = 1, \dots, J\}. \quad (5.1)$$

Пусть, для начала, $m = 1$. В полученной однокритериальной задаче оптимизации критерий будем обозначать через $f(x)$. Для этой задачи известно следующее утверждение (см., например, гл. 4 книги [2]).

Теорема 5.5 (Кун, Таккер). *Пусть функции $f(x)$, $g_1(x), \dots, g_J(x)$ вогнуты на выпуклом множестве D и существует точка $\tilde{x} \in D$, в которой $g_j(\tilde{x}) > 0$, $j = 1, \dots, J$ (выполнено условие Слейтера). Тогда для того, чтобы точка x^0 была решением задачи $\max_{x \in X} f(x)$, необходимо и достаточно существование такого вектора λ^0 , $\lambda^0 \in \mathbb{R}_+^J$, что точка (x^0, λ^0) на множестве $D \times \mathbb{R}_+^J$ является седловой точкой функции Лагранжа*

$$L(x, \lambda) = f(x) + \sum_{i=1}^J \lambda_i g_i(x),$$

т.е. $L(x, \lambda^0) \leq L(x^0, \lambda^0) \leq L(x^0, \lambda)$ при $x \in D$ и $\lambda_j \geq 0$, $j = 1, \dots, J$.

Из теорем Ю и Куна-Таккера, также леммы Карлина сразу получаем следующее утверждение для задачи МКО с множеством X , задаваемым (5.1).

Теорема 5.6. Пусть множество D выпукло, функции $\varphi_i(x)$, $i = 1, \dots, m$, $g_j(x)$, $j = 1, \dots, J$, вогнуты и для множества X выполнено условие Слейтера. Тогда для того, чтобы $x^0 \in S(X)$, необходимо и достаточно существование таких векторов $c \geq 0$, $c \neq 0$, и $\lambda^0 \in \mathbb{R}_+^J$, что точка (x^0, λ^0) является седловой точкой функции Лагранжа

$$L = \langle c, \varphi(x) \rangle + \sum_{j=1}^J \lambda_j g_j(x)$$

на множестве $D \times \mathbb{R}_+^J$.

5.2. Невыпуклые задачи

Требование эффективной выпуклости упрощает исследование задачи МКО, но в то же время сильно ограничивает класс решаемых задач. Перейдем к изучению невыпуклого случая. Рассмотрим задачу (2.3) при X , имеющем вид (5.1). Теперь, однако, выпуклость D и вогнутость рассматриваемых функций не предполагается. В этом случае, по-прежнему, действуют достаточные условия оптимальности, полученные из свойств сверток критериев в §4. Найдем необходимые условия, которые должны отличаться от выпуклого случая.

Далее нам потребуется теорема Моцкина об альтернативе [12], которая следует из леммы Фаркаша — основного утверждения теории линейных неравенств [13]. Мы будем использовать теорему Моцкина в упрощенной форме:

либо система $A\mu > 0$ имеет решение, либо имеет решение система

$$uA = 0, u \geq 0, u \neq 0$$

Введем обозначение

$$J(x^0) = \{j \in \{1, \dots, J\} \mid g_j(x^0) = 0\}$$

Ограничения с номерами из совокупности $J(x^0)$ принято называть активными ограничениями в точке x^0 .

Теорема 5.7 (Да Кунья, Полак, Джоффрион).

Пусть $x^0 \in \text{int } D$, функции $\varphi_i(x)$, $i = 1, \dots, m$, и $g_j(x)$, $j = 1, \dots, J$, дифференцируемы в точке x^0 и выполняется следующее условие регулярности: существует $\tilde{x} \in \mathbb{R}^n$, для которого $\langle \text{grad } g_j(x^0), \tilde{x} \rangle > 0$ для любого $j \in J(x^0)$. Тогда из того, что $x^0 \in S(X)$ следует, что существуют такие $c \neq 0$, $c \in \mathbb{R}_+^m$, $\lambda^0 \in \mathbb{R}_+^J$, что

$$\sum_{i=1}^m c_i \text{grad } \varphi_i(x^0) + \sum_{j=1}^J \lambda_j \text{grad } g_j(x^0) = 0.$$

Доказательство. Рассмотрим вспомогательную систему

$$\begin{cases} \langle \text{grad } \varphi_i(x^0), x \rangle > 0, & i = 1, \dots, m, \\ \langle \text{grad } g_j(x^0), x \rangle > 0, & j \in J(x^0), \end{cases} \quad (5.2)$$

и покажем, что она не имеет решения. Предположим противное, т.е. пусть существует $\hat{x}$ — решение системы (5.2). Рассмотрим семейство точек $x_\varepsilon = x^0 + \varepsilon \hat{x}$, $\varepsilon > 0$. Заметим, что так как $x^0 \in \text{int } D$, то $x_\varepsilon \in D$ при достаточно малых ε . В силу дифференцируемости функций $\varphi_i(x)$, $i = 1, \dots, m$, в точке x^0 , приращение каждой из них можно представить в виде $\varphi_i(x_\varepsilon) - \varphi_i(x^0) = \varepsilon \langle \text{grad } \varphi_i(x^0), \hat{x} \rangle + o(\varepsilon)$. Так как первые m неравенств в системе (5.2) строгие, то при достаточно малых ε имеем

$$\varphi_i(x_\varepsilon) > \varphi_i(x^0). \quad (5.3)$$

Аналогично, в силу дифференцируемости функций $g_j(x)$, $j = 1, \dots, J$, в точке x^0 можно записать $g_j(x_\varepsilon) - g_j(x^0) = \varepsilon \langle \text{grad } g_j(x^0), \hat{x} \rangle + o(\varepsilon)$. Так как неравенства второй группы в системе (5.2) строгие, то при достаточно малых ε получим, что

$$g_j(x_\varepsilon) > g_j(x^0) = 0 \quad \text{для } j \in J(x^0). \quad (5.4)$$

В силу непрерывности функций $g_j(x)$, $j = 1, \dots, J$, при достаточно малых ε имеем

$$g_j(x_\varepsilon) > 0 \quad \text{для } j \notin J(x^0). \quad (5.5)$$

Но тогда неравенства (5.3)–(5.5) означают, что $x_\varepsilon \in X$ и $x_\varepsilon \succ_S x^0$, что противоречит тому, что $x^0 \in S(X)$. Следовательно, система (5.2) решения не имеет. Отсюда по теореме Моцкина об альтернативе должна существовать такая пара $(v, w) \geq 0$, $(v, w) \neq 0$, $v \in \mathbb{R}^m$, $w = (w_j \mid j \in J(x^0))$, что

$$\sum_{i=1}^m v_i \text{grad } \varphi_i(x^0) + \sum_{j \in J(x^0)} w_j \text{grad } g_j(x^0) = 0.$$

Предположим, что $v = 0$, тогда должен существовать такой вектор $w \geq 0$, $w \neq 0$, что

$$\sum_{j \in J(x^0)} w_j \text{grad } g_j(x^0) = 0.$$

Умножив последнее равенство на вектор $\tilde{x}$ из условия регулярности, получим, что

$$\sum_{j \in J(x^0)} w_j \langle \text{grad } g_j(x^0), \tilde{x} \rangle = 0.$$

Но тогда $w_j = 0, j \in J(x^0)$, что противоречит требованию $(v, w) \neq 0$, поскольку $v = 0$. Поэтому $v \neq 0$. Положив $c = v$ и

$$\lambda_j^0 = \begin{cases} w_j, & j \in J(x^0), \\ 0, & j \notin J(x^0), \end{cases}$$

получим требуемое утверждение. $\square$

Замечание 5.1. *Результат теоремы 5.7 представляет собой локальное необходимое условие слабой эффективности. Оно принципиальное отличается от глобального необходимого условия, полученного в теоремах 5.3 и 5.5. Используемая здесь линейная свертка $\langle c, y \rangle$ может не иметь в точке x_0 глобального максимума на X (см. точку y^0 на рис. 4.2 б).*

Глава 6. Анализ устойчивости в задачах МКО

Рассмотрим вопрос об устойчивости множеств Y , Y_P , $P(Y)$, $S(Y)$ по отношению к возмущениям множества X и отображения $\varphi(x)$ в задаче многокритериальной оптимизации

$$y \rightarrow \max, \quad y = \varphi(x), \quad x \in X \subset W.$$

Будем считать, что в $\mathbb{R}^m$ задана метрика $\rho_y(y, y')$. В этом случае можно ввести понятие устойчивости интересующих нас множеств: решение некоторой задачи принято называть *устойчивым*, если оно непрерывно зависит от исходных данных, данном случае от множества X и отображения $\varphi(x)$. Исследование устойчивости важно как с точки зрения постановки задачи (должен не быть смысла в построении множества, которое резко меняется при малом изменении исходных данных), так и с вычислительной точки зрения, поскольку использование компьютера связано с ошибками округления, которые могут привести к существенному отклонению аппроксимации от аппроксимируемого множества.

6.1. О понятии устойчивости в МКО

Для того чтобы сформулировать понятие устойчивости, будем считать, что множество допустимых решений X и отображение $\varphi(x)$ зависят от набора параметров u , принадлежащих пространству U . Будем считать, что задано отображение, ставящее в соответствие каждому набору параметров $u \in U$ множество $X(u)$. Вместо функции $\varphi : W \rightarrow \mathbb{R}^m$ рассмотрим функцию $\hat{\varphi} : W \times U \rightarrow \mathbb{R}^m$. Таким образом, задача многокритериальной оптимизации, зависящая от параметров, имеет вид

$$y \rightarrow \max, \quad y = \hat{\varphi}(x, u), \quad x \in X(u), \quad u \in U.$$

Интересующие нас множества Y , Y_P , $P(Y)$, $S(Y)$ теперь зависят от набора параметров u , обозначим их $Y(u)$, $Y_P(u)$, $P(Y(u))$, $S(Y(u))$. Пусть в пространстве U введена метрика, с помощью которой можно определить понятие окрестности точки $u \in U$. Тогда исследование устойчивости — это анализ непрерывности зависимости указанных множеств от параметров в окрестности некоторой точки $\hat{u} \in \text{int } U$, такой что

$$X = X(\hat{u}), \quad \varphi(x) = \hat{\varphi}(x, \hat{u}).$$

Таким образом, точка $\hat{u}$ задает невозмущенную задачу МКО. Поскольку $X(u)$, $Y(u)$ и $Y_P(u)$ задаются многозначными точечно-множественными отображениями, для изучения устойчивости необходимо определить понятие непрерывности многозначного отображения.

Рассмотрим отображение $F : R^p \rightarrow 2^{R^q}$, где R^p — линейное метрическое пространство, а 2^{R^q} — пространство подмножеств пространства R^q . Пусть в R^q и R^p заданы метрики $\rho_q(\cdot, \cdot)$ и $\rho_p(\cdot, \cdot)$.

Определение 6.1. *Отображение F называется полунепрерывным сверху в точке $\hat{x} \in R^p$, если для любой последовательности $\{x^k\}$ такой, что $x^k \rightarrow \hat{x}$ и любой такой последовательности $\{y^k\}$, $y^k \in F(x^k)$, что $y^k \rightarrow \hat{y}$, имеем $\hat{y} \in F(\hat{x})$.*

Определение 6.2. *Отображение F называется полунепрерывным снизу в точке $\hat{x} \in R^p$, если из того, что $x^k \rightarrow \hat{x}$ и $\hat{y} \in F(\hat{x})$, следует, что найдется такая последовательность $\{y^k\}$, что $y^k \in F(x^k)$ при достаточно больших k и $y^k \rightarrow \hat{y}$.*

Определение 6.3. *Отображение F называется непрерывным в точке $\hat{x} \in R^p$, если оно полунепрерывно сверху и снизу в этой точке.*

Определение 6.4. *Отображение F называется непрерывным в метрике Хаусдорфа, если*

$$h_y(F(x), F(\hat{x})) \rightarrow 0 \quad \text{при} \quad x \rightarrow \hat{x},$$

где $h_y(A_1, A_2)$ — метрика Хаусдорфа, т.е.

$$h_y(A_1, A_2) = \max\left\{\sup_{y \in A_1} \rho_y(y, A_2), \sup_{y \in A_2} \rho_y(y, A_1)\right\}.$$

Замечание 6.1. *Хотя определение непрерывности многозначного отображения F отличается от определения непрерывности в метрике Хаусдорфа, в случае компактных множеств в конечномерных пространствах эти определения совпадают.*

Для того чтобы были ясны особенности проблемы устойчивости в задаче МКО, рассмотрим устойчивость задачи скалярной оптимизации

$$y = f(x) \rightarrow \max, \quad x \in X.$$

Ее возмущение имеет вид

$$y = \hat{f}(x, u) \rightarrow \max, \quad x \in X(u), \quad u \in U.$$

По-прежнему предполагаем, что U — метрическое пространство и $\hat{u} \in \text{int } U$.

Введем обозначения $W(u) = \max_{x \in X(u)} \hat{f}(x, u)$ и

$$M(u) = \{x \in X(u) \mid \hat{f}(x, u) = W(u)\}.$$

Имеет место следующее утверждение¹⁾.

Теорема 6.1. *Пусть отображение $X(u)$ непрерывно в точке $\hat{u} \in \text{int } U$. Пусть функция $\hat{f}(x, u)$ непрерывна в некоторой окрестности $X(\hat{u}) \times \hat{u}$. Тогда отображение $M(u)$ полунепрерывно сверху в точке $\hat{u}$.*

¹⁾Теоремы 6.1–6.3 доказаны в книге Sawaragi Y., Nakayama H., Tanino T. Theory of Multiobjective Optimization. Academic Press, 1985.

6.2. Устойчивость паретовой и слейтеровой границ

Рассмотрим вопрос об устойчивости множеств $P(Y)$ и $S(Y)$.

Теорема 6.2. Пусть множество $Y(\hat{u})$ непусто и отображение $Y(u)$ непрерывно в точке $\hat{u} \in \text{int } U$. Пусть $P(Y(\hat{u})) = S(Y(\hat{u}))$. Тогда отображение $P(Y(u))$ полунепрерывно сверху в точке $\hat{u}$.

Доказательство. Доказательство основывается на определении непрерывности многозначного отображения. Пусть имеются последовательности $\{u^k\}$ и $\{y^k\}$ такие, что $u^k \rightarrow \hat{u} \in \text{int } U$ и $y^k \in P(Y(u^k))$, причем $y^k \rightarrow \hat{y}$. Требуется доказать, что $\hat{y} \in P(Y(\hat{u}))$. Сразу заметим, что из непрерывности $Y(u)$ следует, что $\hat{y} \in Y(\hat{u})$.

Докажем, что $\hat{y} \in P(Y(\hat{u}))$. Предположим противное, т.е. пусть $\hat{y} \notin P(Y(\hat{u}))$. По условию $P(Y(\hat{u})) = S(Y(\hat{u}))$, поэтому $\hat{y} \notin S(Y(\hat{u}))$. Следовательно, найдется такая точка $y' \in Y(\hat{u})$, что $y' > \hat{y}$. Но отображение $Y(u)$ непрерывно (в том числе и полунепрерывно снизу) в точке $\hat{u}$, поэтому найдется такая последовательность $\{\bar{y}^k\}$, что $\bar{y}^k \in Y(u^k)$ и $\bar{y}^k \rightarrow y'$. Так как $y' > \hat{y}$, то для достаточно больших k будет выполняться $\bar{y}^k > y^k$. С другой стороны, $y^k \in P(Y(u^k))$. Получили противоречие, так как точки $\bar{y}^k \in Y$ доминируют паретовские точки. Значит, сделанное предположение неверно и $\hat{y} \in P(Y(\hat{u}))$. $\square$

Замечание 6.2. Для широких классов возмущений, встречающихся на практике, требование $P(Y(\hat{u})) = S(Y(\hat{u}))$ является необходимым условием для полунепрерывности сверху многозначного отображения.

Теорема 6.3. Пусть множество $Y(\hat{u})$ непусто и отображение $Y(u)$ непрерывно в точке $\hat{u} \in \text{int } U$. Пусть множества $Y(u)$ замкнуты и равномерно ограничены в некоторой окрестности точки $\hat{u}$. Тогда отображение $P(Y(u))$ полунепрерывно снизу в точке $\hat{u}$.

Доказательство. Пусть заданы последовательность $\{u^k\}$, $u^k \rightarrow \hat{u}$, и точка $\hat{y} \in P(Y(\hat{u}))$. Поскольку отображение $Y(u)$ непрерывно (и тем самым полунепрерывно снизу), существует такая последовательность $\{\tilde{y}^k\}$, что $\tilde{y}^k \in Y(u^k)$ и $\tilde{y}^k \rightarrow \hat{y}$.

Поскольку множества $Y(u)$ компактны в некоторой окрестности $\hat{u}$, то из теоремы 2.2 о непустоте множества $P(Y)$ и НМ-свойстве паретова строгого порядка на $P(Y) \in Y$ следует, что при достаточно больших k , таких что множества $Y(u^k)$ компактны, существуют точки $\{\bar{y}^k\} \in P(Y(u^k))$, удовлетворяющие соотношению $\bar{y}^k \geq \tilde{y}^k$.

Поскольку множества $Y(u)$ равномерно ограничены, последовательность $\bar{y}^k$ ограничена, поэтому из нее можно выделить сходящуюся подпоследовательность. Переходя к подпоследовательностям, в силу полунепрерывности сверху отображения $Y(u)$ получим, что $\bar{y}^k \rightarrow y' \in Y(\hat{u})$. Перейдя к пределу, получим, что $y' \geq \hat{y}$. Поскольку $\hat{y} \in P(Y)$, то $y' = \hat{y}$. Таким образом, построена последовательность точек $\{\bar{y}^k\}$, $\bar{y}^k \in P(Y(u^k))$, сходящихся к $\hat{y}$. Из определения полунепрерывности следует утверждение теоремы. $\square$

Доказанные теоремы используют непрерывную зависимость множества Y от параметров $\hat{u} \in U$. Рассмотрим условия непрерывной зависимости $Y(u)$ от u в точке $\hat{u}$.

6.3. Устойчивость множества достижимых критерияльных векторов

Итак, рассматриваются метрические пространства $\mathbb{R}^n$, $\mathbb{R}^m$ и U . Прежде всего, сформулируем аналог теорем 6.2 и 6.3.

Теорема 6.4. Пусть $X(\hat{u})$ непусто и компактно. Пусть отображение $X(u)$ непрерывно в точке $\hat{u} \in \text{int } U$, а функция $\hat{\varphi}(x, u)$ непрерывна в некоторой окрестности $X(\hat{u}) \times \hat{u}$. Тогда отображение $Y(u)$ непрерывно в $\hat{u}$.

Доказательство этого утверждения следует сразу из определений полунепрерывности сверху и снизу многозначного отображения $Y(u)$ в точке $\hat{u}$.

Как видим, в этой теореме не требуется выполнения каких-либо специфических условий непрерывности $Y(u)$ (ср. с теоремой 6.1). Для того чтобы получить количественную оценку возмущения множества Y , рассмотрим случай непрерывной по Липшицу функции $\hat{\varphi}(x, u)$, т.е. предположим, что найдутся такие числа $K_x > 0$ и $K_u > 0$, что в некоторой окрестности $O_\varepsilon(\hat{u})$ точки $\hat{u} \in \text{int } U$ имеет место

$$\rho_y(\hat{\varphi}(x', u), \varphi(x'', u)) \leq K_x \rho_x(x', x'')$$

и в некоторой окрестности $X(\hat{u})$ имеет место

$$\rho_y(\hat{\varphi}(x, u'), \varphi(x, u'')) \leq K_u \rho_u(u', u'').$$

Тогда

$$\begin{aligned} h_y(Y(u), Y(\hat{u})) &= h_y(\hat{\varphi}(X(u), u), \hat{\varphi}(X(\hat{u}), \hat{u})) \leq \\ &\leq h_y(\hat{\varphi}(X(u), u), \hat{\varphi}(X(\hat{u}), u)) + h_y(\hat{\varphi}(X(\hat{u}), u), \hat{\varphi}(X(\hat{u}), \hat{u})) \leq \\ &\leq K_x h_x(X(u), X(\hat{u})) + K_u \rho_u(u, \hat{u}), \end{aligned}$$

В качестве примера рассмотрим случай линейных критериев, т.е. $\varphi(x, u) = W(u)x$. Тогда

$$Y(u) = \{u \in \mathbb{R}^m | y = W(u)x, x \in X(u)\}.$$

Оценим расстояние между множеством $Y(u)$ и невозмущенным множеством $Y(\hat{u})$:

$$h_y(Y(u), Y(\hat{u})) \leq \|W(u) - W(\hat{u})\| \cdot \sup_{X(\hat{u})} \|x\| + \|W(u)\| \cdot h_x(X(u), X(\hat{u})).$$

Линейность критериев позволяет разделить возмущение отображения и возмущение множества. Видно, что для того, чтобы даже малое возмущение матрицы $W(\hat{u})$ не приводило к сильным возмущениям $Y(\hat{u})$, требуется ограниченность множества X .

Изучим вопрос устойчивости множества X . Рассмотрим в качестве примера один конкретный тип множества X . Пусть $X = C \cap X_A$, где $X_A = \{x \in \mathbb{R}^n | Ax \leq b\}$, в котором возмущаются матрица A и вектор b , а $C \subseteq \mathbb{R}^n$ — компактное выпуклое множество $\mathbb{R}^n$, которое не подвергается возмущению.

Справедливо следующее утверждение.

Теорема 6.5. *Пусть существует такой вектор $x^* \in C$, что $x^* \in \text{int } X_A$. Тогда множество X зависит непрерывно от элементов матрицы A и вектора b .*

Заметим, что в этом случае имеет место количественная оценка возмущений многогранного множества X , которая позволяет дать количественную оценку возмущения множества Y . Этот вопрос, однако, лежит за пределами нашего курса²⁾.

Рассмотрим теперь вопрос о возмущении множества Y_P . Как легко заметить, если конус $\mathbb{R}_+^n$ не возмущается (а это легко проконтролировать в вычислительных алгоритмах), то имеет место оценка

$$h(Y_P(u), Y_P(\hat{u})) \leq h(Y(u), Y(\hat{u})).$$

Таким образом, из устойчивости множества достижимых критериальных векторов следует устойчивость и его оболочки Эджворта-Парето. Поскольку непрерывная зависимость множества X от возмущений параметров его описания является естественным требованием в задачах оптимизации, скалярной или многокритериальной, приведенный результат показывает, что множества Y и Y_P в рассматриваемом случае являются устойчивыми при выполнении естественных предположений. Отсюда, кстати, следует устойчивость идеальной точки.

В то же время, паретова и слейтерова границы обычно не являются устойчивыми, если они не совпадают. Оказывается, что несовпадение этих границ часто встречается на практике. Доказать факт совпадения границ (если он имеет место) можно, однако это требует огромных вычислительных затрат. При этом наиболее сложна ситуация, когда некоторые паретовские грани близки к слейтеровским. Таким образом, аппроксимация множеств Y и Y_P является вычислительным смысле более естественной задачей, чем аппроксимация множеств $P(Y)$ и $S(Y)$, для решения которых приходится использовать сложные методы регуляризации задачи аппроксимации этих множеств³⁾.

²⁾ А.В. Лотов Об оценке устойчивости и числе обусловленности множества решений системы линейных неравенств // Ж. вычисл. матем. и матем. физ., т.24, №12, 1984.

³⁾ Н.М. Попов Об аппроксимации множества Парето методом сверток // Вестник Моск. ун-та. Вычисл. матем. и киберн., №2, 1982.

ГЛАВА 2. МЕТОДЫ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

Глава 7. О методах многокритериальной оптимизации

Теория многокритериальной оптимизации служит основой методов поддержки принятия решений в случае нескольких критериев выбора. Сами по себе утверждения МКО не являются средством поддержки принятия решений: для построения многокритериальных методов необходимы дополнительные исследования, осуществляемые в рамках теории принятия решений. Более того, понятие решения задачи МКО в теории и методах МКО являются различными. В теории МКО в качестве решения принято рассматривать недоминируемое множество в пространстве критериев или эффективное по Парето множество в пространстве решений. В многокритериальных методах под решением задачи МКО понимается единственная точка эффективного по Парето множества, предпочтительная для лица, принимающего решение (ЛПР). Хотя иногда в методах МКО требуется найти малое число решений, интересных с точки зрения ЛПР, в дальнейших рассуждениях в данном параграфе ограничимся случаем единственного решения. Поскольку все точки недоминируемого множества в пространстве критериев или эффективного по Парето множества в пространстве решений равноценны с точки зрения теории МКО, главную роль в многокритериальных методах поддержки принятия решений играет ЛПР — именно на основе учета предпочтений ЛПР находится то единственное эффективное по Парето решение, которое считается результатом процесса выбора. Методы поддержки принятия решений предназначены для того, чтобы помочь ЛПР в выборе наиболее предпочтительного эффективного по Парето решения. Для того, чтобы уметь правильно конструировать такие методы, необходимо иметь хотя бы приблизительное представление о том, как человек принимает решения.

7.1. О выборе решений человеком

Современное понимание человеческого поведения далеко продвинулось с начала XX века, когда наибольшее распространение имела простая психологическая концепция человека, основанная на схеме “стимул-реакция”. Согласно этой концепции, человек непосредственно реагирует на различные внешние воздействия (стимулы). Такое представление, однако, очень быстро показало свою ограниченность, поэтому были предложены более сложные теории, базирующиеся, в частности, на утверждении о том, что человек принимает решения на основе имеющейся у него целостной картины мира. К сожалению, эти представления пока являются недостаточно конструктивными для их непосредственного применения в методах поддержки принятия решений, поэтому при разработке этих методов часто (но далеко не

всегда!) используется более удобная концепция функции полезности (ценности), в рамках которой человек решает задачу максимизации некоторой функции, и нужно просто помочь ему найти ее решение. При реализации такой концепции достаточно построить математическую модель ситуации и используемую функцию полезности, а также сформулировать ограничения на возможные решения. Многие психологи критикуют этот подход как упрощенный, утверждая, что человек не всегда способен осознать и, тем более, достаточно логично выразить свои предпочтения. Экспериментальные исследования подтверждают эту точку зрения. В связи с этим специалисты по принятию решений в последнее время все чаще обращаются к более сложным моделям человека, разработанных психологами. В частности, используется понятие концептуальной, или ментальной, модели окружающего мира, используемой человеком при прогнозировании последствий своих действий.

Один из подходов к построению ментальной модели основывается на представлении о восприятии внешней среды человеком¹⁾. Принято считать, что восприятие состоит из следующих этапов: **восприятия внешней среды в ощущениях, преобразования ощущений в образы, создания логических представлений о мире**. С таким представлением о восприятии связана модель сознания человека, в которой выделяются: **уровень логического мышления, уровень образного мышления, подсознательный уровень** (рис. 7.1). Экспериментально подтверждено, что все три уровня участвуют в принятии решений, причем влияние их может быть несогласованным. В таком случае реакция человека на вопросы может отклоняться от ожидаемых логичных ответов.

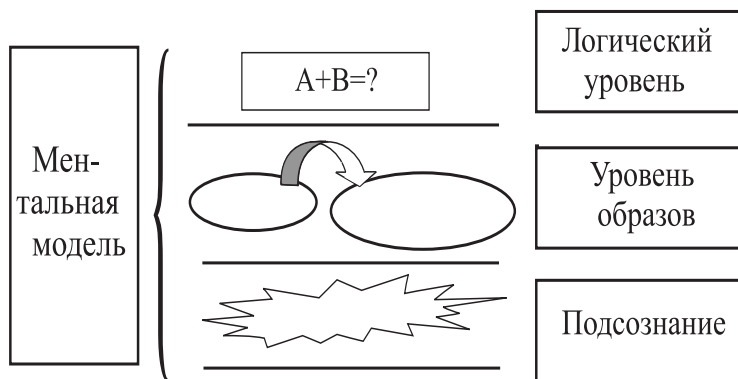


Рис. 7.1.

Приведем простой пример с использованием бинарного отношения, возникающего в результате ответов на вопросы о том, “что лучше?”. Если основываться на представлении о том, что ответы логичны, то следует ожидать, что это отношение будет транзитивно. Оказывается, однако, что даже

¹⁾Б.Ф. Ломов Методологические и теоретические проблемы психологии. М.: Наука, 1984.

в случае двух критериев достаточно легко добиться нетранзитивности ответов [9].

Пусть имеются два критерия, один из которых существенно важнее, чем другой. Человеку предлагается некий исходный вариант А, а затем вариант В, в котором значение более важного критерия незначительно ухудшается по сравнению с А, а менее важного — значительно улучшается. Человек выбирает В. Далее, действуя аналогичным образом, человеку предлагают варианты С, D, ..., Е, которые он последовательно выбирает (ряд должен быть достаточно длинным). Затем человеку опять предлагают вариант А, о котором он к этому времени уже давно забыл. Человек замечает, что значение более важного критерия в варианте Е после суммирования всех уступок значительно меньше, чем в варианте А. Поэтому человек выбирает вариант А. Тем самым, он демонстрирует нетранзитивность своих предпочтений.

Описанный выше пример является одним из самых простых примеров, показывающих, что человек может быть не в состоянии строить логичные ответы даже в относительно простых ситуациях. Возникает вопрос о том, где пределы возможностей человека в переработке информации в более сложных ситуациях. Психологические исследования показали, что объем информации, который человек способен одновременно держать в так называемой быстрой памяти, составляет около 7 единиц. Так, если называть отдельные буквы, то человек запоминает в среднем 7 букв, если слова — 7 слов, если фразы — 7 фраз и т.д. Поэтому при принятии решений человек может оперировать лишь с небольшим числом альтернатив, и то только в том случае, если он воспринимает их как целое (например, если альтернатива — это автомобиль, а не столбец с техническими параметрами). Кроме того, возникает вопрос о том, какие мысленные операции являются достаточно простыми для человека, а какие — нет, т.е. на какие вопросы ЛПР может ответить легко, а на какие — не может.

7.2. Сложность вопросов в многокритериальных методах

Рассмотрим вопросы, которые могут встретиться ЛПР в процедурах поддержки принятия решений при нескольких критериях. От ЛПР может потребоваться осуществить некоторые мысленные операции, которые можно разбить на следующие три группы: операции с критериями, с их значениями и с группами альтернатив. Эксперименты и наблюдения за поведением ЛПР позволили оценить сложность этих операций для человека [8]. Будем использовать следующие обозначения:

С — сложная операция (вопрос является слишком сложным для человека, так что он часто дает недостоверные или противоречивые ответы);

Д — допустимая операция (человек обычно дает логичные непротиворечивые ответы на вопрос, причем при повторении эксперимента не меняет свои ответы);

МР — операция, допустимая при малой размерности задачи;

НД — операция, неопределенно допустимая (недостаточно исследована, но имеющиеся данные позволяют надеяться на ее допустимость);

НС — операция, неопределенно сложная (недостаточно исследована, но имеющиеся данные говорят о том, что она скорее всего сложная).

В приведенном далее списке под сравнением значений критериев или альтернатив имеется в виду указание, какое из значений (или какая из альтернатив) является более предпочтительным. В результате экспериментов получены, в частности, следующие результаты:

- операции с критериями:
 - назначение весов критериев — С,
 - упорядочивание критериев по важности — НД,
 - декомпозиция, т.е. разбиение критериев на подкритерии — МР;
- операции со значениями критериев:
 - сравнение различных значений одного критерия — Д,
 - построение функции полезности по одному критерию — С,
 - назначение удовлетворительных значений по одному критерию — НД;
- операции с альтернативами:
 - сравнение двух альтернатив, описываемых двумя критериями — Д,
 - сравнение двух многокритериальных альтернатив — МР,
 - выбор из группы альтернатив — С,
 - назначение целевой точки — С,
 - выбор критерия, значения которого хотелось бы улучшить, уступая по другому критерию — Д,
 - определение в количественном виде увеличения значения одного критерия, компенсирующее уменьшение значения другого — Д (однократно), НС (многократно);
- сравнение альтернатив целостным образом:
 - сравнение двух альтернатив целостным образом — НД,
 - сравнение многих альтернатив — С.

Приведенный анализ сложности операций позволяет оценить реализуемость на практике многокритериальных методов поддержки принятия решений, предлагаемых различными специалистами.

Надо заметить, однако, что такие оценки сложности операций являются несколько условными, поскольку в них не учитывается фактор времени. Так, например, сравнение двух альтернатив, различающихся между собой по многим показателям, может быть успешно осуществлено, если для этого затратить достаточное время. Это становится ясно, если вспомнить первый

метод многокритериального анализа альтернатив, который был предложен знаменитым американским ученым и политическим деятелем Б. Франклином²⁾ в 1772 году. Б. Франклин отдавал себе отчет в том, что основная трудность анализа возможного решения связана с ограниченностью человеческой памяти. Он предложил простой способ — постепенно, по мере осознания, занести достоинства и недостатки некоторого решения на лист бумаги. Далее он предложил вычеркивать компенсирующие друг друга свойства решения и принять его, если оно будет иметь некомпенсированные преимущества. Крайне важно, что Б. Франклин подчеркивал необходимость затратить несколько дней на эту процедуру. Такое требование понятно с точки зрения концепции трехуровневой ментальной модели — это время требуется для того, чтобы все уровни ментальной модели пришли к консенсусу. Ясно, что и выбор из небольшого числа многокритериальных альтернатив также должен осуществляться в течение нескольких дней. Точно так же, для решения проблемы выбора из большого числа многомерных альтернатив может потребоваться несколько месяцев. Так что в приведенных оценках сложности операций недостает указания на затраченное время (по-видимому, полагалось, что на обдумывание дается не более нескольких минут). К сожалению, задача анализа влияния затрачиваемого времени на надежность результатов выбора пока не изучена.

7.3. Требования к предпочтениям ЛПР в методах МКО

При построении теории МКО делались различные предположения о свойствах критериев выбора. Возникает вопрос о том, какими свойствами должны обладать предпочтения ЛПР, чтобы было возможно применять результаты теоретических исследований задачи МКО, рассмотренных в предыдущей части книги.

Вообще говоря, в реальных задачах принятия решения ЛПР может принимать во внимание значения бесконечного числа характеристик. Как уже известно из теоретической части курса, в задачах МКО рассматривается случай, когда число рассматриваемых характеристик конечно. Точнее говоря, в рамках МКО требуется, чтобы пространство критериев было m -мерным линейным пространством $\mathbb{R}^m$, т.е. набор значений критериев $\varphi(x)$ был m -мерным вектором. Поэтому для применения методов МКО должно выполняться предположение о том, что рассматриваемый вектор критериев с достаточной полнотой описывает все интересы и предпочтения ЛПР. В этом случае качество каждого решения полностью характеризуется соответствующими значениями критериев, и поэтому выбор решения может быть осуществлен на основе выбора соответствующего вектора критериальных величин: от ЛПР требуется временно “забыть” о существовании самих решений и рассматривать лишь значения критериев. Такие задачи выбора решений принято называть *критериальными*. Это предположение, конечно, несколько искажает реальность: ЛПР интересно взглянуть и на решения.

²⁾Перевод письма Б. Франклина приведен на стр. 105 книги [9].

Но в случае большого или бесконечного числа решений это невозможно. Поэтому в задачах МКО приходится соглашаться с предположением о том, что рассматриваемый набор критериев описывает полностью интересы и предпочтения ЛПР, а информировать ЛПР можно будет только об отдельных отобранных решениях.

Главное предположение, которое делается при разработке методов МКО, состоит в следующем. Предполагается, что предпочтения ЛПР описываются некоторым не известным нам квазипорядком “не хуже”, заданным на рассматриваемом пространстве критериев $\mathbb{R}^m$, который обозначим через $\mathcal{R}_0$. Отсюда сразу следует, что мы не сомневаемся в транзитивности предпочтений ЛПР. Основным предположением при построении методов МКО является утверждение о том, что бинарное отношения Парето $\mathcal{P}$ является подмножеством квазипорядка $\mathcal{R}_0$. Именно поэтому результаты изучения бинарного отношения Парето $\mathcal{P}$ могут быть использованы при разработке методов, направленных на нахождение единственного решения, наиболее предпочтительного для ЛПР.

Для того, чтобы $\mathcal{P} \subset \mathcal{R}_0$ требуется, чтобы квазипорядок $\mathcal{R}_0$ удовлетворял одному важному предположению. Это предположение — независимость критериев по предпочтению, т.е. отсутствие зависимости направления улучшения значений критерия (его увеличение или уменьшение) от значений остальных критериев. Вообще говоря, в задаче выбора решения может оказаться так, что направление улучшения значения некоторого частного критерия зависит от значений других критериев.

Пример. Идеальная высота потолка в жилой комнате зависит от ее длины и ширины. Высоту, меньшую идеальной, желательно увеличивать, большую — уменьшать. Таким образом, желательное направление изменения высоты зависит от длины и ширины комнаты, т.е. предположение о независимости критериев по предпочтению не выполняется.

Перейдем к строгому определению независимости критериев по предпочтению. Пусть в критериальном пространстве $\mathbb{R}^m$ даны четыре точки:

$$\begin{aligned} y^{(1)} &= \begin{cases} y_i = s \\ y_j = y'_j, \end{cases} & j = \{1, \dots, m\} \setminus \{i\}, \\ y^{(2)} &= \begin{cases} y_i = s \\ y_j = y''_j, \end{cases} & j = \{1, \dots, m\} \setminus \{i\}, \\ y^{(3)} &= \begin{cases} y_i = t \\ y_j = y'_j, \end{cases} & j = \{1, \dots, m\} \setminus \{i\}, \\ y^{(4)} &= \begin{cases} y_i = t \\ y_j = y''_j, \end{cases} & j = \{1, \dots, m\} \setminus \{i\}. \end{aligned}$$

Для случая двухкритериальной задачи эти две точки изображены на рис. 7.2. Пара точек $y^{(1)}$ и $y^{(3)}$ (а также $y^{(2)}$ и $y^{(4)}$) совпадает по значению частного критерия y_2 , меняется только значение y_1 .

Определение 7.1. Критерий с номером i называют независимым по предпочтению от остальных критериев, если для всех чисел s и t имеет место следующее свойство: из того, что $y^{(1)} \succeq y^{(3)}$ для некоторого набора $y'_j, j = \{1, \dots, m\} \setminus \{i\}$, следует, что $y^{(2)} \succeq y^{(4)}$ для любого другого набора $y''_j, j = \{1, \dots, m\} \setminus \{i\}$.

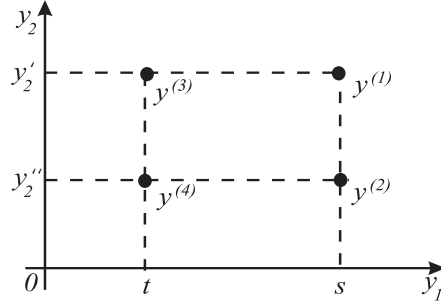


Рис. 7.2.

Для $m = 2$ (см. рис. 7.2) это определение выглядит так: если из $y^{(1)} \succeq y^{(3)}$ следует, что при любом другом значении $y_2 = y_2''$ будет выполнено $y^{(2)} \succeq y^{(4)}$, то критерий y_1 не зависит по предпочтению от критерия y_2 .

Для бинарного отношения Парето любой критерий независим по предпочтению от остальных критериев — предпочтительно увеличение значения критерия вне зависимости от значений остальных. Именно поэтому для применения методов МКО требуется, чтобы для квазиупорядка $\mathcal{R}_0$ любой критерий был независим по предпочтению от остальных критериев. Если при этом для какого-либо критерия предпочтительно уменьшение его значений, то для приведения к нужному виду его достаточно умножить на минус единицу³⁾.

Методы МКО конструируются на основе предположения о том, что квазиупорядок $\mathcal{R}_0$ содержит бинарное отношение Парето.

7.4. Классификация методов МКО

Далее будут рассматриваться методы, предназначенные для нахождения предпочтительного решения, оптимального по Парето, в случае бесконечного множества допустимых решений $X \subseteq W$, заданного совокупностью математических соотношений, и критериями $y \in \mathbb{R}^m$, связанными с решениями заданной вектор функцией $y = \varphi(x)$. При этом, как и ранее, будем рассматривать задачу многокритериальной максимизации.

К настоящему моменту разработаны тысячи различных методов МКО. Существуют различные подходы к их классификации: по типу моделей, по используемым методам скалярной оптимизации, по области приложения и

³⁾Если предпочтительным является какое-то заданное значение некоторого критерия, то для использования МКО его сначала надо заменить на новый критерий, связанный с отклонением значения исходного критерия от предпочтительного значения.

т.д. Наиболее важный из способов классификации — по роли ЛПР. С использованием такой классификации методы поддержки принятия решений в задачах МКО принято разбивать на следующие классы:

1. методы поиска предпочтительного решения без участия ЛПР;
2. методы, в которых ЛПР привлекается для построения правила выбора единственного или небольшого числа решений, оптимальных по Парето, до начала изучения множества допустимых решений;
3. интерактивные (итеративные) процедуры решения задачи с участием ЛПР, в которых ЛПР постепенно формулирует свои предпочтения в процессе изучения допустимых решений;
4. методы, основанные на аппроксимации границы Парето и информировании ЛПР о ней в том или ином виде; в таких методах ЛПР указывает предпочтительное достижимое сочетание значений критериев прямо среди точек аппроксимации границы Парето.

Методы первых двух групп основываются на построении *решающего правила*, т.е. такого отображения $C : 2^W \rightarrow 2^W$, что для каждого $X \subseteq W$ имеем $C(X) \in 2^X$, причем среди отобранных решений находится решение, наиболее предпочтительное для ЛПР. Отличие первой группы от второй состоит в том, что в первой группе методов решающее правило строится без участия ЛПР на основе некоторой аксиоматики или некоторых эвристических принципов, а во втором случае используется информация о предпочтениях ЛПР. В частности, решающее правило часто реализуется на основе концепции функции полезности (ценности), рассмотренной в §1. В этом случае решающее правило имеет вид

$$C(X) = \{x^* \in X \mid U(\varphi(x^*)) = \max_{x \in X} U(\varphi(x))\}. \quad (7.1)$$

Заметим, что идея построения решающего правила широко применяется в методах классификации малого числа альтернатив. Хотя в таких задачах также можно использовать построение функции полезности, исследователи предпочитают разрабатывать другие решающие правила, которые позволяют провести классификацию всех альтернатив. При этом решающие правила стараются разработать так, чтобы они были понятны ЛПР⁴⁾.

Далее каждый из классов методов МКО рассматривается более подробно.

⁴⁾В качестве иллюстрации такого подхода в данной главе приведен эвристический метод ELECTRE.

Глава 8. Методы, основанные на построении решающего правила

8.1. Методы без участия ЛПР

В методах, описанных в данном разделе, предлагается выбрать решающее правило, не привлекая ЛПР к решению этой проблемы. Обоснование выбора базируется на более или менее логичном содержательном объяснении, почему выбирается то или иное правило (обычно максимизация свертки критериев). Далее предлагается найти одно из решений задачи (7.1) и предъявить его ЛПР как наиболее подходящее. Могут, например, использоваться такие свертки, как

- 1) $\psi(y) = \sum_i y_i$,
- 2) $\psi(y) = \min_i y_i$,
- 3) $\psi(y) = \prod_i y_i$ при $y_i > 0$,
- 4) $\psi(y) = -\rho(y, y^*)$, где $\rho(y', y'')$ — каким-то образом определенное расстояние между критериальными точками y' и y'' , y^* — идеальная точка.

Эти функции обычно используются для критериев, имеющих одинаковую размерность. В том случае, когда величины y_i имеют разную размерность, требуется предварительно перейти к безразмерным масштабированным критериям. Это часто предлагается сделать с помощью замены:

$$\bar{y}_i = \frac{y_i - y_i^{\min}}{y_i^* - y_i^{\min}},$$

где $y_i^{\min}$ — минимальное “приемлемое” значение i -го критерия. В качестве $y_i^{\min}$ иногда берется наихудшее значение y_i на множестве Y или $P(Y)$ ¹⁾. Далее находится решение, оптимальное в смысле используемой свертки. Если ЛПР согласен использовать найденное решение, оно считается наилучшим.

Отметим, что такие способы замены предпочтений ЛПР на некоторые искусственные конструкции обычно применяются в случае отсутствия ЛПР или его нежелании принимать участие в выборе решения по каким-то причинам. В любом случае этот способ реализует произвол эксперта, предлагающего свертку. Альтернативный подход состоит в использовании какой-либо системы аксиом. При этом ЛПР, если он имеется, должен согласиться с аксиомами.

¹⁾Отметим, что наихудшее значение y_i на $P(Y)$ трудно найти из-за невыпуклости задачи оптимизации и неустойчивости границы Парето.

Надо отметить, что методы без участия ЛПР в процессе принятия решений, по своей логике не подходят для поддержки принятия решений, поскольку устраняют ЛПР из этого процесса. В то же время, такие методы могут быть приемлемы для ЛПР, не желающего нести ответственность за потенциальные негативные последствия принимаемого решения.

8.2. Методы, основанные на привлечении ЛПР к построению функции полезности

Теоретически наиболее хорошо разработаны методы, основанные на построении функции полезности $U(y)$, отражающей предпочтения ЛПР [6]. Напомним, что вопрос о связи функции полезности с бинарными отношениями уже был рассмотрен в первой части книги. Сразу отметим, что при построении функций полезности ЛПР наряду с обоснованными методами зачастую используются методы и функции, не имеющие теоретического обоснования (эвристические методы).

Отметим одну важную особенность функций полезности. Нетрудно видеть, что если $U(y)$ — функция полезности на Y , отражающая предпочтения некоторого ЛПР, $f(t)$ — строго возрастающая функция на $\mathbb{R}^1$, то функция $\tilde{U}(y) = f(U(y))$ также является функцией полезности, отражающей предпочтения этого ЛПР. Таким образом, если для бинарного отношения предпочтения ЛПР существует хотя бы одна функция полезности, то их существует бесконечно много. Это, однако, не создает проблем, поскольку интерес представляет не сама функция полезности, а определяемые ею поверхности безразличия, т.е. решения уравнения $U(y) = \text{const}$, которые не изменяются при таком преобразовании функции полезности. Для того чтобы понять роль поверхностей безразличия, достаточно вспомнить геометрическую интерпретацию решения задачи $\max_{y \in Y} U(y)$ — это такая точка $y^* \in Y$, через которую проходит линия $U(y) = \text{const}$ с максимальным значением константы.

При $m = 2$ решения уравнения $U(y) = \text{const}$ называются *кривыми безразличия*. Функции полезности принято называть *стратегически эквивалентными*, если они приводят к одной картине поверхностей безразличия.

Пример. Пусть $U(y) = A \prod_{i=1}^m y_i^{\alpha_i}$, $A > 0$, $\alpha_i > 0$, $i = 1, \dots, m$. Если $f(t) = \ln(t)$, то $\tilde{U}(y) = \ln U(y)$. Тогда $\tilde{U}(y) = \ln A + \sum_{i=1}^m \alpha_i \ln y_i$. Функции $U(y)$ и $\tilde{U}(y)$ стратегически эквиваленты.

С практической точки зрения представляет интерес вопрос о том, когда предпочтения ЛПР представимы в наиболее простом виде — в виде аддитивной функции полезности, т.е. $U(y) = \sum_i U_i(y_i)$. Так, как показывает пример, к этому классу предпочтений относится предпочтение, задаваемое мультипликативной функцией $U(y) = A \prod_{i=1}^m y_i^{\alpha_i}$. Для того, чтобы представить предпочтения в виде аддитивной функции полезности, эти предпочте-

ния должны удовлетворять довольно жестким ограничениям.

Рассмотрим случай $m = 2$, когда кривые безразличия можно изобразить на плоскости. Пусть $U(y_1, y_2)$ непрерывно дифференцируема. Тогда уравнение кривой безразличия $U(y_1, y_2) = \text{const}$ имеет вид

$$\frac{\partial U}{\partial y_1} dy_1 + \frac{\partial U}{\partial y_2} dy_2 = 0.$$

Тогда

$$\frac{dy_1}{dy_2} = -\frac{\partial U}{\partial y_2} / \frac{\partial U}{\partial y_1} = -\lambda.$$

Величина

$$\lambda = \frac{\partial U}{\partial y_2} / \frac{\partial U}{\partial y_1}$$

называется *предельной нормой замещения (tradeoff rate)* (см. [6]).

Пусть структура предпочтения аддитивна, т.е. в произвольной точке Z с координатами (y_1, y_2) для функции полезности справедливо представление $U(Z) = U_1(y_1) + U_2(y_2)$. Тогда $\lambda_Z = U'_2(y_2)/U'_1(y_1)$. Таким образом, для касательных к кривым безразличия в точках $A = (y_1^0, y_2^0)$, $B = (y_1^0, y_2^1)$, $C = (y_1^1, y_2^0)$ и $D = (y_1^1, y_2^1)$, изображенных на рис. 8.1, должно выполняться следующее соотношение:

$$\lambda_D = \frac{U'_2(y_2^1)}{U'_1(y_1^1)} = \frac{U'_2(y_2^1)}{U'_1(y_1^0)} \cdot \frac{U'_1(y_1^0)}{U'_2(y_2^0)} \cdot \frac{U'_2(y_2^0)}{U'_1(y_1^1)} = \frac{\lambda_B \lambda_C}{\lambda_A}.$$

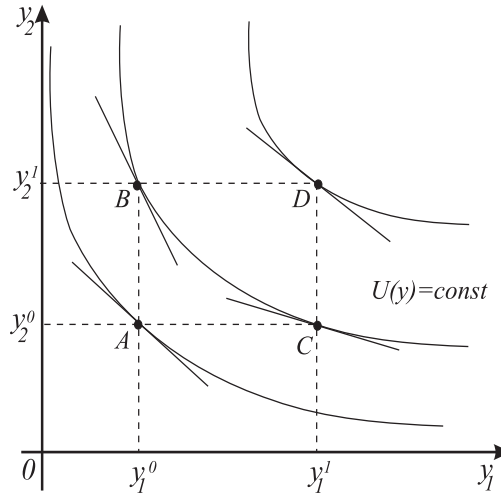


Рис. 8.1.

Получили, что необходимым условием аддитивности функции полезности при $m = 2$ является выполнение равенства $\lambda_A \lambda_D = \lambda_B \lambda_C$ для любых A, B, C и D . Как можно показать, оно является и достаточным. Для того, чтобы обоснованно использовать аддитивные функции полезности, требуется

проверить выполнение полученных условий для всех совокупностей вершин прямоугольников в положительном ортанте.

Пусть функция $U(y)$ непрерывна на связном множестве Y . Рассмотрим случай $m = 3$. Будем называть пару критериев (y_1, y_2) *независимой по предпочтению* от критерия y_3 , если верна импликация $(y'_1, y'_2, y_3^*) \succ (y''_1, y''_2, y_3^*) \Rightarrow (y'_1, y'_2, y_3) \succ (y''_1, y''_2, y_3)$ при любых значениях y_3^* и y_3 , т.е. структура предпочтения в пространстве (y_1, y_2) не зависит от y_3 . Если все три пары критериев независимы по предпочтению от оставшегося критерия, то структура предпочтения аддитивна. Проверка независимости по предпочтению может оказаться непростым делом.

При $m > 3$ аддитивность структуры предпочтений имеет место, если любые наборы критериев независимы по предпочтению от остальных критериев, не вошедших в эти наборы. Проверка всех этих условий бывает весьма затруднительна, поэтому зачастую исследователи, использующие построение функций полезности, без всякого обоснования считают, что эти функции аддитивны. Дело в том, что для аддитивной структуры предпочтений построить поверхности безразличия гораздо проще, чем в общем случае.

8.3. Построение поверхностей безразличия в случае аддитивной структуры предпочтений

В данном разделе описана обоснованная процедура построения поверхностей безразличия для аддитивной структуры предпочтений, предложенная в [6]. Начнем с процедуры построения кривых безразличия для аддитивной структуры предпочтений в случае двух критериев, т.е. будем рассматривать функции полезности, представимые в виде $U(y) = U_1(y_1) + U_2(y_2)$, $y \in \mathbb{R}^2$. Эта процедура может быть проиллюстрирована графически (см. рис. 8.2).

Поскольку функция полезности не является однозначно определенной, прежде всего необходимо указать точку нулевой полезности и единицу измерения. Для этого ЛПР должен указать минимально допустимые значения критериев $y^{min} = (y_1^{min}, y_2^{min})$, для которых имеет смысл проводить построение поверхностей безразличия. Соответствующую точку на рис. 8.2 обозначим через A . Положим

$$U(A) = U_1(A) = U_2(A) = 0.$$

Затем определяется значение y_1^1 , при котором отклонение от минимального является существенным для ЛПР. Если положить $U_1(y_1^1) = 1$, то будет задан масштаб для функции полезности. Обозначим точку (y_1^1, y_2^{min}) через B .

но.

Процедуру построения аддитивной функции полезности можно считать сложной для человека, так как операции сравнения критериальных точек осуществляется большое число раз. В результате этого человек, которому задаются данные вопросы, может ошибаться в своих оценках или давать противоречивые ответы. Скажем, может оказаться, что $U(K)$ не равно $U(H)$, что может означать отсутствие аддитивности или наличие ошибок в ответах ЛПР. Более того, для проверки аддитивности предпочтений нужно бы проверять условие типа $\lambda_A \lambda_D = \lambda_B \lambda_C$ для различных прямоугольников на критериальной плоскости.

Когда данная процедура переносится на многомерный случай, она становится значительно сложнее. Уже на первом шаге число точек единичной полезности равно размерности пространства, а поверхность безразличия представляет собой грань симплекса. На следующих шагах ситуация еще усложняется. Соответственно возрастает сложность алгоритма, усугубляемая отсутствием графической поддержки. При этом возможность ошибки ЛПР возрастает многократно, тем более что при $m \geq 3$ условия независимости по предпочтениям имеют качественный вид и их проверить практически невозможно.

Отметим, что иногда для того, чтобы избежать использования описанной выше теоретически обоснованной процедуры построения поверхностей безразличия предлагают строить аддитивную функцию полезности в виде

$$U(y) = \sum_{i=1}^m c_i U_i(y_i), \quad (8.1)$$

где $c_i > 0$ — веса, задаваемые ЛПР, $U_i(y_i)$ — функции, также задаваемые ЛПР в диалоге с компьютером на основе непосредственного назначения величин полезности для конечного числа значений одного критерия. Такой подход не является теоретически обоснованным, поскольку аддитивность функции полезности еще не означает возможность использования какого-либо конкретного вида (8.1). Аддитивность означает существования каких-то заранее неизвестных функций критериев, но теория не отвечает на вопрос, какие функции надо использовать.

Рассмотрим пример. Пусть предпочтение задается функцией $U(y) = \sqrt{y_1 y_2}$, $y_1, y_2 > 0$. Ясно, что это предпочтение аддитивно. При применении описанного подхода ЛПР, быть может, и нарисует однокритериальные функции полезности $U_1(y_1) = \sqrt{y_1}$ и $U_2(y_2) = \sqrt{y_2}$, но, положив в связи с одинаковой важностью критериев $c_1 = c_2 = 1$, получит функцию $\hat{U}(y) = \sqrt{y_1} + \sqrt{y_2}$, которая не совпадает с настоящей функцией $U(y) = \sqrt{y_1 y_2}$.

Надо подчеркнуть, что описанный подход привлекателен для многих в связи с его простотой и мнимой близостью к таким теоретическим понятиям как аддитивность предпочтения. Все же, такой подход теоретически не обоснован и является эвристическим.

Далее опишем некоторые другие эвристические методы построения правил выбора, основанные скорее не на абстрактных, а практических сообра-

жениях.

8.4. Современные эвристические подходы

Целевой подход

Целевой подход основан на том, что лицо, принимающее решение, должно назначить цель — некоторую (обычно недостижимую) критериальную точку $\hat{y}$, отражающую его желания. В задачах многокритериальной максимизации, рассматриваемых нами, требованиям ЛПР удовлетворяет конус точек $Z = \{z \in \mathbb{R}^m | z \geq \hat{y}\}$. В качестве функции полезности используется функция $U(y) = -\rho(y, Z)$, где $\rho(y, Z)$ — отклонение критериального вектора y от Z . Решающее правило состоит в выборе допустимого решения, минимизирующего $\rho(y, Z)$ на множестве $Y = \varphi(X)$. Точнее говоря, находится такое решение, последствия которого наиболее близки к конусу точек, удовлетворяющих требованиям ЛПР. Близость при этом понимается в смысле используемого понятия отклонения от конуса $\rho(y, Z)$.

Достоинством подхода является его простота: от ЛПР требуется лишь один раз выразить свои предпочтения — указать цель $\hat{y}$. Далее решение находится автоматически. С другой стороны, если последствия найденного решения далеки от цели (часто дело обстоит именно так в связи с тем, что ЛПР из-за отсутствия поддержки в процессе назначения цели указывает ее без знания реальных возможностей), это может вызвать недовольство у ЛПР. Кроме того, если целевая точка сильно отклоняется от совокупности достижимых критериальных векторов Y , то полученное решение больше зависит от того, какая функция $\rho(y, Z)$ используется, чем от самой цели. В связи с тем, что ЛПР обычно не участвует в выборе этой функции (эта задача слишком сложна для человека, не имеющего математического образования), то в таком случае оптимальное решение фактически определяется экспертом, выбравшем ту или иную функцию.

Благодаря своей простоте целевой подход является наиболее распространенным методом МКО. Для повышения обоснованности выбора целевой точки $\hat{y}$ полезно проинформировать ЛПР об идеальной точке y^* .

Об использовании линейных функций полезности

Зачастую исследователи вместо построения поверхностей безразличия сложных нелинейных функций полезности (или самих таких функций) пытаются ограничиться в качестве функции полезности простой линейной сверткой критериев $U(y) = \sum_i c_i y_i$, коэффициенты c_i которой (веса или важности критериев) должен указать ЛПР. Эта функция имеет существенные недостатки. Прежде всего, непонятно, почему предпочтения ЛПР должны выражаться в такой простой форме. Функции полезности, рассматриваемые в теории многокритериальной полезности, обычно вогнуты, и это отражает существо дела — веса, отражающие важность критериев, должны

зависеть от значений самих критериев. При недостаточной величине одного из критериев его важность должна быть велика, а полезность такого несбалансированного набора не может быть большой. Скажем, недостаточную скорость самолета нельзя компенсировать его отличной маневренностью. В линейной же функции малое значение одного критерия *всегда* можно компенсировать избыточным значением другого.

Далее, как было показано в главе, посвященной теории МКО, линейные функции могут использоваться только для эффективно-выпуклых задач, в противном же случае часть недоминируемых точек не может быть получена на основе максимизации линейной функции. Из-за этого некоторые недоминируемые критериальные точки, быть может, наиболее предпочтительные для ЛПР, остаются за рамками анализа.

Рассмотрим очень простой пример. Абитуриент осуществляет выбор одного из трех университетов, руководствуясь двумя критериями — уровнем преподавания и условиями для занятий спортом. Пусть для первого университета (A1) значение первого критерия y_1 равно 10, второго критерия y_2 равно 1, для второго университета (A2) — $y_1 = 5$ и $y_2 = 5$, для третьего университета (A3) — $y_1 = 1$ и $y_2 = 10$. Таким образом, первый университет характеризуется отличным уровнем преподавания, но имеет плохие условия для занятия спортом, второй университет имеет вполне удовлетворительное качество преподавания и удовлетворительные условия занятия спортом, третий университет — неудовлетворительное качество преподавания, но отличные условия занятия спортом. На рис. 8.3 в плоскости критериев изображены критериальные точки A_1 , A_2 и A_3 , соответствующие университетам A1, A2 и A3. Решить задачу — выбрать одну из точек A_1 , A_2 и A_3 .

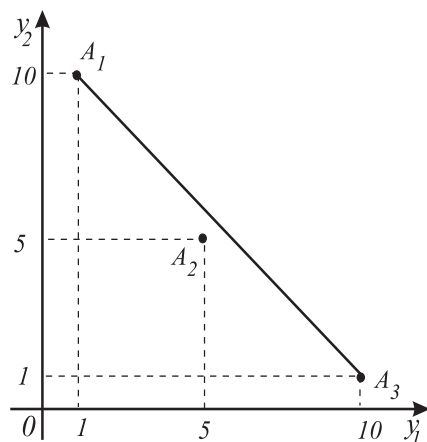


Рис. 8.3.

Соединим точки A_1 , A_3 отрезком прямой. Рассмотрим функции полезности, представляющие собой линейные свертки критериев $U(y) = c_1y_1 + c_2y_2$ с различными положительными весами c_1, c_2 . Поскольку линии уровня ли-

нейной функции полезности представляют собой прямые, а точка A_2 лежит ниже отрезка A_1, A_3 , при любых положительных весах c_1, c_2 решением задачи могут быть только точки A_1 и A_3 . Таким образом, университет A_2 , характеризующийся сбалансированным набором условий, не будет получен абитуриентом для внимательного рассмотрения, а ведь именно он мог оказаться наиболее предпочтительным.

Заметим, что часто требуется назначить веса в случае близких критериальных точек. В таком случае использование линейной функции полезности вполне разумно — веса линейных функций соответствуют градиенту нелинейной функции полезности, который меняется мало при малом изменении точки в пространстве критериев.

Метод анализа иерархий

Метод анализа иерархий (Analytic Hierarchical Process, АНР) базируется на линейной свертке критериев. Этот популярный метод был предложен известным американским ученым Т. Саати для анализа политических ситуаций, в которых невозможно использовать математическое описание проблем выбора решения²⁾. В связи с тем, что метод АНР стал широко использоваться и в задачах многокритериальной оптимизации, кратко рассмотрим его основные идеи.

Метод анализа иерархий основан на модификации процедуры назначения весов. Рассмотрим простейший случай применения метода. Хотя метод АНР применим и при бесконечном числе вариантов, мы ограничимся его применением в случае выбора одного из конечного числа N альтернативных вариантов, каждый из которых характеризуется набором из m критериев $y_1, y_2, \dots, y_m$. Будем считать, что значения критериев для каждого из вариантов известны заранее. Хотя эти величины в методе АНР также могут быть оценены (и обычно оцениваются в случае политических или социальных проблем) с помощью того же метода, здесь для упрощения описания эту возможность рассматривать не будем. Заметим, что в случае использования математических моделей для расчета значений критериев эти значения можно считать заданными, поскольку в связи с конечностью числа вариантов для их расчета требуется провести ограниченный объем вычислений.

Предполагается, что предпочтения можно представить в линейном виде, т.е. $U(y) = \sum_{i=1}^m W_i y_i$, так что задача состоит в том, чтобы определить веса³⁾ критериев $W = (W_1, W_2, \dots, W_m)$. По мнению Т. Саати, основная проблема заключается в том, что человеку трудно назначить веса непосредственно, поэтому он предлагает получать от ЛПР не сами значения весов, а показатели *относительной важности критериев* a_{ij} , $i, j = 1, \dots, m$. Значения a_{ij} должны выбираться из некоторой шкалы, например, $\{1, 2, \dots, 9\}$, причем

²⁾Т. Саати Принятие решений. Метод анализа иерархий. М.: Радио и связь, 1993.

³⁾Вместо привычных $c_1, \dots, c_m$ здесь сохраняются обозначения Т. Саати, т.е. используются $W_1, \dots, W_m$.

число 9 означает, что критерий i много важнее критерия j , а число 1 — что критерии примерно эквивалентны по важности. Автоматически полагается $a_{ji} = 1/a_{ij}$ и $a_{ii} = 1$. Как не трудно заметить, от человека требуется дать $m(m-1)/2$ ответов на вопросы о величинах $a_{ij}, i < j$. На основе его ответов в методе оцениваются m величин $W_i, i = 1, \dots, m$. Матрица $A = \|a_{ij}\|^{m \times m}$ содержит избыточную информацию, которая используется в методе для контроля логичности ответов ЛПР о величинах a_{ij} .

Расчет величин $W_i, i = 1, \dots, m$ и проверка логичности осуществляются на основе введения понятия непротиворечивых ответов ЛПР. Непротиворечивым является такая матрица величин a_{ij} , для которой найдется некоторый набор весов $W_i, i = 1, \dots, m$, которому она соответствует, т.е.

$$a_{ij} = \frac{W_i}{W_j}, \quad i, j = 1, \dots, m. \quad (8.2)$$

В этом случае естественно считать, что набор весов $W_i, i = 1, \dots, m$, отражает интересы ЛПР. При выполнении (8.2) матрицу величин a_{ij} обозначим через A^0 , т.е.

$$A^0 = \left\| \frac{W_i}{W_j} \right\|^{m \times m}.$$

Нетрудно заметить, что если это так, то матрица A^0 , удовлетворяет соотношению $A^0 W = mW$. Таким образом, W — собственный вектор, а $\lambda_1 = m$ — собственное число матрицы A^0 . При этом строки матрицы A^0 пропорциональны, так что ее ранг равен 1. Из этого следует, что уравнение $A^0 v = 0$ имеет $m-1$ линейно независимых решений, так что остальные собственные числа матрицы A^0 равны нулю.

Вернемся к реальной ситуации. Обычно ЛПР дает противоречивые ответы, т.е. строки матрицы $A = \|a_{ij}\|^{m \times m}$ не пропорциональны, поэтому ее собственные значения отличаются от собственных значений матрицы A^0 . В методе АНР матрица A рассматривается как некоторое возмущение матрицы A^0 , поэтому ожидается, что собственные числа матрицы A будут все же близки к собственным числам матрицы A^0 . Так как по построению, матрица A положительна, то по теореме Фробениуса-Перрона [1] у нее существует максимальное положительное собственное число $\lambda_{max} > 0$ и положительный собственный вектор W . Собственное число λ_{max} рассматривается как возмущение собственного числа $\lambda_1 = m$ невозмущенной матрицы A^0 . В связи этим собственный вектор матрицы A , соответствующий собственному числу λ_{max} , используется в качестве аппроксимации искомого набор весов, отражающих интересы ЛПР.

Величина $\Delta = \frac{\lambda_{max} - m}{m - 1}$ используется в качестве характеристики отклонения матрицы A от невозмущенной матрицы A^0 , т.е. она является оценкой непоследовательности ЛПР. Если величина Δ достаточно мала (скажем, по сравнению с этой величиной для случайной матрицы размера $m \times m$, удовлетворяющей $a_{ji} = 1/a_{ij}$), то полученные результаты удовлетворительны, а указанный собственный вектор матрицы A может использоваться в качестве вектора весов.

Кроме недостатков, свойственных всем методам, основанным на линейной свертке критериев, метод АНР имеет и специфические недостатки. В частности, было показано, что введение дополнительного неэффективного варианта может привести к такому изменению весов, что это приведет к изменению решения задачи. Поэтому метод АНР является эвристическим, и его надо использовать с большой осторожностью.

Стоит отметить, что метод АНР получил широкое распространение, уступая только целевому подходу, благодаря своей простоте и личным способностям автора, его умению организовать рекламу метода. Т. Саати удалось внедрить метод среди лиц, которые до этого не имели представления о многокритериальных методах поддержки принятия решений и полагались исключительно на экспертов и собственную интуицию. Существенным фактором для таких лиц является то, метод АНР не требует построения математических моделей, для чего необходимы знания и умения, отсутствующие у них.

Для того чтобы устранить недостатки, возникающие из-за использования линейной функции полезности, было предложено рассмотреть функцию полезности вида $U(y) = \sum_{i=1}^m W_i U_i(y_i)$, т.е. сначала для каждого критерия построить однокритериальную функцию полезности U_i , а затем с помощью метода АНР найти веса критериев. Такой подход позволяет устранить недостатки, характерные для линейных функций полезности, но все же остается эвристическим подходом, не гарантирующим правильность представления предпочтений (см. пример в п. 8.2).

Отметим, что в последнее время были разработаны более сложные варианты метода анализа иерархий, однако, они лежат за пределами нашего курса.

Глава 9. Метод ELECTRE

Метод ELECTRE является одним из первых методов многокритериальной классификации альтернатив (outranking). Он был предложен известным французским ученым Б. Руа в 70-е годы XX века и породил целое направление в области методов поддержки выбора из конечного числа альтернатив при нескольких критериях. В методе ELECTRE вместо функции полезности строится правило в виде бинарного отношения, которое позволяет выделить подмножество альтернатив из исходной совокупности. Альтернативы заданы значениями своих показателей (критериев выбора), которые могут быть получены различными путями, в том числе и с использованием математического моделирования.

Рассмотрим N альтернатив, каждая из которых характеризуется значениями m критериев, причем в качестве критериев могут выступать как количественные, так и качественные показатели. Значение i -го критерия для выбора j -й альтернативы будем обозначать y_{ji} , $i = 1, \dots, m$, $j = 1, \dots, N$.

Прежде всего, в методе ELECTRE требуется перейти от количественных показателей к качественным. Это осуществляется разбиением множества количественных значений на конечное число классов, каждому из которых присваивается качественная характеристика. После этого каждый из альтернативных вариантов оказывается задан наборами из m качественных показателей.

9.1. Общая схема метода ELECTRE

Общая схема метода состоит из четырех этапов.

1. Назначение весов. Лицо, принимающее решение, определяет веса каждого из критериев $W_1, W_2, \dots, W_m$.

2. Построение индекса согласия. Для каждой пары альтернатив j и k множество критериев $I = \{1, 2, \dots, m\}$ разбивается на три группы:

$$I_{jk}^+ = \{i \in I \mid y_{ji} \succ y_{ki}\}, \quad I_{jk}^- = \{i \in I \mid y_{ji} \prec y_{ki}\}, \quad I_{jk}^o = \{i \in I \mid y_{ji} \approx y_{ki}\}.$$

Множество I_{jk}^+ включает те критерии, по которым j -я альтернатива лучше k -й, множество I_{jk}^- , состоит из критериев, по которым j -я альтернатива хуже k -й, а множество I_{jk}^o состоит из тех критериев, по которым j -я и k -я альтернативы эквивалентны. *Индекс согласия* с тем, что альтернатива j лучше альтернативы k определяется следующим образом:

$$j k = \frac{\sum_{i \in I_{jk}^+} W_i + \alpha \sum_{i \in I_{jk}^o} W_i}{\sum_{i=1}^m W_i},$$

где α — параметр, $\alpha \in \{1, 0.5, 0\}$ (выбор параметра α зависит от того, какая модификация метода реализуется).

3. Построение индекса несогласия. Для каждой пары j и k индекс несогласия с тем, что альтернатива j лучше альтернативы k , определяется по формуле

$$d_{jk} = \frac{1}{100} \max_i \left\{ \begin{array}{l} \text{интервал превосходства } k\text{-й альтернативы} \\ \text{над } j\text{-й по } i\text{-му критерию} \end{array} \right\},$$

где интервал превосходства k -й альтернативы над j -й по i -му критерию определяет количество последовательных переходов из класса в класс, которое необходимо осуществить для того, чтобы j -й вариант стал эквивалентен k -му по i -му критерию, умноженное на цену одного такого перехода. Предполагается, что величины d_{jk} не могут превосходить единицу.

4. Построение решающего правила. На основе чисел $p \in (0, 1]$ и $q \in [0, 1)$, определяемых ЛПР, на множестве альтернатив строится следующее бинарное отношение: j -я альтернатива признается лучше альтернативы k , при условии того, что $c_{jk} \geq p$ и $d_{jk} \leq q$. Сразу можно заметить, что при $\alpha = 0.5$, $p = 1$ и $q = 0$ указанное бинарное отношение совпадает с бинарным отношением Слейтера. Действительно, в этом случае j -я альтернатива доминирует k -ю тогда и только тогда, когда $c_{jk} = 1$ и $d_{jk} = 0$, т.е. $i \in I_{jk}^+$ для всех $i = 1, \dots, m$. При $p < 1$ и $q > 0$ может оказаться, что и другие пары альтернатив связаны введенным бинарным отношением.

После того, как бинарное отношение построено, лицу, принимающему решение, представляется множество взаимно недоминируемых альтернатив, на котором построенное бинарное отношение обладает НМ-свойством. Это множество зачастую называют ядром, но ядром по фон Нейману-Моргенштерну оно может не являться, поскольку это бинарное отношение не обязательно транзитивно. Так, в таком “ядре” могут присутствовать альтернативы, доминируемые альтернативами, не включенными в “ядро”. Далее ЛПР выбирает окончательное решение из этого “ядра”.

Таким образом, метод ELECTRE позволяет сократить число анализируемых вариантов, облегчая тем самым выбор ЛПР. Покажем, как работает этот метод, на примере задачи о покупке автомобиля.

9.2. Пример использования метода ELECTRE

Допустим, мы хотим купить автомобиль, выбрав один из семи вариантов ($N = 7$). Каждый автомобиль будет оцениваться по четырем показателям ($m = 4$): цене, комфортности салона, скорости и внешнему виду. Цена и скорость являются количественными критериями, поэтому разобьем область их значений на классы, присвоив каждому классу свой идентификатор и перейдя тем самым к качественным показателям.

Возможные значения показателей для нашего примера представлены в таблице 1, а в таблице 2 с учетом введенных обозначений перечислены значения этих критериев для каждого из семи автомобилей (в таблице 2 представлена матрица $\|y_{ji}\|^{N \times m}$).

Таблица 1

	Показатели	Значения	Код
1	цена	≤ 27000	25
		28000 - 32000	30
		33000 - 37000	35
		38000 - 42000	40
		43000 - 47000	45
2	комфортность	высокая	H
		средняя	M
		низкая	L
3	скорость	быстрая	F
		медленная	S
4	внешний вид	красивый	B
		приемлимый	A

Таблица 2

	1	2	3	4
1	45	H	F	B
2	40	H	S	B
3	40	M	F	B
4	35	M	F	A
5	35	M	S	B
6	35	L	F	B
7	25	L	S	A

Теперь применим описанный метод.

1. Выберем следующие веса: $W_1 = 5$, $W_2 = 3$, $W_3 = 1$, $W_4 = 1$.

2. Считая, что $\alpha = 0.5$, вычислим индексы согласия для каждой пары вариантов выбора автомобилей. Рассмотрим сначала автомобили 1 и 2. Автомобиль 1 лучше автомобиля 2 по третьему показателю (по скорости), и совпадает с ним по второму и четвертому (по комфортности и по качеству внешнего вида), поэтому $I_{jk}^+ = \{3\}$, $I_{jk}^- = \{1\}$, $I_{jk}^o = \{2, 4\}$. Согласно заданным весам

$$_{12} = \frac{W_3 + 0.5(W_2 + W_4)}{W_1 + W_2 + W_3 + W_4} = \frac{1 + 0.5(3 + 1)}{10} = 0.3.$$

Аналогично

$$_{21} = \frac{W_1 + 0.5(W_2 + W_4)}{W_1 + W_2 + W_3 + W_4} = \frac{5 + 2}{10} = 0.7.$$

Как нетрудно видеть, при $\alpha = 0.5$ сумма симметричных индексов согласия равна единице. Сравнивая все остальные автомобили попарно, получим матрицу согласия

$$C = \frac{1}{10} \begin{pmatrix} - & 3 & 4 & 4.5 & 4.5 & 4 & 5 \\ 7 & - & 6 & 4 & 4 & 3.5 & 4.5 \\ 6 & 4 & - & 3 & 3 & 4 & 5 \\ 5.5 & 6 & 7 & - & 5 & 6 & 4 \\ 5.5 & 6 & 7 & 5 & - & 7 & 4.5 \\ 6 & 6.5 & 6 & 4 & 3 & - & 3.5 \\ 5 & 5.5 & 5 & 6 & 5.5 & 6.5 & - \end{pmatrix}.$$

3. Для определения индексов несогласия необходимо задать цену перехода значения в соседний класс по каждому из критериев. В нашем примере будем считать, что понижение цены автомобиля на один уровень оценивается в 20 баллов, увеличение скорости — в 25 баллов, повышение удобства салона и качества внешнего вида — по 20 баллов. Вычислим индекс несогласия для первого и второго варианта выбора автомобиля. Первый автомобиль хуже второго только по цене, причем их разделяет лишь один переход, который оценивается в 20 баллов, поэтому $d_{12} = 0.2$. Найдём теперь d_{21} . Так как автомобиль 1 превосходит автомобиль 2 по скорости, причем на один класс, и цена перехода составляет 25 баллов, то $d_{21} = 0.25$. Рассчитав все остальные индексы несогласия, получим матрицу несогласия

$$D = \frac{1}{10} \begin{pmatrix} - & 2 & 2 & 4 & 4 & 4 & 8 \\ 2.5 & - & 2.5 & 2 & 4 & 4 & 6 \\ 2 & 2.5 & - & 2 & 2.5 & 2 & 6 \\ 2 & 2.5 & 2 & - & 2.5 & 2 & 4 \\ 2.5 & 2 & 2.5 & 2.5 & - & 2.5 & 4 \\ 4 & 4 & 2 & 2 & 2.5 & - & 4 \\ 4 & 4 & 2.5 & 2.5 & 2 & 2.5 & - \end{pmatrix}$$

4. Для построения решающего правила выберем $p = 0.6$, $q = 0.2$ и определим, между какими вариантами автомобилей существует доминирование. Для варианта 1 в 1-й строке матрицы несогласия D ищем элементы, не превосходящие 0.2. Таких элементов два, они соответствуют автомобилям 2 и 3. Однако индексы согласия с тем, что первый автомобиль лучше второго и третьего, равны, соответственно, $c_{12} = 0.3$ и $c_{13} = 0.4$. Эти числа меньше 0.6, поэтому первый автомобиль не доминирует никакой другой автомобиль. То же самое получается и для второго автомобиля.

Третья строка матрицы D содержит три элемента, равных 0.2 — первый, четвертый и шестой. Среди чисел, стоящих на соответствующих позициях в третьей строке матрицы C , есть одно, не меньшее 0.6 — это первый элемент. Следовательно, третий автомобиль доминирует первый.

Поступая дальше аналогичным образом, находим все остальные пары, связанные отношением доминирования. На рис. 9.1 *a* такие пары соединены стрелками: стрелка идет от лучшего варианта к худшему. Как видно, построенное бинарное отношение не обладает свойством транзитивности. Так, четвертый автомобиль доминирует третий, третий доминирует первый, но четвертый не доминирует первый, так что надо быть осторожным при построении “ядра”. Для нашего примера в случае $p = 0.6$, $q = 0.2$ в множество взаимно недоминируемых альтернатив бинарного отношения войдут недоминируемые автомобили 4, 5, 7, а также автомобиль 1, который включается в “ядро” для того, чтобы сохранялось НМ-свойство.

Найдем множество взаимно недоминируемых альтернатив бинарного отношения для рассматриваемого примера при других значениях p и q . В случае, когда $p = 1$, $q = 0$, “ядро” будут составлять все семь вариантов автомобилей. Будем теперь постепенно уменьшать p и увеличивать q . При $p = 0.9$, $q = 0.1$ “ядро” опять будут составлять все семь автомобилей. Та же самая ситуация будет и при $p = 0.8$, $q = 0.2$. Если взять $p = 0.7$ и $q = 0.2$, то четвертый автомобиль доминирует третий, и, следовательно, в “ядро” войдут все варианты, кроме третьего. В случае, когда $p = 0.6$, $q = 0.2$, “ядро” будут составлять автомобили 1, 4, 5, 7. Если же $p = 0.6$, $q = 0.25$, то возникающие в этом случае отношения доминирования представлены на рис. 9.1 *b*, причем “ядро” состоит из автомобилей 1, 5, 7.

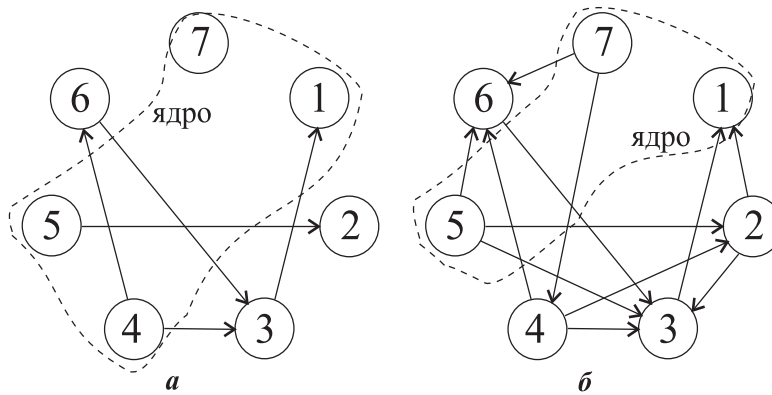


Рис. 9.1.

Очень интересна ситуация, когда $p = 0.5$, $q = 0.25$. Для автомобилей 5 и 4 индексы согласия и несогласия имеют значения $c_{54} = 0.5 \geq p$, $d_{54} = 0.25 \leq q$, т.е. автомобиль 5 доминирует автомобиль 4, но в то же время $c_{45} = 0.5 \geq p$ и $d_{45} = 0.25 \leq q$, т.е. автомобиль 4, в свою очередь, доминирует автомобиль 5. Таким образом, бинарное отношение не только не является транзитивным, но и может приводить к циклам доминирования, причем состоящим всего из двух вариантов, что уже совсем недопустимо. Хорошо, если все элементы цикла доминируются какими-то другими, не входящими в этот цикл вариантами, так что его можно просто исключить из рассмотрения. В нашем случае “ядро” будет состоять из автомобилей 2 и 7. Однако в общем случае неясно, что делать, если образующие цикл варианты являются недоминируемыми. Ясно, для того, чтобы подобных ситуаций не возникало, величина p должна быть достаточно велика.

9.3. Выводы

Подведем итоги. Итак, от ЛПР в процессе реализации метода ELECTRE требуется получить

- 1) веса критериев,
- 2) цены перехода из класса в класс для построения индексов несогласия,
- 3) числа p и q для построения бинарного отношения.

При этом результат зависит от того, какие p и q будут выбраны. Так, если p близко к единице, а q к нулю, то с помощью метода ELECTRE можно исключить только доминируемые альтернативы. При других значениях p и q появляются другие возможности доминирования и могут возникнуть ситуации, когда доминируемая альтернатива попадает в “ядро”.

Так как ЛПР довольно сложно сразу назначить p и q разумным образом, то был предложен следующий подход. В качестве начальных значений выбираются $p = 1$ и $q = 0$, затем они постепенно меняются. При этом ЛПР

получает информацию о том, какие альтернативы попадают в “ядро”, а также диаграммы, изображающие отношения доминирования между альтернативами, аналогичные тем, что представлены на рис. 9.1. Когда изменения параметров начинают приводить к противоречиям, то процесс останавливается, и ЛПР выбирает наиболее приемлемый для себя вариант значений p и q из рассмотренных ранее. Можно также исключать часть альтернатив при больших значениях p и малых q , а затем продолжать анализ с меньшим p и большим q только для тех вариантов, которые не были исключены на предыдущих шагах. Разработаны и более сложные варианты метода ELECTRE, которые получили название ELECTRE II и ELECTRE III (см. [9]).

Отметим, что существует очень важное отличие метода ELECTRE от рассмотренных ранее методов: решающее правило, с помощью которого осуществляется выбор между имеющимися альтернативами, не определяется заранее, а меняется с учетом мнения ЛПР. В этом процессе человек меняет параметры алгоритма в соответствии с тем, какими свойствами обладает решаемая задача и добивается наиболее приемлемого для себя результата. Это является характерным свойством итеративных методов, которые мы будем изучать в дальнейшем.

Глава 10. Общее представление об итеративных процедурах

Итеративные человеко-машинные многокритериальные методы возникли в 60-х годах XX века. Они принципиально отличались от других многокритериальных методов того времени тем, что человеку было необходимо взаимодействовать с компьютерной программой (решающее правило можно построить и без использования компьютера). Методы такого типа получили название *интерактивных*, или *диалоговых*, процедур. Применение таких процедур показало, что их основная особенность состоит не в интерактивности (сейчас взаимодействие с компьютером — правило, а не исключение), а в *итеративности*, т.е. в том, что эти процедуры основаны на итерациях, в которых перемежаются действия человека и работа некоторой компьютерной программы, решающей вспомогательную задачу, обычно задачу оптимизации. Итеративная процедура строится таким образом, что человек анализирует результаты, полученные компьютером на очередной итерации, и высказывает свои предпочтения, которые реализуются в виде параметров задачи, решаемой на следующей итерации.

Итак, рассматривается задача (2.3). В соответствии с общей теорией МКО, для поиска эффективного решения можно использовать некоторую свертку критериев. Итеративный метод в этом случае состоит в том, что ЛПР (прямо или косвенно) задает параметры свертки, которые передаются компьютеру для решения соответствующей задачи оптимизации.

Опишем эту процедуру формально. Пусть каким-то образом выбрана функция свертки $\psi_\lambda : \mathbb{R}^m \rightarrow \mathbb{R}^1$, зависящая от вектора параметров λ . Тогда каждая итерация состоит из двух шагов — компьютерного и человеческого:

- 1) При некотором λ компьютер решает задачу оптимизации свертки, т.е. задачу поиска $\tilde{x} \in X$, для которого $\psi_\lambda(\varphi(\tilde{x})) = \max_{x \in X} \psi_\lambda(\varphi(x))$, а также, возможно, рассчитывает некоторую вспомогательную информацию;
- 2) ЛПР анализирует $\varphi(\tilde{x})$ и вспомогательную информацию ¹⁾ и выражает свои предпочтения путем изменения вектора параметров λ .

Процедура продолжается до тех пор, пока человек не будет удовлетворен полученным решением или, быть может, просто устанет. Возможно использование формализованных правил остановки процесса.

Здесь приведена простейшая форма итерации, в некоторых методах шаги разбиты на части.

10.1. Простейшие итеративные методы

Рассмотрим итеративную процедуру, прежде всего приходящую в голову начинающим исследователям.

¹⁾Заметим, что при желании ЛПР может получить и само решение $\tilde{x}$.

Итеративное назначение весов в линейной свертке

В этой процедуре используется максимизация простейшей линейной свертки критериальных функций $\sum_{i=1}^m c_i^{(k)} \varphi_i(x)$. Итерации процедуры выглядят следующим образом.

Начальная итерация. Находится идеальная точка y^* и выбираются произвольные значения весов, например, $c_i^{(0)} \equiv 1, i = 1, \dots, m$.

(k+1)-я итерация. Перед началом этой итерации должны быть заданы положительные веса $c_i^{(k)}, i = 1, \dots, m$.

шаг 1. Компьютер решает задачу поиска $\max_{x \in X} \sum_{i=1}^m c_i^{(k)} \varphi_i(x)$, определяя точку максимума $x^{(k+1)}$.

шаг 2. ЛПР сравнивает $\varphi(x^{(k+1)})$ с идеальной точкой y^* и, быть может, точками, найденными на предыдущих итерациях $\varphi(x^{(l)}), l = 1, \dots, k$. Если решение удовлетворяет ЛПР, то процедура завершена. В противном случае ЛПР назначает новые веса $c_i^{(k+1)}, i = 1, \dots, m$, после чего данная итерация завершается, и осуществляется переход к следующей итерации.

Итерации продолжаются до тех пор, пока не будет достигнут результат, удовлетворяющий ЛПР.

Проанализируем этот метод. Прежде всего заметим, что получаемые решения $(x^{(k)})$ являются оптимальными по Парето. Согласно исследованиям возможностей человека во взаимодействии с компьютером, человек способен указать, значения какого критерия он хотел бы увеличить в критериальной точке $\varphi(x^{(k)})$, а какого уменьшить. Однако в данной процедуре от него требуется значительно больше — изменить веса. Это совсем другая задача, и она трудна для ЛПР, поскольку последствия изменения весов непредсказуемы. Например, ЛПР не может знать, на сколько нужно увеличить вес какого-либо критерия для того, чтобы значение критерия стало удовлетворительным. Таким образом, этот логически простой метод является сложным для ЛПР.

Рассмотрим еще один традиционный итеративный метод, нашедший свое применение в практических задачах.

Метод уступок

шаг 0. На предварительном шаге ЛПР ранжирует частные критерии вектор-функции $\varphi(x)$ в порядке убывания их важности. Перенумеровав после этого критерии, можно считать, что первый критерий — самый важный.

шаг 1. Решая задачу поиска $\max_{x \in X} \varphi_1(x)$, находим точку $\tilde{x}^{(1)}$, наилучшую по первому критерию.

шаг 2. ЛПР назначает уступку Δ_1 по первому критерию.

шаг 3. Решается задача поиска $\max_{x \in X, \varphi_1(x) \geq \varphi_1(\tilde{x}^{(1)}) - \Delta_1} \varphi_2(x)$. Решение $\tilde{x}^{(2)}$ максимизирует второй критерий, удовлетворяя дополнительному ограничению, при выполнении которого уступка по первому критерию ограничена ЛПР.

Далее делается дополнительная уступка по второму критерию и т.д., пока не дойдем до последнего критерия. После этого можно вернуться к уступке по первому критерию и т.д. до тех пор, пока не будет получено решение, удовлетворяющее ЛПР. Заметим, что на практике метод уступок завершается после однократного перебора всех критериев.

Отметим, что метод уступок при нулевых уступках совпадает с лексикографическим методом решения задач МКО, в котором сначала максимизируется самый важный критерий, далее — второй по важности (в надежде, что есть какие-то резервы для этого) и т.д. Лексикографическая процедура, вообще говоря, некорректна с вычислительной точки зрения, поскольку малые возмущения параметров исходной задачи могут привести к существенному изменению решения задачи. Хотя и были разработаны устойчивые вычислительные методы²⁾ решения этой задачи, использование лексикографической процедуры, не устойчивой к исходным данным, не может быть рекомендовано для решения задач МКО.

Метод уступок требует от ЛПР ответов на сложные вопросы — необходимо определять уступки, не имея информации о последствиях этого шага. В то же время, концепция ограничений, накладываемых на значения критериев, использованная в этом методе, оказалась весьма полезной и удобной. Были предложены альтернативные методы, базирующиеся на применении ограничений. Далее рассмотрен метод STEM, в котором метод назначения ограничений на значения критериев реализуется в виде, более простом для ЛПР.

Понимание трудностей прямого назначения параметров свертки в традиционных методах, описанных выше, позволила выработать фундаментальное понятие теории итеративных процедур — понятие *структуризованных методов*. К ним относят методы, основанные на том, что ЛПР не определяет параметры задачи оптимизации (например, веса или уступки), а отвечает на вопросы о своих предпочтениях. К таким вопросам относится, например, вопрос о том, какой из двух критериальных векторов более предпочтителен. Другой тип вопросов о предпочтениях — это вопросы о том, значения какого критерия ЛПР желает увеличить, а какого можно уменьшить. Методы, в которых человеку прямо задаются вопросы о параметрах задачи оптимизации, получили название *неструктуризованных методов*, к ним относятся описанные выше методы. Рассмотрим пример структуризованного метода.

²⁾В.В. Федоров. Численные методы максимина. М.: Наука, 1979.

10.2. Метод Джоффриона-Дайера-Файнберга

Метод Джоффриона-Дайера-Файнберга (GDF) основан на использовании идеи известного метода Фрэнка-Вулфа (см. [14]) решения выпуклых задач оптимизации, состоящего в следующем (см. рис. 10.1). Пусть на выпуклом множестве $X \subset \mathbb{R}^n$ нужно найти максимум достаточно гладкой вогнутой скалярной функции $f(x)$.

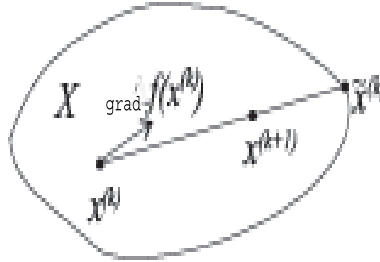


Рис. 10.1.

Пусть точка $x^{(k)}$ — текущее приближение к решению задачи оптимизации. Находим $\text{grad } f(x^{(k)})$ и решаем вспомогательную задачу

$$\max_{x \in X} \langle \text{grad } f(x^{(k)}), x \rangle,$$

заменяя тем самым задачу максимизации нелинейной функции задачей максимизации линейной функции на выпуклом множестве. Пусть максимум достигается в точке $\tilde{x}^{(k)}$. Соединяем точки $x^{(k)}$ и $\tilde{x}^{(k)}$ отрезком прямой и ищем $\max_{x \in [x^{(k)}, \tilde{x}^{(k)}]} f(x)$. Полученная в результате точка $x^{(k+1)}$ рассматривает как новое приближение. Далее переходим к новой итерации. Преимущество этого метода состоит в том, что он довольно быстро сходится.

Перейдем к задаче многокритериальной оптимизации (2.3). Предположим, что функции $\varphi_i(x), i = 1, \dots, m$, вогнуты и дифференцируемы, а у ЛПР имеется представление о своих предпочтениях в виде вогнутой дифференцируемой функции полезности $U(y)$. Проблема состоит в том, чтобы найти $\max_{x \in X} U(\varphi(x))$, где функция $U(y)$ не известна исследователю. В методе Джоффриона-Дайера-Файнберга для решения этой задачи предлагается применить итеративную процедуру, основанную на методе Фрэнка-Вулфа. Это позволяет не строить функцию полезности во всем пространстве.

Пусть $x^{(k)}$ — текущая точка. Найдем аппроксимацию $\text{grad } U(\varphi(x))|_{x=x^{(k)}}$. Отметим, что для дифференцируемых функций $U(y)$ и $\varphi(x)$ имеет место $\frac{\partial U}{\partial x_j} = \sum_{i=1}^m \frac{\partial U}{\partial y_i} \frac{\partial \varphi_i}{\partial x_j}$. Производные $\frac{\partial \varphi_i}{\partial x_j}$ можно определить, так как функции $\varphi_i(x)$ заданы. Производные $\frac{\partial U}{\partial y_i}$ можно найти, используя процедуру аппроксимации поверхностей безразличия функции полезности в точке $y^{(k)} =$

$\varphi(x^{(k)})$ критериального пространства точно так же, как это делалось при построении аддитивной функции полезности в §8.

Дадим описание алгоритма.

Перед началом $(k+1)$ -й итерации должны быть заданы $x^{(k)}$ и $y^{(k)} = \varphi(x^{(k)})$.

(k+1)-я итерация

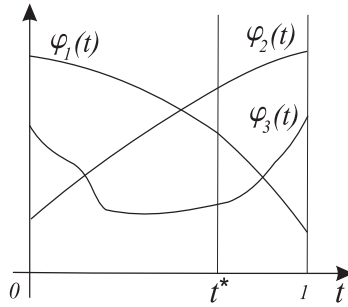
шаг 1. ЛПР строит $\text{grad } U(y^{(k)})$ — градиент функции полезности в точке $y^{(k)}$.

шаг 2. Компьютер ищет $\tilde{x}^{(k)}$ как решение задачи

$$\max_{x \in X} \langle \text{grad } U(\varphi(x^{(k)})), x \rangle$$

. При этом $\text{grad } U(\varphi(x^{(k)}))$ находится как градиент суперпозиции функций $U(y)$ и $y = \varphi(x)$.

шаг 3. На экране компьютера изображается график, на оси абсцисс которого изображен отрезок $[x^{(k)}, \tilde{x}^{(k)}]$, параметризованный переменной t : $t = 0$ в точке $x^{(k)}$ и $t = 1$ в точке $\tilde{x}^{(k)}$; по оси ординат откладываются функции $\varphi_i(t)$, $i = 1, \dots, m$ (см. рис. 10.2). ЛПР указывает предпочтительную точку t^* , которая интерпретируется как точка максимума функции $U(t)$ на отрезке $[x^{(k)}, \tilde{x}^{(k)}]$. Точка $x^{(k+1)} = t^* \tilde{x}^{(k)} + (1 - t^*)x^{(k)}$ считается результатом итерации.



Об. 10.2.

Процедура считается завершенной, когда $\tilde{x}^{(k)} = x^{(k)}$ или ЛПР указывает точку $t^* = 0$, т.е. отказывается передвинуть решение из текущей точки.

Таким образом, в методе Джоффриона-Дайера-Файнберга человек решает две задачи: строит градиент функции полезности и указывает максимум функции полезности на отрезке. Обе эти задачи являются сложными для человека в том случае, когда число критериев превышает два. Процедура построения градиента функции полезности, основанная на многократном попарном сравнении критериальных уступок, является сложной, особенно если учесть необходимость применять такую процедуру на каждой

итерации. Одномерный выбор, когда человек обязан указать лучшую среди бесконечного числа многокритериальных альтернатив, также непрост, хотя в данном случае ситуация сильно упрощается за счет использования графики. В целом можно сказать, что в этом методе ЛПР рассматривается как автомат, способный измерять градиент функции полезности и находить оптимальную точку при одномерной многокритериальной оптимизации.

10.3. Требования к структуризованным итеративным методам

В целях анализа качества структуризованных итеративных методов была сформулирована система требований, которым эти методы, вообще говоря, должны удовлетворять. В рамках этих требований предполагается, что ЛПР имеет функцию полезности $U(y)$, не известную нам. Обозначим через $\tilde{x}$ предпочтительную для ЛПР точку на множестве X , т.е. решение задачи $\max_{x \in X} U(\varphi(x))$. Методы должны обеспечивать:

- 1) сходимость процесса к предпочтительной точке, т.е. $x^{(k)} \rightarrow \tilde{x}$, где $x^{(k)}$ — решения, получаемые на соответствующих итерациях при логичном поведении ЛПР, при котором ответы даются в соответствии с функцией полезности;
- 2) сходимость должна быть достаточно быстрой;
- 3) наличие у ЛПР права на ошибку, т.е. в случае неправильного ответа ЛПР должна иметься возможность в дальнейшем искупить ошибку своими правильными действиями. У ЛПР должна быть возможность найти разумное решение даже в случае многократных ошибок. Ошибки могут замедлять работу процедуры, но решение должно быть получено в соответствии с предпочтениями ЛПР;
- 4) вопросы, задаваемые ЛПР, должны быть простыми.

Проанализируем процедуру Джоффриона-Дайера-Файнберга. Она является структуризованной, поскольку человек отвечает на вопросы, связанные с его предпочтениями относительно альтернативных вариантов решения. Если ЛПР действует в соответствии со своей полезностью, эта процедура обеспечивает сходимость к предпочтительному решению. Так как метод Фрэнка-Вулфа быстро находит точку, близкую к оптимальной, метод Джоффриона-Дайера-Файнберга сходится быстро. Выполняется и требование о праве на ошибку. Единственная проблема метода — сложность вопросов, задаваемых ЛПР. Именно поэтому процедура не нашла широкого применения.

Глава 11. Некоторые итеративные методы МКО

В этом и двух следующих параграфах рассмотрено несколько методов МКО, заложивших основы для дальнейшего совершенствования методов. Надо отметить, что имеется несколько основных идей, на основе которых строятся итеративные методы. Прежде всего, это

- использование градиентов (например, метод Джоффриона-Дайера-Файнберга),
- использование весовых коэффициентов,
- использование целевых точек,
- использование ограничений.

Мы уже рассмотрели структуризованный метод, основанный на использовании градиентов. Перейдем к методам других групп.

11.1. Процедура Зайонца-Валлениуса

Как уже говорилось, человеку трудно непосредственно назначить веса критериев. Метод Зайонца-Валлениуса — это, по существу, попытка задать человеку вопросы о предпочтениях таким образом, чтобы на основе его ответов можно было бы уменьшить неопределенность, связанную с весами. Рассматривается линейная задача МКО

$$y \rightarrow \max, \quad y = Cx, \quad X = \{x \in \mathbb{R}^n \mid Ax \leq b\}. \quad (11.1)$$

Предполагается, что функция полезности также линейна, т.е. $U(y) = \sum_{i=1}^m \lambda_i y_i$, $\lambda_i > 0$, где λ_i — неизвестные веса, не зависящие от y . Веса λ_i нормированы, а именно $\sum_{i=1}^m \lambda_i = 1$. Считается, что перед очередной итерацией известен конус, которому принадлежит вектор весов λ . Задача итерации заключается в том, чтобы сузить этот конус. Для сужения конуса используются результаты сравнения точек в критериальном пространстве.

На начальной итерации кладется $\Lambda^{(0)} = \text{int } \mathbb{R}_+^m$. Перед $(k+1)$ -й итерацией известен конус $\Lambda^{(k)} \subset \Lambda^{(0)}$. Итерация состоит из следующих шагов.

(k+1)-я итерация

шаг 1. Выбирается некоторый вектор $\lambda^{(k)} \in \Lambda^{(k)}$ и решается задача линейного программирования: $\max_{x \in X} \langle \lambda^{(k)}, Cx \rangle$. Помимо нахождения точки $\tilde{x}^{(k)}$, оптимальной для данного вектора $\lambda^{(k)}$, рассчитываются вершины $x^{(k,p)}$, $p = 1, \dots, P^{(k)}$, соседние с $\tilde{x}^{(k)}$.

шаг 2. ЛПР отвечает на вопросы о том, что предпочтительнее, $\varphi(\tilde{x}^{(k)})$ или $\varphi(x^{(k,p)})$ для всех $p = 1, \dots, P^{(k)}$. ЛПР может ответить “лучше”, “хуже”, “эквивалентны”. В соответствии с этими ответами формируются дополнительные ограничения на вектор $\lambda \in \mathbb{R}_+^m$:

если $\varphi(\tilde{x}^{(k)}) \succ \varphi(x^{(k,p)})$, то $\langle \lambda, \varphi(\tilde{x}^k) \rangle > \langle \lambda, \varphi(x^{(k,p)}) \rangle$,

если $\varphi(\tilde{x}^{(k)}) \prec \varphi(x^{(k,p)})$, то $\langle \lambda, \varphi(\tilde{x}^k) \rangle < \langle \lambda, \varphi(x^{(k,p)}) \rangle$,

если $\varphi(\tilde{x}^{(k)}) \sim \varphi(x^{(k,p)})$, то $\langle \lambda, \varphi(\tilde{x}^k) \rangle = \langle \lambda, \varphi(x^{(k,p)}) \rangle$.

Неравенства порождают систему ограничений $D^{(k)}\lambda > 0$, а равенства — систему $D_0^{(k)}\lambda = 0$.

шаг 3. Конус сужается: $\Lambda^{(k+1)} = \Lambda^{(k)} \cap \{\lambda | D^{(k)}\lambda > 0\} \cap \{\lambda | D_0^{(k)}\lambda = 0\}$.

В методе используется следующее правило остановки. Если все соседние вершины множества X оказались хуже, чем $\tilde{x}^{(k)}$, то, в силу линейности задачи, это означает, что найдена оптимальная вершина.

В реальном процессе остановка часто происходит раньше, в момент, когда $\Lambda^{(k+1)}$ становится пустым (это означает, что человек давал противоречивые ответы). Для исключения этого недостатка авторы процедуры сначала устранили условие эквивалентности как наиболее чувствительное к ошибкам. Этого оказалось недостаточно. Тогда, прежде чем предлагать сравнение точек, авторы стали проверять, не может ли привести задаваемый вопрос к противоречию. Устранив вопросы, которые помогают контролировать правильность предположения о линейности функции полезности и непротиворечивость ЛПР в рамках этого предположения, авторы сделали процедуру работоспособной.

Анализ метода. Если человек не дает противоречивых ответов (или порождающие их вопросы исключены), то процедура работает нормально, и метод довольно быстро сходится к предпочтительной точке. Вопросы являются сложными, так как происходит многократное сравнение двух многомерных альтернатив. Заметим, что при этом нет права на ошибку.

Несмотря на свои недостатки и ограниченность области применения (требуется линейность функции полезности!), процедура оказала большое влияние на развитие методов МКО.

11.2. Метод Штойера

Как и в предыдущем методе, рассматривается линейная задача МКО, однако линейность функции полезности уже не обязательна. Линейная свертка используется в методе лишь для аппроксимации предпочтений.

Предполагается, что перед началом $(k+1)$ -й итерации известны такие величины $p_i^{(k)}$, $q_i^{(k)}$, что $\lambda_i \in [p_i^{(k)}, q_i^{(k)}]$, $i = 1, \dots, m$.

(k+1)-я итерация.

шаг 1. Генерируется большое число (порядка $N = 50m$) наборов весов $\lambda^{(l)} = (\lambda_1^{(l)}, \dots, \lambda_m^{(l)})$, $l = 1, \dots, N$, равномерно распределенных на

$[p^{(k)}, q^{(k)}]$, где $p^{(k)} = (p_1^{(k)}, \dots, p_m^{(k)})$, $q^{(k)} = (q_1^{(k)}, \dots, q_m^{(k)})$.

Точки нормируются, т.е. $\sum_{i=1}^m \lambda_i^{(l)} = 1$, $l = 1, \dots, N$.

шаг 2. Происходит фильтрация точек так, чтобы осталось порядка $L = 3m$ точек $\tilde{\lambda}^{(l)}$, $l = 1, \dots, L$, достаточно хорошо представляющих множество $[p^{(k)}, q^{(k)}]$.

шаг 3. Решаются задачи $\max_{x \in X} \langle \tilde{\lambda}^{(l)}, Cx \rangle$ для всех $\tilde{\lambda}^{(l)}$, $l = 1, \dots, L$, в результате чего определяются соответствующие точки максимума $\tilde{x}^{(k,l)}$.

шаг 4. Рассматривая векторы $\varphi(\tilde{x}^{(k,l)})$, $l = 1, \dots, L$, ЛПР выбирает наиболее предпочтительную из точек $\tilde{x}^{(k,l)}$, $l = 1, \dots, L$, и соответствующий вектор $\tilde{\lambda}^{(l)}$, который обозначается через $\lambda^{(k+1)}$.

шаг 5. Строится новый параллелепипед $[p^{(k+1)}, q^{(k+1)}]$, где

$$p^{(k+1)} = \lambda^{(k+1)} - \frac{r^k}{2}, q^{(k+1)} = \lambda^{(k+1)} + \frac{r^k}{2}.$$

Здесь $r \in (0, 1)$ — число, характеризующее скорость сжатия (задается заранее).

Анализ метода. Явный недостаток метода состоит в том, что сжатие параллелепипеда происходит со скоростью, не зависящей от ответов ЛПР, поэтому процедура может и не сойтись к наиболее предпочтительному решению (при слишком быстром сжатии оно может быть просто отсечено). Однако если сходимости имеет место, то при быстром сжатии она достаточно быстрая. С правом на ошибку ситуация не очень ясна, так как какую-то ошибку можно исправить, а какую-то ошибку — нет (многое зависит от скорости сжатия параллелепипеда). Вопросы сложные, поскольку ЛПР нужно выбирать из $3m$ многокритериальных альтернатив (причем достаточно много раз).

Метод Штойера-Чу для нелинейных задач

Метод предназначен для решения нелинейных задач МКО и использует идеи метода Штойера. В нелинейном случае предлагается использовать чебышевские свертки, связанные с идеальной точкой — метод Штойера-Чу отличается от метода Штойера только тем, что вместо линейной функции с весами используется функция Чебышева $\rho_\lambda(y, y^*) = \max_{i=1, \dots, m} (\lambda_i |y_i^* - y_i|)$. При этом сохраняются все трудности метода Штойера и добавляются новые, связанные с нелинейностью задачи.

11.3. Методы с целевыми точками

Эти методы, предназначенные для общего случая множества X и вектор-функции φ , основаны на минимизации некоторого расстояния от точек множества $Y = \varphi(X)$ до целевых множеств $Z = \hat{y} + \mathbb{R}_+^m$.

Метод 1.

На нулевой итерации в качестве целевой берется идеальная точка. Перед $(k+1)$ -й итерацией целевая точка $\hat{y}^{(k)}$ должна быть уже найдена.

(k+1)-я итерация.

шаг 1. Компьютер ищет точку $\tilde{x}^{(k)}$, являющуюся решением задачи

$$\min_{x \in X} \rho(\varphi(x), Z^{(k)}), \text{ где } Z^{(k)} = \hat{y}^{(k)} + \mathbb{R}_+^m.$$

шаг 2. ЛПР сравнивает $\varphi(\tilde{x}^{(k)})$, $\hat{y}^{(k)}$ и y^* , после чего назначает новую целевую точку $\hat{y}^{(k+1)}$.

Если ЛПР недоволен значением j -й координаты вектора $\varphi(\tilde{x}^{(k)})$, то при назначении целевой точки ему рекомендуется увеличить значение соответствующей координаты $\hat{y}^{(k+1)}$.

Метод 2.

В этом методе применяется расстояние

$$\rho_\mu(y, y^*) = \sum_{i=1}^m \mu_i \frac{y_i^* - y_i}{y_i^* - y_i^{\min}},$$

где величины $y_i^{\min}$ могут быть выбраны различными способами, скажем, $y_i^{\min} = \min_{y \in Y} y_i$, или $y_i^{\min} = \min_{y \in P(Y)} y_i$. Величины $y_i^{\min}$ также могут быть назначены ЛПР.

На нулевой итерации используются исходные веса $\mu_i^{(0)} \equiv 1$. Перед $(k+1)$ -й итерацией должен быть известен k -й вариант весов $\mu_i^{(k)}$, $i = 1, \dots, m$, свертки $\rho_\mu(y, y^*)$.

(k+1)-я итерация.

шаг 1. Минимизируется расстояние с весами $\mu^{(k)}$ от точек множества Y до идеальной точки y^* , т.е. находится точка $\tilde{x}^{(k)}$, являющаяся решением задачи $\min_{x \in X} \rho_\mu(\varphi(x), y^*)$.

шаг 2. ЛПР сравнивает $\varphi(\tilde{x}^{(k)})$ и y^* , после чего назначает новые веса $\mu_i^{(k+1)}$ свертки $\rho_\mu(y, y^*)$.

Проанализируем эти два метода. Обе описанные процедуры не структуризованы, поэтому говорить о сходимости не приходится. Отметим, однако, что эти методы обладают положительным свойством — человек имеет право на ошибку. В то же время, назначение целей и весов не являются простыми задачами, поэтому на практике ЛПР обычно предпочитает однократно назначить целевую точку. При этом метод вырождается в

рассмотренный ранее эвристический метод задания предпочтения в форме назначения целевой точки.

Отметим, что существует модификация первого метода — назначение многих целевых точек.

11.4. Метод STEM

Метод STEM (STEP Method) является одним из первых итеративных методов. Он был предложен в конце 60-х годов XX века группой исследователей, которая включала будущего академика РАН О.И. Ларичева (1934–2003) и троих французских специалистов — Р. Бенайюна, Ж. Монгольфье и Ж. Терни. Метод STEM нашел многочисленных последователей, которые модифицировали его различным образом. Сначала мы рассмотрим вариант, близкий к тому, который был предложен с самого начала, а потом поговорим о его модификации.

Рассматривается задача МКО общего типа (2.3). Идея метода STEM состоит в том, чтобы задавать ЛПР как можно более простые вопросы относительно его предпочтений. Ответы используются для изменения ограничений на значения критериев типа $\varphi_j(x) \geq l_j, j = 1, \dots, m$, а также для расширения списка $J \subseteq \{1, \dots, m\}$ критериев, по которым уже достигнуто удовлетворительное значение. Изменения ограничений, однако, должно осуществляться в более удобной для ЛПР форме, чем в методе уступок.

Предполагается, что выполнено техническое ограничение на идеальную точку $y^* > 0$. На начальной итерации алгоритма кладется $l_j^{(0)} = -L, j = 1, \dots, m$, где L — большое число, $J^{(0)} = \emptyset$. В результате k -й итерации алгоритма должны быть определены $l_j^{(k)}, j = 1, \dots, m$ и $J^{(k)}$, где $l_j^{(k)}$ — критериальные ограничения, используемый на $(k+1)$ -й итерации, $j = 1, \dots, m$, а $J^{(k)}$ — множество тех номеров критериев, для которых уровень $l_j^{(k)}$ является удовлетворительным.

На $(k+1)$ -й итерации осуществляется следующая последовательность действий.

(k+1)-я итерация

шаг 1. С помощью компьютера решается серия m задач оптимизации

$$\max_{x \in X} \left\{ \varphi_i(x) \mid \varphi_j(x) \geq l_j^{(k)}, j = 1, \dots, m \right\}, i = 1, \dots, m,$$

и находятся соответствующие точки максимума¹⁾. На основе этой информации некоторым образом с учетом $J^{(k)}$ рассчитываются веса $\lambda_i^{(k) > 0}$ (предложено несколько вариантов расчета весов, на которых

¹⁾Заметим, что на первой итерации, когда $l_j^{(0)} = -L, j = 1, \dots, m$, при этом находятся решения, оптимальные по частным критериям. Они позволяют рассчитать идеальную точку y^* .

мы не будем останавливаться) и определяется $x^{(k)}$ — решение задачи оптимизации

$$\max_{x \in X, \varphi_j(x) \geq l_j^{(k)}, j=1, \dots, m} \sum_{i=1}^m \lambda_i^{(k)} \varphi_i(x),$$

являющееся оптимальным по Парето решением задачи МКО. На этом первый шаг процедуры завершается.

шаг 2. ЛПР получает информацию о $\varphi(x^{(k)})$, $J^{(k)}$, $l^{(k)}$ и y^* , на основе которой ЛПР усиливает ограничения и, может быть, расширяет множество номеров критериев, имеющих удовлетворительные значения. Пусть ЛПР не устраивает более всего значение критерия с номером i^* . Тогда он кладет

$$l^{(k+1)} = \begin{cases} l^{(k)}, & i \neq i^* \\ l_{i^*}^{(k)} + \Delta, & i = i^*, \end{cases}$$

где Δ — некоторое положительное число, которое ЛПР определяет заранее. Далее, ЛПР может расширить множество $J^{(k+1)}$. Итерация окончена.

Рассмотрим этот метод с точки зрения четырех требований к итеративным методам.

1. Анализ сходимости невозможен, так как непонятно, как определяется Δ . Более того, выбор Δ — неструктуризованный элемент метода. Эта величина может оказаться слишком большой, в результате чего наиболее предпочтительное решение будет отсечено.
2. Нет права на ошибку — отсеченное наиболее предпочтительное решение вернуть невозможно.
3. Вопросы типа классификации критериев, задаваемые ЛПР, просты. Это и есть основное достоинство метода, особенно учитывая то, что это был первый метод, в котором внимание уделялось именно простоте вопросов. В то же время, процедура включает не только простые вопросы (вопрос о назначении Δ является сложным).

Модификация метода STEM

Профессор Корнелльского университета (США) Д.П. Лаукс, известный специалист в области поддержки принятия решений в сфере водного хозяйства, заметил, что постоянное сужение множества допустимых значений критериев — недостаток метода STEM и предложил модификацию метода, не имеющую этого недостатка.

В модифицированном методе на первой итерации наряду с идеальной точкой y^* находится некоторое решение x^0 , оптимальное по Парето, и соответствующий недоминируемый критериальный вектор, который рассматривается ЛПР. Если ЛПР не удовлетворен решением, то требуется указать, значение какого критерия y_{i^*} не удовлетворительно в наибольшей степени и по каким критериям $y_{j_1}, y_{j_2}, \dots, y_{j_k}$ можно идти на уступки, а также и величины уступок $\Delta_1, \dots, \Delta_k$, на которые можно пойти для увеличения значения y_{i^*} . Далее на основе решения задачи

$$y_{i^*} \rightarrow \max, \quad x \in X, \quad \varphi_{j_l}(x) \geq \varphi_{j_l}(x^0) - \Delta_l, \quad l = 1, \dots, k,$$

компьютер рассчитывает новое решение $x^{(1)}$ и новую недоминируемую по Слейтеру критериальную точку $\varphi(x^{(1)})$, после чего ЛПР снова определяет наименее удовлетворительный критерий и т.д. Таким образом, исключается один из главных недостатков метода STEM — отсутствие права на ошибку.

Поскольку назначаемые величины уступок являются, по существу, пробными, не приводящими к фатальным ошибкам, это делает их назначение не столь важным и, соответственно, более простым шагом. Таким образом, был получен относительно простой метод, предназначенный для использования на практике.

Примером применения этого метода служит разработка одной водохозяйственной системы в Алжире, осуществленная Д.П. Лауксом совместно с алжирскими специалистами еще в 70-х гг. XX века. Лица, принимающие решение, находились в Алжире, а суперкомпьютер, который мог решать соответствующую задачу оптимизации — в Корнелльском университете, США. Поскольку компьютерные сети в то время отсутствовали, передача данных, необходимых для формулировки очередной задачи оптимизации, осуществлялась по обычному телефону. Алжирцы, получив от Д.П. Лаукса решение, оптимальное по Парето, и соответствующие значения критериев, анализировали их и формировали критериальные уступки, которые передавали в Корнелльский университет по телефону. Профессор Лаукс вводил их в компьютер, решал задачу оптимизации и передавал новое решение и значения критериев в Алжир. В результате за несколько итераций было найдено решение, оптимальное по Парето, которое удовлетворило ЛПР в Алжире.

Глава 12. Итеративные методы визуализации кривых на границе Парето

В последние годы все большее распространение получает идея о том, что многокритериальную информацию нужно предоставлять человеку в графическом виде. До начала массового использования персональных компьютеров это направление сдерживалось сложностью вывода информации. После появления персональных компьютеров и разработки графических пакетов технические проблемы исчезли, поэтому, начиная с середины 80-х годов прошлого века, стало быстро развиваться новое направление итеративных процедур, связанное с визуализацией информации. Некоторые примеры итеративных процедур, использующих визуализацию, рассмотрены в этом параграфе. В них изучается выпуклая задача многокритериальной максимизации.

12.1. Метод проецирования на границу Парето

Этот метод использует некоторые идеи метода Джоффриона-Дайера-Файнберга (GDF) и является одним из первых многокритериальных итеративных методов, основанных на визуализации. В отличие от метода GDF, все текущие решения оптимальны по Парето, а градиент функции полезности проецируется на границу Парето. Далее ЛПР указывает предпочтительную точку на этой проекции. Из метода GDF взяты способы определения градиента функции полезности и поддержки выбора точки одномерной кривой в пространстве критериев.

Пусть уже известен градиент функции полезности в текущей недоминируемой критериальной точке $y^{(k)} = \varphi(x^{(k)})$. Рассмотрим точку $\tilde{y}^{(k+1)} = \varphi(\tilde{x}^{(k)})$, которая является решением задачи максимизации линейной функции, заданной градиентом функции полезности. Далее, в отличие от метода GDF, отрезок $[y^{(k)}, \tilde{y}^{(k)}]$ проецируется на границу Парето. Для двумерного случая это показано на рис. 12.1

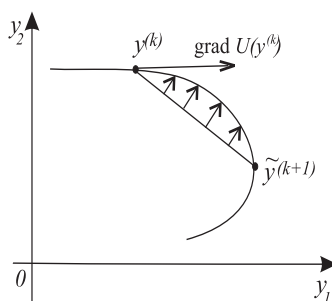


Рис. 12.1.

Далее компьютер изображает график изменения критериальных величин вдоль линии проекции, параметром которой является переменная t . Он аналогичен графику, используемому в методе GDF, за исключением того, что все точки графика теперь являются недоминируемыми. Благодаря этому ЛПР информируется только о недоминируемых точках (заметим, что многие считают, что демонстрация доминируемых точек дезориентирует ЛПР).

Как и в методе GDF, ЛПР выбирает t^* — такое значение параметра t , при котором достигается наилучшее сочетание значений критериев; после этого кладется $y^{(k+1)} = t^* \hat{y}^k + (1 - t^*) y^{(k)}$.

О недостатках и достоинствах этого метода в рамках четырех требований, сформулированных выше, можно сказать то же самое, что и о методе GDF — определение градиента функции полезности является крайне сложной задачей. Метод обладает и другим недостатком метода GDF — нахождение t^* также является непростой задачей, а многократное решение такой задачи слишком сложно для ЛПР. Требования права на ошибку и быстрой сходимости метода выполняются.

12.2. Метод Корхонена-Лааксо

Главное отличие метода Корхонена-Лааксо от предыдущего состоит в том, что в нем отсутствуют трудности, связанные с построением градиента функции полезности. Для этого вместо градиента используется другое направление в критериальном пространстве, которое строится на основе идей целевого программирования.

Перед началом $(k+1)$ -й итерации должна быть задана точка $y^{(k)} \in P(Y)$.

(k+1)-я итерация

шаг 1. ЛПР задает целевую точку $\hat{y}^{(k)}$.

шаг 2. Точки $\hat{y}^{(k)}$ и $y^{(k)}$ соединяются отрезком, который проецируется на границу Парето, после чего на дисплее изображается рисунок, аналогичный рисунку для одномерной оптимизации в методе GDF.

шаг 3. ЛПР выбирает t^* , и, следовательно, $y^{(k+1)}$.

Обсуждая недостатки и достоинства этого метода, можно сказать, что авторам удалось избежать самой сложной проблемы предыдущего метода, поскольку теперь не нужно строить градиент функции полезности. В остальном метод сохраняет свою сложность для ЛПР — задача выбора t^* не проста. К тому же приходится многократно решать сложную задачу назначения целевой точки. Достоинства метода (как и метода GDF) — быстрая сходимость при правильных действиях ЛПР и наличие у ЛПР права на ошибку.

12.3. “Бег по множеству Парето”

На основе идей метода Корхонена-Лааксо в конце 80-х годов прошлого века Корхоненом и Валлениусом был предложен метод “Бег по множеству Парето” (Pareto Race). Основная особенность этого метода состоит в том, что вместо отдельных итераций рассматривается непрерывный процесс изменения целевой точки $\hat{y}$ и текущего положения y . Изменением точки $\hat{y}$ управляет ЛПР, который делает это с помощью простых средств, например, перемещения движков прокруток. Перемещение текущей точки y также осуществляется непрерывно: она перемещается по границе Парето вдоль проекции $[y, \hat{y}]$ со скоростью, задаваемой ЛПР: ЛПР имеет возможность увеличивать и уменьшать скорость или даже приостанавливать процесс движения. Таким образом, в руках ЛПР оказывается средство управления точкой на границе Парето, напоминающее управление автомобилем: направление задается положением точки $\hat{y}$, а скорость — непосредственно кнопками. Текущее положение представляется обычной столбчатой диаграммой, на которой отображаются величины значений критериев.

Итак, метод “Бег по множеству Парето” позволяет свободно передвигаться по границе Парето, последовательно получая точки этой границы. Основным недостатком метода является отсутствие у ЛПР целостного представления о границе Парето при $m > 2$. В связи с этим ЛПР даже при явно осознаваемых предпочтениях может не понимать, куда и как имеет смысл перемещаться. Такого недостатка лишены методы, направленные на информирование ЛПР о границе Парето в целом.

Глава 13. Методы, направленные на информирование ЛПР о границе Парето

Важным классом методов решения задач многокритериальной оптимизации являются методы, основанные на предварительном построении (аппроксимации) границы Парето. Смысл такого подхода состоит в том, чтобы дать ЛПР информацию о границе Парето, не задавая вопросов о предпочтениях, а затем предложить указать предпочтительную точку границы Парето. Единственное, что ЛПР должен сделать заранее — указать направление улучшения значений критериев. Конечно, такое направление должно существовать для каждого из критериев, для чего сразу требуется предположить независимость по предпочтениям. Если требуется, ЛПР может выбрать несколько “интересных” точек границы Парето и получить несколько решений для дальнейшего детального анализа.

Для реализации этого подхода надо разработать методы аппроксимации границы Парето и способы информирования ЛПР об этой аппроксимации. Начнем с обсуждения второго вопроса. В двухкритериальном случае граница Парето может быть представлена в наглядном графическом виде на плоскости. Если в задаче конечное число возможных решений, то множество $Y = \varphi(X)$ является совокупностью точек на плоскости, среди которых можно визуальным образом выделить недоминируемые (для множества точек Y , изображенного на рис. 13.1 а, недоминируемые точки выделены графически). В случае бесконечного числа решений можно аппроксимировать и изобразить множество $Y = \varphi(X)$, причем его границы Парето и Слейтера также можно достаточно легко выделить визуальным образом (см. рис. 13.1 б, где изображено множество Y и графически выделена его граница Парето). В последнем случае можно также аппроксимировать и изобразить только границу Парето или границу Слейтера, помня при этом об их возможной неустойчивости.

В случае $m > 2$ проблема является значительно более сложной. Начнем со случая конечного числа возможных решений.

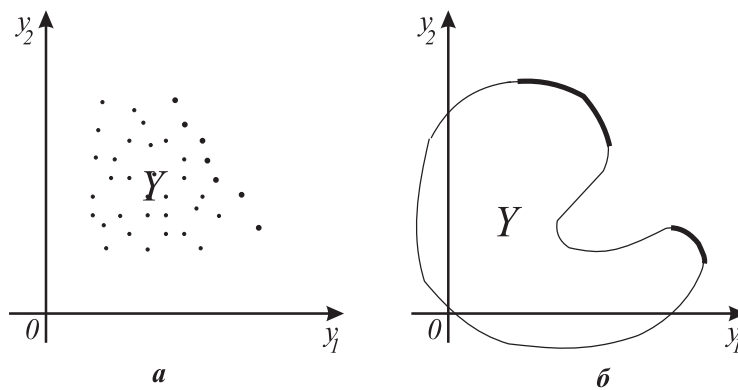


Рис. 13.1.

13.1. Конечное число решений

Существует несколько подходов к визуализации конечного числа критериальных точек при $m > 2$.

Параллельные отрезки (профили альтернатив, value path)

Для начала критерии нормируются так, чтобы их значения находились в отрезке $[0, 1]$. Затем на дисплее компьютера изображаются m параллельных отрезков (по числу критериев), каждый из которых соответствует одному критерию. Для визуализируемой критериальной точки на каждом из отрезков отмечаются соответствующие значения критерия, а затем проводится ломаная линия, которая соединяет значения, соответствующие этой точке, на смежных отрезках (см. рис. 13.2).

Изучая критериальные точки, представленные в виде ломаных, ЛПР может сравнить эти точки для большого числа критериев (до нескольких десятков). В то же время, трудно сравнить между собой более дюжины точек, так что доминируемые точки разумно отсеивать заранее и не изображать.

Для того чтобы упростить исследование, можно отдельные наиболее интересные точки изобразить ломаными ярких цветов на фоне одноцветных ломаных, соответствующих остальным точкам. Это дает возможность рассмотреть несколько ломаных на фоне остальных.

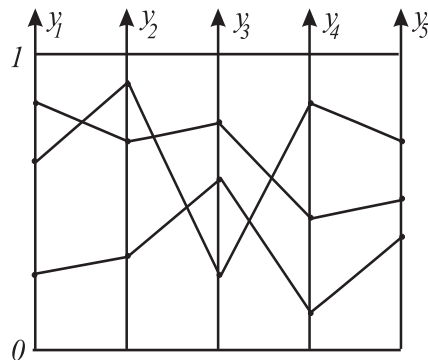


Рис. 13.2.

Радарные диаграммы

Метод радарных диаграмм является модификацией метода параллельных отрезков. В этом случае отрезки направлены в разные стороны, что, по нашему мнению, только затрудняет исследование, хотя, конечно, выглядит более эффектно (см. рис. 13.3). Все замечания, высказанные относительно метода параллельных отрезков, верны и в случае радарных диаграмм.

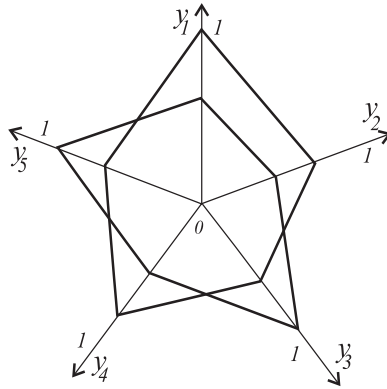


Рис. 13.3.

Подводя итог, можно сказать, что графические методы представления конечного числа альтернатив при $m > 2$ пригодны для анализа относительно небольшого числа альтернатив. Поэтому на практике, если число недоминируемых альтернатив превосходит 10, они представляются ЛПР в виде списка, который наряду со значениями критериев для таких альтернатив может содержать и описание самих альтернатив. По аналогии таким подходом, значительная часть методов аппроксимации границы Парето для задач с бесконечным числом решений направлена на ее аппроксимацию конечным числом точек, которые затем представляются ЛПР в виде списка.

13.2. Аппроксимация границы Парето конечным числом точек

Рассмотрим методы аппроксимации границы Парето конечным числом точек при бесконечном множестве X . Существует несколько подходов для решения этой задачи. Они базируются на использовании весов, ограничений или целей, а также на некоторых других идеях.

Использование весов

Рассмотрим линейную задачу многокритериальной оптимизации (11.1). Из теории сверток следует, что, при заданном векторе $\lambda > 0$, решая задачу линейного программирования

$$\langle \lambda, y \rangle \rightarrow \max \quad (13.1)$$

$$y = Cx, \quad x \in X, \quad (13.2)$$

можно найти точку границы Парето. Меняя $\lambda \in \mathbb{R}^m$, $\lambda > 0$, можно найти все недоминируемые вершины и грани. Впервые эту идею сформулировали известные американские ученые С. Гасс и Т. Саати в 1955 году для случая двух критериев. Прежде всего отметим, что паретова граница для линейной задачи при $m = 2$ есть ломаная, состоящая из конечного числа вершин,

соединенных между собой ребрами. Чтобы ее найти, достаточно найти все недоминируемые вершины и соединить соседние вершины отрезками прямой.

Для нахождения всех вершин С. Гасс и Т. Саати предложили решить задачу параметрического линейного программирования

$$\begin{aligned} \lambda y_1 + (1 - \lambda)y_2 &\rightarrow \max \\ y &= Cx, \quad Ax \leq b, \end{aligned}$$

где λ — параметр, удовлетворяющий соотношению $0 \leq \lambda \leq 1$. Решение задачи параметрического линейного программирования состоит в последовательном поиске таких значений параметра λ , при которых происходит переход из вершины в вершину, а также в расчете новых вершин. Благодаря этому удастся ограничиться конечным числом значений параметра λ вместо рассмотрения его всевозможных неотрицательных значений.

Решение задачи осуществляется следующим образом. Сначала решается задача оптимизации при $\lambda_1 = 1$ и находится $y_1^{(0)} = \max_{x \in X} y_1$. При этом $y_2^{(0)}$ равно максимальному значению y_2 , которое может быть найдено при решении этой задачи оптимизации. Затем определяется такое число $0 < \lambda < 1$, при котором оптимальное решение переходит в соседнюю недоминируемую вершину $y^{(1)}$, и находится эта вершина. Ясно, что $y_1^{(1)} < y_1^{(0)}$ и $y_2^{(1)} > y_2^{(0)}$. Эту операцию можно легко осуществить в рамках симплекс-метода решения задач линейного программирования. Далее находится меньшая величина λ , при которой осуществляется переход из $y^{(1)}$ в некоторую соседнюю недоминируемую вершину $y^{(2)}$, и так далее. Таким образом, последовательно находятся все недоминируемые вершины. Соседние вершины соединяются отрезками прямой. В 1955 г. такая граница изображалась на бумаге, в настоящее время — на экране компьютера. ЛПР получает представление о возможных значениях критериев и, главное, о том, на какую величину требуется ухудшить значение одного из критериев для того чтобы улучшить значение другого (так называемое объективное, или критериальное замещение, по-английски — objective tradeoff). Критериальное замещение можно легко оценить в любой точке границы Парето, просто взглянув на ее изображение.

Отметим, что описанный двухкритериальный метод нашел своих последователей, и в настоящее время его использование можно часто встретить в прикладных, особенно эколого-экономических, исследованиях.

В 70-х годах XX века появилось обобщение этой идеи на случай $m > 2$. Отметим, что решив при заданном векторе λ задачу линейного программирования (13.1)–(13.2), мы получим точку границы Парето, которая является недоминируемой вершиной многогранного множества Y . Можно легко сформулировать необходимое и достаточное условие недоминируемости вершины множества Y , которое дает возможность ввести понятие эффективного базиса в пространстве решений. Для того чтобы найти все недоминируемых вершины, нужно перебрать эффективные базисы. Метод имеет следующие этапы:

- 1) находится произвольная недоминируемая вершина $y^{(0)} \in P(Y)$,
- 2) находятся соседние вершины в $\mathbb{R}^m$,
- 3) выясняется, являются ли они недоминируемыми и не были ли рассмотрены ранее,
- 4) если среди них есть новая недоминируемая вершина, осуществляется переход в эту вершину.

В отличие от двумерного случая, здесь возникают неоднозначность выбора соседней вершины и вопрос о том, каким образом обойти все вершины. Эта задача является сложной с вычислительной точки зрения, особенно учитывая неизбежные ошибки округления. Мы не будем углубляться в эти вопросы, подробно рассмотренные в книге [14].

Остановимся на проблеме информирования ЛПР об уже найденном списке вершин. Отметим, что С. Гасс и Т. Саати находили как вершины, так и ребра, причем граница Парето изображалась целиком. В многомерных методах описанного типа, однако, находятся только вершины, которые представляются ЛПР в виде списка. Недостаток такого подхода к информированию ЛПР состоит в том, что, во-первых, невозможно изучить список, включающий в себя многие тысячи недоминируемых вершин (а число вершин превосходит тысячи уже в самых простых задачах), и, во-вторых, для ЛПР важны не сами вершины, а форма границы, т.е. критериальное замещение, которое не может быть получено из такого списка. Далее, вершины обычно расположены неравномерно, так что они могут не давать представления о значениях критериев на всей границе Парето. Для того чтобы исправить недостатки метода, было предложено по найденным вершинам выяснять, какие из них лежат на одной грани, и выписывать уравнения многомерных граней. Авторы такой идеи надеялись, что благодаря этому ЛПР сможет получить более полную информацию о границе Парето. С этим трудно согласиться — понимание границы Парето, представленной в форме совокупности многомерных граней, чрезвычайно затруднительно для ЛПР. Даже в случае относительно малого числа вершин и граней (скажем, не более нескольких сотен) анализ такой информации практически невозможен.

Подход на основе ограничений

Задачу МКО общего типа (2.3) для начала рассмотрим при $m = 2$. Пусть множество имеет вид, представленный на рис. 13.4. С помощью решения задачи скалярной оптимизации находятся точка $y^0 = (y_1^0, y_2^0)$, являющаяся критериальной точкой с максимально возможным значением критерия y_1 и точка $y^2 = (y_1^2, y_2^2)$, являющаяся критериальной точкой с максимально возможным значением критерия y_2 . Тогда идеальная точка $y^* = (y_1^*, y_2^*)$ находится как $y_1^* = y_1^0, y_2^* = y_2^2$. Отрезок $[y_2^1, y_2^*]$, содержащий границу Парето, разбивается на K частей и решается серия задач

$$y_1 \rightarrow \max \\ y = \varphi(x), \quad x \in X, \quad y_2 \geq a^k,$$

где $a^k = y_2^0 + k(y_2^* - y_2^0)/K, k = 1, \dots, K - 1$. Решая каждую из задач, получаем точку $P(Y)$. Действуя таким образом, можно найти список точек, аппроксимирующих границу Парето. На рис. 13.4 видно, что решения различных задач могут совпадать, так что число найденных точек может быть меньше, чем число решенных задач. При желании, можно решить задачу оптимизации второго критерия при варьировании значения первого и получить дополнительные точки.

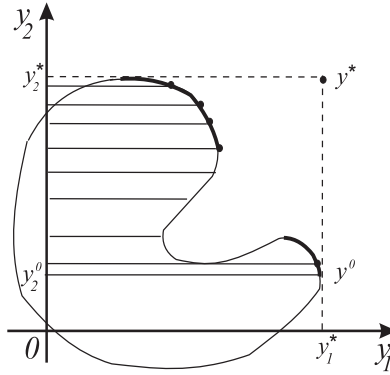


Рис. 13.4.

В многомерном случае прежде всего надо решить проблему нахождения параллелепипеда, содержащего границу Парето. В отличие от двумерного случая, здесь уже недостаточно решить m задач оптимизации отдельных критериев, поскольку в многомерном случае граница Парето может выйти за пределы параллелепипеда, заданного решениями задач оптимизации отдельных критериев.

Пусть эта сложная задача решена и найдена такая точка y^{**} , часто называемая надиром, что $[y^{**}, y^*]$ — минимальный параллелепипед, содержащий границу Парето. Далее оптимизируется один критерий (например, первый), а остальные варьируются. Как и в двухмерном случае, можно поочередно оптимизировать значения каждого из критериев, варьируя значения остальных. В итоге находится список точек недоминируемого множества, который представляется ЛПР. Отметим, что полученные точки могут представлять границу Парето очень неравномерно. Сложность понимания информации ЛПР, представленной в виде длинного списка недоминируемых точек, также затрудняет использование метода ограничений при $m > 2$.

Веса в нелинейном случае

Решается задача $\rho_\lambda(y, y^*) \rightarrow \min, \quad y = \varphi(x), \quad x \in X$. В качестве функции $\rho_\lambda(y, y^*)$ можно взять, например, взвешенную функцию Чебышева

$$\rho(y, y^*) = \max_{i=1, \dots, m} \lambda_i |y_i^* - y_i|.$$

На множестве неотрицательных весов λ_i , сумма которых равна единице, выбирается сетка и решается задача для всех точек сетки, в результате чего находится большое число недоминируемых точек. Важно, что этот метод пригоден и в случае нелинейных невыпуклых моделей. В этом подходе также возникают проблемы неравномерности расположения точек и сложности понимания ЛПР информации в виде огромного списка. Отметим, что в многоэкстремальных задачах решение даже одной задачи (т.е. при фиксированных весах) может вызывать затруднение. Это замечание относится и к методу ограничений.

Другие идеи

Среди других подходов следует отметить методы, основанные на покрытии множества X шарами переменного радиуса, расчете вектор-функции $\varphi(x)$ в центрах шаров и выборе недоминируемых критериальных точек. Радиусы шаров покрытия определяются константами Липшица функций $\varphi_i(x), i = 1, \dots, m$. Такие методы применимы при относительно малых размерностях пространства решений и при наличии оценок констант Липшица, что не всегда возможно. Как и в других методах, получаемый список критериальных точек может вызвать затруднения у ЛПР.

Глава 14. Аппроксимация множества достижимых критериальных векторов и его ОЭП

В этом и следующем параграфах рассматривается подход к визуализации многомерной ($m > 2$) границы Парето на основе предварительной аппроксимации множества достижимых критериальных векторов или его ОЭП. Этот подход описан подробно в книге [11]. Подход состоит в аппроксимации множеств в достаточно простом виде. Далее двумерные сечения аппроксимации используются для быстрой диалоговой визуализации границы Парето.

В данном параграфе рассматриваются проблемы аппроксимации. Ясно, аппроксимация многомерных множеств требует различных подходов в выпуклом и невыпуклом случаях. Поэтому рассмотрим эти случаи отдельно — сначала выпуклый случай, а затем уже невыпуклый.

В том случае, когда аппроксимируемые множества выпуклы, обычно удается решить задачу максимизации линейной свертки критериев:

$$\langle c, y \rangle \rightarrow \max, \quad y = \varphi(x), \quad x \in X.$$

Для линейных многокритериальных задач, например, такая задача скалярной оптимизации может быть решена при практически любой размерности пространства решений. При ее решении осуществляется расчет значения опорной функции множества $Y = \varphi(X) \subset \mathbb{R}^m$ для направления $c \in \mathbb{R}^m$, т.е. вычисление функции

$$g_Y(c) = \max_{y \in Y} \langle c, y \rangle.$$

При расчете опорной функции будем считать, что $\|c\| = 1$.

14.1. О полиэдральной аппроксимации выпуклых компактных тел

Напомним, что под *полиэдральной аппроксимацией* выпуклого множества понимают поиск таких матрицы G и вектора g , что множество решений системы $Gy \leq g$ достаточно хорошо описывает аппроксимируемое множество. Интенсивная разработка методов полиэдральной аппроксимации выпуклых множеств связана с тем, что выпуклые множества часто встречаются в практических задачах. Например, множество достижимых критериальных векторов выпукло в линейных многокритериальных задачах, имеющих весьма широкое применение в экономических и экологических исследованиях. Важно, что вопрос об аппроксимации выпуклых тел достаточно хорошо изучен теоретически. В частности, разработаны методы полиэдральной аппроксимации выпуклых компактных тел, асимптотически оптимальные по простоте построенной аппроксимации и числу измерений характеристик аппроксимируемого множества.

Кроме возможности подсчитать значение опорной функции, при полиэдральной аппроксимации удобно предположить, что аппроксимируемое множество компактно и телесно. Таким образом, можно рассмотреть вопрос о полиэдральной аппроксимации множеств достижимых критериальных векторов, являющихся многомерными выпуклыми компактными телами (ВКТ), для которых можно рассчитать значение опорной функции для любого нормированного вектора $c \in \mathbb{R}^m$. В дальнейшем мы будем рассматривать задачу аппроксимации ВКТ, абстрагируясь от того факта, что нас интересует задача многокритериальной оптимизации. Благодаря этому описываемая методика может быть применена для решения и других прикладных задач.

Аппроксимация ВКТ осуществляется на основе построения последовательностей аппроксимирующих многогранников с растущим числом вершин. Точнее говоря, далее рассматриваются итерационные методы аппроксимации многомерного ВКТ, под которыми понимаются такие методы построения последовательности телесных многогранников $P^0, P^1, \dots, P^k, \dots$ с растущим числом вершин, в которых последующий многогранник строится на основе предыдущего с использованием расчетов опорной функции аппроксимируемого множества Y . При этом будем рассматривать такие методы, что

$$h(P^N, Y) \rightarrow 0,$$

где $h(A_1, A_2)$ — метрика Хаусдорфа, определенная в §6. Важным отличием рассматриваемых методов аппроксимации, основанных на построении последовательности многогранников, от методов приближенного описания с помощью отдельно взятых выпуклых тел, является возможность аппроксимации ВКТ с любой степенью точности. За это преимущество, однако, приходится платить дорогой ценой: как показывают практика и теоретические оценки, сложность описания аппроксимирующего многогранника быстро растет с увеличением точности аппроксимации и ростом размерности аппроксимируемого тела. Тем не менее, приходится идти на построение аппроксимирующих последовательностей в связи с тем, что в многокритериальных задачах представляет интерес форма паретовой границы аппроксимируемого множества, а не только область, где это множество находится. Поэтому применение итерационных методов связано с необходимостью разработки методов, оптимальных с точки зрения сложности описания аппроксимирующих многогранников. Кроме того, поскольку каждое вычисление опорной функции аппроксимируемого тела может быть достаточно сложным и требовать большого времени, представляют интерес методы, оптимальные с точки зрения числа вычислений опорной функции.

Такой подход разумен даже в том случае, когда аппроксимируемое множество само является многогранником, поскольку за небольшое число шагов может быть построен аппроксимирующий многогранник со значительно меньшим числом вершин.

Итерационные методы различаются способом построения последующего многогранника по предыдущему. Рассмотрим методы, использующие так называемую *схему восполнения*. В этой схеме используется построение внут-

ренных аппроксимирующих многогранников, вершины которых должны принадлежать границе $\text{Fr } Y$ аппроксимируемого тела Y .

Очередная $(k+1)$ -я итерация метода, основанного на схеме восполнения, состоит из двух шагов (предполагается, что уже построен многогранник P^k).

(k+1)-я итерация

шаг 1. Некоторым образом выбирается точка $\hat{y} \in \text{Fr } Y$.

шаг 2. Строится $P^{k+1} = \text{conv}\{\hat{y}, P^k\}$ в требуемом виде.

Конкретные методы, основанные на схеме восполнения, различаются способами решения на каждом шаге следующих задач:

- 1) способом выбора присоединяемой точки $\hat{y} \in \text{Fr } Y$,
- 2) видом задания $P^{k+1} = \text{conv}\{\hat{y}, P^k\}$ и способом его построения в требуемом виде.

Среди методов, основанных на схеме восполнения, можно выделить адаптивные (активные) методы, в которых при выборе точки $\hat{y} \in \text{Fr } Y$ используется информация о многограннике P^k . В этом случае выбор $\hat{y}$ адаптирован к форме Y в той мере, в какой P^k аппроксимирует Y . Адаптивные методы являются более эффективными, нежели неадаптивные (пассивные), основанные, как правило, на расчете опорной функции для некоторой априорной сетки на сфере направлений и не учитывающие конкретную форму аппроксимируемого тела. В качестве характерного примера адаптивных итерационных методов рассмотрим метод уточнения оценок, получивший наибольшее распространение к настоящему времени.

14.2. Метод уточнения оценок

Метод уточнения оценок (УО) основан на схеме восполнения и использует тот факт, что аппроксимирующий многогранник может быть задан двумя способами: как выпуклая оболочка вершин и как множество решений системы линейных неравенств. Для выбора присоединяемой точки $\hat{y} \in \text{Fr } Y$ с учетом формы P^k в методе УО используется представление P^k в виде множества решений системы линейных неравенств. Для этого на шаге 2 метода восполнения решается достаточно сложная задача построения $P^{k+1} = \text{conv}\{\hat{y}, P^k\}$ в виде системы линейных неравенств. Поскольку такая система линейных неравенств находится на основе вершин многогранников, в методе УО аппроксимирующий многогранник строится сразу в обеих возможных формах — в виде выпуклой оболочки и в виде пересечения полупространств, задаваемых неравенствами. Дадим более точное описание метода УО.

Перед началом $(k+1)$ -й итерации должен быть построен выпуклый телесный многогранник P^k с вершинами, принадлежащими границе аппроксимируемого тела Y . Многогранник P^k должен быть также задан виде множества решений некоторой системы линейных неравенств. Обозначим через $U(P^k)$ конечное множество единичных внешних нормалей к гипергра-

ням многогранника P^k . Очевидно, что множество $U(P^k)$ задано, поскольку многогранник P^k задан в виде множества решений системы линейных неравенств.

(k+1)-я итерация

шаг 1. На основе расчета $(g_Y(u) - g_{P^k}(u))$ для всех $u \in U(P^k)$ находится та нормаль $\hat{u} \in U(P^k)$, на которой достигается

$$\max_{u \in U(P^k)} (g_Y(u) - g_{P^k}(u)).$$

В качестве $\hat{y}$ выбирается такая точка $\text{Fr } Y$, что $\langle \hat{u}, \hat{y} \rangle = g_Y(\hat{u})$.

шаг 2. Определяется $U(P^{k+1})$ на основе построения $\text{conv}\{\hat{y}, P^k\}$ в виде множества решений системы линейных неравенств.

Метод УО проиллюстрирован на рис. 14.1, на котором изображены тело Y , многогранники P^k и P^{k+1} , а также внешний аппроксимирующий многогранник $\hat{P}^{k+1}$, получаемый попутно. Предполагается, что перед началом работы метода задан исходный выпуклый телесный многогранник P^0 с вершинами, принадлежащими границе аппроксимируемого тела. Методы построения многогранника P^0 здесь рассматриваться не будут.

Теоретическое исследование метода УО показало, что для тел, имеющих хотя бы одну точку гладкости, этот метод строит многогранники, последовательность которых оптимальна по порядку числа вершин. Были получены и более общие результаты о свойствах схем восполнения вообще и метода УО (см. [11]).

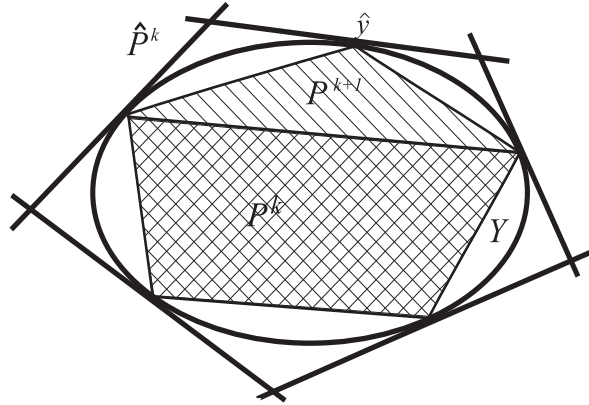


Рис. 14.1.

Алгоритмическая реализация метода УО основана на решении конечного числа задач выпуклой оптимизации (линейного программирования в случае линейных моделей) на шаге 1 и нахождении выпуклой оболочки многогранника P^k и точки $\hat{y}$ на шаге 2. Если первая задача является хорошо изученной и для ее решения имеется стандартное программное обеспечение, то задача, решаемая на втором шаге, требует особого внимания.

14.3. Построение выпуклой оболочки точки и многогранника

Рассмотрим второй шаг схемы восполнения, т.е. построение $\text{conv}\{\hat{y}, P^k\}$. Опишем на трехмерном примере общую схему решения этой задачи.

Пусть имеется многогранник $ABCDEF$, и требуется построить грани выпуклой оболочки этого многогранника и точки G (см. рис. 14.2 а). Нетрудно понять, что в эту новую выпуклую оболочку войдут все грани исходного многогранника, не видимые из точки G (т.е. $ABCD$, BCE , CDE , DEF и ADF), и не войдут грани, видимые из точки G (ABF и BEF , заштрихованные на рис. 14.2 а). Кроме того, выпуклая оболочка будет иметь новые грани. Этими новыми гранями являются грани минимального конуса с вершиной G , содержащего многогранник $ABCDEF$ (см. рис. 14.2 б).

Каждая такая новая грань лежит в гиперплоскости, проходящей через ребро многогранника, лежащее на границе между видимой и невидимой (из точки G) частями поверхности многогранника. Эта граница состоит из ребер AB , BE , EF и FA . Таким образом, каждая новая грань выпуклой оболочки проходит через пересечение пары соседних граней многогранника, одна из которых видима, а другая — невидима из точки G . Например, новая грань BEG проходит через ребро BE , которое является пересечением видимой грани BEF и невидимой грани BCE .

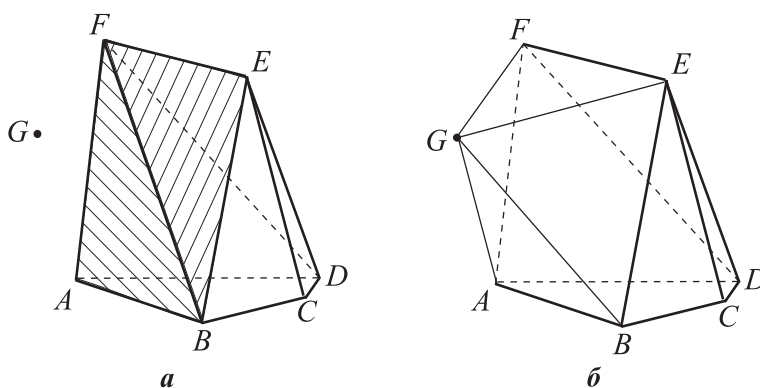


Рис. 14.2.

Описанная схема легко обобщается на случай пространства произвольной размерности m , если учесть, что граница между видимой и невидимой из точки G частями многогранника в $\mathbb{R}^m$ состоит из граней размерности $m - 2$. Эта схема построения выпуклой оболочки многогранника и точки часто называется методом “под-над” (“beneath-beyond”).

Для реализации схемы “под-над” требуется решить три проблемы: 1) определить, является ли данная гипергрань видимой из присоединяемой точки; 2) определить, являются ли две гиперграни соседними; 3) преобразовать представление многогранника к новому виду. Способы решения

перечисленных проблем зависят от формы представления многогранника и определяют конкретный вид метода.

Специфика метода уточнения оценок, требующая получить описание выпуклой оболочки в виде множества решений системы линейных неравенств, приводит к тому, что реализация схемы “под-над” в методе УО связана с использованием методов свертывания систем линейных неравенств. Мы не рассматриваем этот сложный вопрос в данном курсе (см. [11]).

Разработанные методы аппроксимации выпуклых тел многогранника оказались надежным средством аппроксимации множества достижимых критериальных векторов в случае его выпуклости, ограниченности и разумной размерности пространства критериев (до семи-восьми).

Отметим, что метод УО может быть использован и для аппроксимации множества Y_P . Для этого достаточно взять в качестве исходного аппроксимирующего многогранного множества неположительный ортант с вершиной на границе Парето. Далее все операции производятся без изменения. Все теоретические и экспериментальные утверждения также верны в этом случае.

14.4. Гибридные методы аппроксимации ОЭП

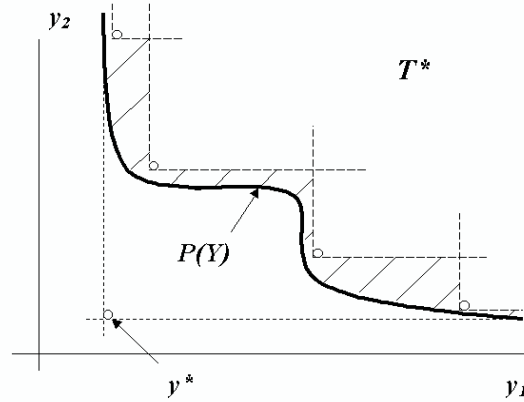
Обратимся теперь к невыпуклым нелинейным задачам. Ограничимся задачами аппроксимации ОЭП, наиболее востребованной в МКО. Для аппроксимации ОЭП используется гибридный метод, обладающий следующими свойствами:

- математическая модель может быть задана в виде черного ящика (скажем, вычислительного модуля, свойства которого не могут быть изучены аналитически);
- дается стохастическая оценка качества аппроксимации.

В данном методе в качестве аппроксимирующих множеств предлагается использовать совокупность конусов с вершинами, близкими к границе Парето. Точнее говоря, аппроксимация ОЭП задается конечной совокупностью T точек множества Y , в какой-то степени близких к границе Парето. Рассмотрим задачу многокритериальной минимизации. Тогда множество

$$T^* = \cup \{y + \mathbb{R}_+^m : y \in T\}$$

задает внутреннюю аппроксимацию множества Y^* . На рис. 14.3 изображены идеальная точка y^* , граница Парето $P(Y)$, четыре точки ее аппроксимации, а также множество T^* , заданное объединением конусов с вершинами в этих точках. Заметим, что T^* является ОЭП совокупности точек из T (здесь рассматривается задача многокритериальной минимизации). Использование в качестве аппроксимации множества T^* позволяет осуществить аппроксимацию и визуализацию невыпуклой ОЭП.



Об. 14.3.

Важнейшая особенность описываемого метода — использование вероятностных оценок качества текущей аппроксимации. Именно это позволяет отказаться от использования информации для оценки качества аппроксимации. Поэтому начнем изложение с методов статистического оценивания качества текущей аппроксимации ОЭП, а только затем перейдем к методам аппроксимации и визуализации. В заключение кратко обсудим перспективы использования описанных методов в многопроцессорных системах.

Статистическое оценивание качества аппроксимации ОЭП

Итак, начнем с вопроса об оценивании качества аппроксимации множества Y^* множеством T^* . Прежде всего, сформулируем требования к оценке качества аппроксимации. Во-первых, эта оценка должна быть *монотонной*, т.е. если внутренняя аппроксимация T_2^* более точна, чем T_1^* для $T_1^* \subset T_2^*$, то оценка качества T_2^* должна быть выше, чем для T_1^* . Во-вторых, она должна быть *действительной*, т.е. различать хорошие и плохие аппроксимации. В-третьих, оценка должна быть *реализуема* в реальных задачах.

Для стохастической оценки качества аппроксимации используется величина полноты аппроксимации. Под *полнотой* η_T некоторой аппроксимации T^* множества Y^* , заданной базой T , будем понимать вероятность того, что из $x \in X$ следует $\varphi(x) \in T^*$. Условие

$$\eta_T \geq \eta \quad (14.1)$$

означает, что в аппроксимации T^* представлена не менее чем η -я доля множества X (т.е. при достаточно больших η , $0 < \eta < 1$, множество T^* не очень бедно по сравнению с Y^*). Задача оценки качества T^* может быть сформулирована как проверка выполнения условия (14.1) при достаточно большом числе $\eta < 1$.

Оценка полноты аппроксимации в методе проводится эмпирически, на основе генерирования случайных точек из X , поэтому такая проверка может гарантировать выполнение утверждения (14.1) только с определенной

вероятностью $\chi < 1$. Таким образом, вместо (14.1) рассчитывается число η , для которого выполняется вероятностная оценка

$$P\{\eta_T > \eta\} \geq \chi, \quad (14.2)$$

где $0 < \chi < 1$. Вопрос состоит в том, как выбрать число N случайных точек $x \in X$, чтобы получить оценку типа (14.2) для заданного χ . Рассмотрим этот вопрос подробнее.

Пусть используется выборка $H_N = \{x_1, \dots, x_N\}$, состоящая из N случайных равномерно распределенных точек множества X . Пусть $\eta_T^N = n/N$, где n — число таких элементов выборки, для которых выполняется включение $\varphi(x) \in T^*$. Величина выборочной полноты аппроксимации η_T^N является несмещенной оценкой полноты аппроксимации η_T . Определение величины η и числа N случайных точек $x \in X$, достаточных для получения оценки (14.2), может быть основано на следующей теореме.

Теорема 14.1. *Имеет место следующая оценка:*

$$P\{\eta_T > \eta_T^N - \Delta(\chi, N)\} \geq \chi,$$

где

$$\Delta(\chi, N) = [\ln(1 - \chi)^{-1}/(2N)]^{1/2}. \quad (14.3)$$

Смысл этой теоремы состоит в том, что величина η отличается от η_T^N на величину $\Delta(\chi, N)$, связанную с величиной N согласно (14.3). Если задать положительную величину Δ , то по заданному значению χ можно найти число генерируемых точек, достаточное для получения требуемой точности — это минимальное целое число, удовлетворяющее неравенству

$$N(\Delta, \chi) \geq \ln(1 - \chi)^{-1}/(2\Delta^2). \quad (14.4)$$

Так, скажем, для $\Delta = 0.1$ и $\chi = 0.9$ формула (14.4) дает $N = 126$, а для $\Delta = 0.01$ и $\chi = 0.99$ получаем $N = 23026$. Таким образом, пользователь имеет широкие возможности варьирования этих величин в зависимости от продолжительности расчета величины критериального вектора. При этом для оценки качества аппроксимации наряду с интервальной оценкой полноты (14.2) можно просто использовать несмещенную оценку величины η_T , которой является выборочная полнота аппроксимации $\eta_T^{(N)}$. Другая полезная оценка — расстояние от совокупности $\varphi(x)$, $x \in H_N$, до множества T^* .

Пусть теперь пространство $\mathbb{R}^m$ нормированное, $Q \subset \mathbb{R}^m$. Сразу уточним, что расстояние между точками y^1 и y^2 в пространстве $\mathbb{R}^m$ в данном методе характеризуется метрикой

$$\rho(y^1, y^2) = \max_i \{\lambda_i |y_i^1 - y_i^2|\},$$

где λ_i — некоторые положительные веса. В этом случае шары пространства $\mathbb{R}^m$ являются параллелепипедами со сторонами, параллельными осям координат. Для $\varepsilon > 0$ через $(Q)_\varepsilon$ обозначим открытую ε -окрестность этого

множества. Тогда для каждого $\varepsilon > 0$ можно определить величину $\eta_T^N(\varepsilon)$, равную доле точек $x \in H_N$, для которых $\varphi(x) \in (T^*)_\varepsilon$. Функцию $\eta_T^N(\varepsilon)$ назовем функцией выборочной полноты аппроксимации. Она характеризует качество аппроксимации более полно, чем просто величина $\eta_T^N = \eta_T^N(0)$. Отметим, что функция $\eta_T^N(\varepsilon)$ является монотонно неубывающей функцией ε и может быть легко рассчитана. Существует конечная величина $\varepsilon_{max} = \min\{\varepsilon | \eta_T^N(\varepsilon) = 1\}$, называемая радиусом полного покрытия. В качестве оценки качества аппроксимации T^* можно взять значения функции при различных ε , например $\eta_T^N(0)$, а также радиус полного покрытия.

При неформальном анализе качества аппроксимации пользователь может рассмотреть график функции $\eta_T^N(\varepsilon)$ целиком. Величина $\eta_T^N(\varepsilon)$ является несмещенной оценкой величины $\eta_T(\varepsilon)$, смысл которой состоит в том, что $(1 - \eta_T(\varepsilon))$ — вероятность получения точки вне множества $(T^*)_\varepsilon$ при генерировании случайных точек в множестве X . Таким образом, функция $\eta_T^N(\varepsilon)$ показывает, каких усилий стоит улучшение аппроксимации ОЭП с помощью генерирования случайных точек.

Экспериментальное исследование описанной статистической оценки показало ее пригодность в случае относительно малой размерности пространства решений $\mathbb{R}^n$ и достаточно медленно меняющихся функций φ . В случае большой размерности пространства решений и (или) быстро меняющихся функций, когда множество эффективных решений имеет исключительно малую меру и не может быть найдено с помощью генерирования случайных точек, величина выборочной полноты аппроксимации оказывается неспособной различать точные и неточные аппроксимации (скажем, оказывается, что $\eta_T^N(0) = 0$ как для точных, так и для неточных аппроксимаций), т.е. оценка не является действенной. В связи с этим в методе в случае достаточно больших размерностей пространства решений используется альтернативная оценка качества аппроксимации, развивающая описанные идеи, но в то же время действенная как в случае большой размерности вектора решений, так и в случае быстро меняющихся функций.

Как и прежде, сгенерируем выборку H_N из N случайных равномерно распределенных точек множества X , а затем решим такую задачу скалярной оптимизации, что ее решение приближает вектор $\varphi(x)$ к границе Парето. Заменяем рассмотрение конкретных задач оптимизации изучением отображения $\Phi : X \rightarrow X$, ставящего в соответствие случайной точке $x \in X$ некоторую такую “улучшенную” точку $x' \in X$, что вектор $\varphi(x')$ “более близок” к $P(Y)$, чем $\varphi(x)$ (например¹⁾, $\varphi(x') \leq \varphi(x)$). Определим величину $\eta_\Phi^N(\varepsilon)$, равную доле точек $x \in H_N$, для которых $\varphi(\Phi(x)) \in (T^*)_\varepsilon$. Функция $\eta_\Phi^N(\varepsilon)$, называемая обобщенной выборочной полнотой, также характеризует качество аппроксимации, обладает свойством монотонности и может быть легко рассчитана. В том случае, когда корректно предположение о существовании вероятности

$$\eta_\Phi(\varepsilon) = P\{x \in X \Rightarrow \varphi(\Phi(x)) \in (T^*)_\varepsilon\}, \quad (14.5)$$

¹⁾Напомним, что здесь рассматривается задача многокритериальной **минимизации**.

на основе $\eta_{\Phi}^N(\varepsilon)$ можно построить доверительные интервалы для $\eta_{\Phi}(\varepsilon)$ при всех $\varepsilon \geq 0$.

Функция $\eta_{\Phi}^N(\varepsilon)$, являющаяся выборочной полнотой аппроксимации, используется также и сама по себе — она либо представляется пользователю в виде графика, либо просто происходит автоматическая проверка величины ее некоторых значений, скажем радиуса полного покрытия. Отметим, что величина $1 - \eta_{\Phi}(\varepsilon)$, где $\eta_{\Phi}(\varepsilon)$ определяется по формуле (14.5), по-прежнему характеризует вероятность получения точек вне ε -окрестности множества T^* в описанном процессе, т.е. трудоемкость улучшения текущей аппроксимации. Примеры функций $\eta_{\Phi}^N(\varepsilon)$ приведены на рис. 14.4.

Реализация отображения Φ , как уже говорилось, может быть основана на решении задачи локальной однокритериальной оптимизации, где в качестве целевой функции берется некоторая свертка $\psi(\varphi(x))$ векторного критерия $\varphi(x)$. Отметим, что свертка критериев оптимизации $\psi(y)$ может быть выбрана в соответствии с возможностями и целями пользователя. В результате применения градиентной процедуры решения этой задачи находится локальный экстремум $x' \in X$, зависящий от начальной точки $x^0 \in H_N$. Полагаем $x' = \Phi(x^0)$.

Логически наиболее подходящей сверткой является расстояние от точки x' до множества T^* , величину которой следует максимизировать. Реализация такой идеи, однако, затруднительна из-за того, что множество T^* задается как объединение большого числа (нескольких сотен или даже тысяч) конусов. Поэтому на практике используются два альтернативных подхода. В рамках первого из них точка $x' = \Phi(x^0)$ строится в результате решения задачи локальной минимизации свертки

$$\psi(y) = \max_j \{\lambda_j(\varphi_j(x) - \tilde{y}_j^*)\} + \delta \sum_{j=1}^m \varphi_j(x), \quad (14.6)$$

где $\tilde{y}_j^*$ — текущая аппроксимация идеальной точки, δ — малое число, $\delta > 0$. Величины λ_j определяются исходной точкой $x^k \in H_N$:

$$\lambda_j = [\varphi_j(x^k) - \tilde{y}_j^*]^{-1}, j = 1, 2, \dots, m,$$

при $\varphi_j(x^k) > \tilde{y}_j^*$. Благодаря индивидуальному выбору величин λ_j происходит движение текущей точки из $\varphi(x^k)$ в направлении точки $\tilde{y}_j^*$, что способствует ее смещению в сторону границы Парето текущей аппроксимации.

В рамках второго метода точка $x' = \Phi(x^0)$ строится в результате решения задачи локальной минимизации свертки

$$\psi(y) = -\min_j \{\lambda_j(y_j^0 - \varphi_j(x))\} + \delta \sum_{j=1}^m \varphi_j(x), \quad (14.7)$$

где y^0 — некоторая такая точка, что $y \leq y^0$ для всех $y \in P(Y)$, δ — малое число, $\delta > 0$. Величины λ_j также определяются (при $\varphi_j(x^k) < y_j^0, j = 1, 2, \dots, m$) исходной точкой $x^k \in H_N$:

$$\lambda_j = [y_j^0 - \varphi_j(x^k)]^{-1}, j = 1, 2, \dots, m.$$

Благодаря индивидуальному выбору величин λ_i происходит движение текущей точки из $\varphi(x^k)$ в направлении, противоположном y^0 .

Поскольку в методах локальной оптимизации решение зависит от исходной точки $x^0 \in X$, при решении такой задачи находится локальный минимум $x' = \Phi(x^0)$. Заметим, что поверхности уровня сверток (14.6) и (14.7) совпадают с границей конусов доминирования в пространстве критериев, что обеспечивает улучшение (по Парето) качества решения в процессе минимизации свертки.

Надо, однако, отдавать себе отчет в том, что расчет решения $x' = \Phi(x^0)$ обычно представляет собой весьма трудоемкую задачу, значительно более сложную, чем расчет $\varphi(x)$. Так, в рассмотренной далее задаче выбора оборудования для охлаждения стали ($n = 325$) для решения одной задачи оптимизации в среднем требовалось около 2000 расчетов значения функции $\varphi(x)$ (не считая расчета ее градиента). Поэтому при оценке обобщенной полноты аппроксимации приходится ограничиваться малым числом точек в выборке N и грубой оценкой величины $\eta_{\Phi}(\varepsilon)$. Эта грубая оценка, однако, крайне полезна, поскольку именно обобщенная полнота является наиболее важной характеристикой качества аппроксимации.

Приведем пример графиков обобщенной выборочной полноты для двух аппроксимаций ОЭП в пятикритериальной задаче оптимизации, критерии которой были основаны на использовании функций, имеющих существенные нелинейности, в частности несколько крутых пиков.

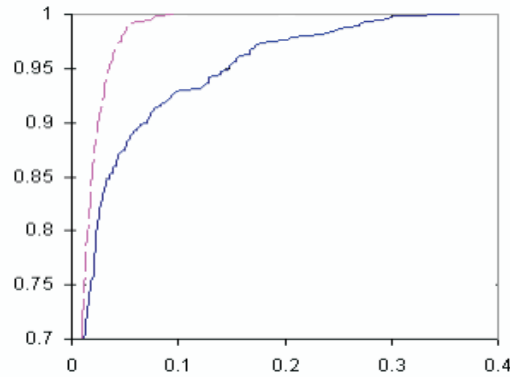


Рис. 14.4.

На рис. 14.4 представлены графики выборочной обобщенной полноты $\eta_{\Phi}^N(\varepsilon)$ для грубой аппроксимации (итерация 1, сплошная линия) и точной (итерация 7, штриховая линия) аппроксимации. Величина ε измеряется в относительных единицах — по отношению к диаметру границы Парето. Видно, что в случае итерации 1 приемлемая величина обобщенной полноты (скажем, 0.95) достигается только при $\varepsilon = 0.15$, т.е. 5% новых точек отстают

от аппроксимации дальше, чем 15% диаметра паретовой границы. В случае итерации 7 обобщенная полнота в 0.95 достигается уже при $\varepsilon = 0.03$. При этом радиус полного покрытия составляет около 0.35 для итерации 1 и меньше 0.1 для итерации 7. Таким образом, для итерации 7 все новые точки оказались на расстоянии от аппроксимации, составляющем не более 10% диаметра паретовой границы, причем 95% точек оказались на расстоянии не более чем в 3%. В этом примере применение обобщенной полноты позволяет распознать различие в качестве двух аппроксимаций в достаточно сложной задаче многокритериальной оптимизации.

Гибридный метод аппроксимации ОЭП

Начнем с описания общей концепции гибридных методов аппроксимации ОЭП для нелинейных задач. Гибридные методы²⁾, основываются на комбинировании методов различных типов, что позволяет решать задачи с разными свойствами критериальных функций. Методы, составляющие гибридные методы, являются итеративным, т.е. в них осуществляется пошаговая процедура улучшения качества аппроксимации.

Однофазный метод аппроксимации

В однофазном методе на каждой итерации осуществляется оценка качества полученной ранее аппроксимации T^* на основе построения выборочной функции полноты $\eta^N(\varepsilon)$, получаемой с помощью генерирования выборки H_N из N случайных равномерно распределенных точек множества допустимых решений X , и на основе расчета критериальных векторов для точек из H_N . При автоматической остановке метода проверяется выполнение требований по радиусу полного покрытия $\varepsilon_{max} < \varepsilon_0$, где ε_0 — заданное положительное число. В диалоговом режиме пользователю предоставляется график $\eta^N(\varepsilon)$, после чего он сам решает, останавливать ли алгоритм. Если условие остановки не выполняется, то из совокупности точек исходной базы аппроксимации и критериальных точек, полученных при оценке полноты, формируется промежуточная совокупность точек, из которой путем выделения недоминируемых точек создается новая база аппроксимации ОЭП. Достоинством метода является его простота и наличие способа оценки качества аппроксимации, а также высокая эффективность в случае малой размерности пространства решений и критериальных функций с относительно медленным ростом.

Двухфазные методы аппроксимации

В двухфазных методах на каждой итерации улучшение качества полученной ранее аппроксимации T^* осуществляется также на основе генерирования выборки H_N из N случайных равномерно распределенных точек множества допустимых решений X . Отличием является то, что эти точки

²⁾Березкин В. Е., Каменев Г. К., Лотов А. В. (2006) Гибридные адаптивные методы аппроксимации невыпуклой многомерной паретовой границы // ЖВМиМФ. 2006. Т. 46(11). С. 2009-2023

улучшаются с помощью отображения Φ , которое в двухфазных методах задается с помощью решения задачи локальной оптимизации одной из сверток критериев, приведенных выше, или их комбинации. Рассчитывается обобщенная выборочная полнота $\eta_{\Phi}^N(\varepsilon)$, по которой проверяется то же правило остановки, что и в однофазном методе. Если требуемая точность не достигнута, новая аппроксимация строится точно так же, как и в однофазном методе.

Достоинством двухфазного метода является наличие действенного способа оценки качества аппроксимации в случае большой размерности пространства решений (до нескольких сотен переменных). Отметим, что, вообще говоря, можно сформулировать большое число различных двухфазных оптимизационных методов, отличающихся методами свертки и их параметрами. Недостатком двухфазных методов является необходимость расчета большого числа критериальных точек при вычислении значений градиентов критериальных функции и при поиске локальных оптимумов вдоль выбранных направлений спуска. В том случае, когда заданы формулы расчета градиента, описанная проблема частично теряет свою остроту.

Вид метода, на котором основана локальная оптимизация, не имеет принципиального значения. В том случае, когда критериальные функции рассчитываются вычислительными модулями, приходится осуществлять оптимизацию только с использованием значений этих функций (simulation-based optimization).

Трехфазные методы аппроксимации

Трехфазные методы отличаются от двухфазных тем, что при генерировании выборки H_N используется не равномерное распределение точек на множестве X , а распределение, зависящее от аппроксимации множества парето-эффективных решений $P(X)$. При этом выборка является объединением двух выборок, одна из которых генерируется равномерно на всем множестве X , а вторая генерируется равномерно на аппроксимации множества $P(X)$. При этом первая выборка содержит малое число точек, а вторая — большое число. Подобная идея давно используется в методах неравномерного случайного поиска оптимума единственного критерия (сжатия области поиска), в частности, в широко используемом в настоящее время методе “имитируемого остывания стали” (simulated annealing). В отличие от традиционного метода simulated annealing, в предлагаемом методе множество, на котором проводится поиск, определяется адаптивно, т.е. на основе предыдущих расчетов, а не задается как заранее определенная окрестность точки текущего оптимума. В остальном метод не отличается от двухфазного метода оценки качества аппроксимации границы Парето.

Важным аспектом трехфазного метода является способ аппроксимации множества $P(X)$: оно аппроксимируется объединением шаров с центрами в прообразах точек базы и радиусом $\tau > 0$. Для расчета величины радиуса τ применяется теория экстремальных статистик, описание и применение которой для расчета величины радиуса τ выходит за пределы нашего курса лекций. Благодаря тому, что генерация большинства случайных точек осу-

ществляется в пределах аппроксимации множества парето-эффективных решений, метод оказывается особенно эффективным в случае большой размерности пространства решений (до нескольких сотен переменных) и сложной структуры экстремумов критериальных функций. В частности, они позволяют уменьшить число находимых локальных экстремумов, не имеющих отношения к границе Парето. Это приводит к значительному ускорению процесса аппроксимации.

Генетический метод

Генетические методы представляют собой результат реализации эволюционных концепций, основанных на идеях смешивания разумных решений. Генетические алгоритмы могут потребовать меньшего числа расчетов критериальных точек по сравнению с двух- и трехфазными методами.

В гибридных методах хорошо показал себя генетический метод “оштукатуривания” (plastering), примененный к уже достаточно точной аппроксимации, построенной с помощью двух- и трехфазного методов. В методе “оштукатуривания” из прообразов имеющейся базы покрытия T случайно (но с учетом некоторых ограничений) выбираются пары точек (“родители”). Далее, между точками каждой такой пары проводится отрезок и на этом отрезке случайным образом выбирается заранее заданное или случайное число $q > 1$ новых точек множества допустимых решений (“популяция наследников”). Для всех точек-наследников вычисляются значения $\varphi(x)$. Полученные критериальные точки используются для оценки качества имеющейся аппроксимации (в том числе рассчитывается радиус полного покрытия ε_{max}). Далее метод не отличается от однофазного, за исключением того, что число точек в базе аппроксимации в методе оштукатуривания может стать слишком велико, поэтому из полученной базы по какому-либо правилу (например, близость в пространстве решений) должно быть исключено значительное количество точек.

Эксперименты показывают, что наиболее подходящее место этого метода — завершающие шаги процесса аппроксимации, когда, наряду с уточнением аппроксимации, требуется получить достаточно большое число точек базы аппроксимации для большей выразительности изображений границы Парето. Этим описываемый гибридный метод отличается от большинства гибридных методов, в которых эволюционные методы используются с самого начала.

Использование параллельных вычислений

В задачах с большой размерностью вектора решений и сложной структурой критериальных функций построение достаточно точной аппроксимации границы Парето требует вычисления большого числа критериальных точек. Если же время одного такого расчета составляет минуты, то неизбежно обращение к параллельным вычислениям, которые позволяют распределить большой объем вычислений на большое число процессоров. Обычно при этом возникают сложные проблемы организации параллельного счета. Большим достоинством описанного гибридного метода аппрок-

симации ОЭП является наличие естественного параллелизма. Как вычисление критериальной точки, так и решение задач локальной оптимизации могут осуществляться параллельно без какой-либо предварительной адаптации алгоритма. Более того, выделение недоминируемого подмножества для полученных критериальных точек иногда (например, в однофазном методе) также можно осуществлять параллельно. Хотя обработка результатов всех расчетов итерации (оценка полноты, максимального отклонения и т. д.) требует полного объема информации, это не сильно сказывается на эффективности расчетов, поскольку требует относительно малого объема вычислений.

Интересно, что предложенный метод аппроксимации ОЭП для нелинейных систем может быть реализован не только на компьютерных кластерах с известным числом процессоров, но и в компьютерных сетях с неопределенным числом процессоров (и, возможно, других компьютерных ресурсов, таких как средства хранения информации), географически разделенных, но связанных между собой сетью. Такие средства расчета весьма недороги, поскольку используют время простоя компьютеров, которое практически бесплатно, но потенциально весьма мощны. Отметим, что некоторые черты таких средств делают программирование на них значительно более сложным, чем на обычных компьютерных кластерах: процессоры могут сильно различаться между собой, число доступных процессоров может резко меняться во времени, связь между отдельными процессорами может быть плохой и, наконец, процессоры могут выключаться без предварительного предупреждения. В гибридном методе, описанном здесь, благодаря использованию случайных точек множества X эти проблемы исчезают: результат статистического анализа качества аппроксимации мало зависит от того, будет ли получен результат расчета от отдельного процессора. То же самое относится к результатам локальной оптимизации: отдельный расчет не имеет принципиального значения.

Глава 15. Визуализация границы Парето

15.1. Двумерные сечения многомерных множеств

В этом параграфе описываются методы, позволяющие осуществить визуализацию границы Парето для задач МКО после завершения аппроксимации множества достижимых критериальных векторов Y или его оболочки Эджворта-Парето Y_P простыми фигурами. В описываемых методах граница Парето визуализируется на основе изображения двумерных сечений многомерного множества или карт решений — специальным образом сформированных наборов его двумерных сечений.

Напомним определение двумерного сечения некоторого множества Z пространства $\mathbb{R}^m$. Пусть нас интересует сечение множества Z плоскостью, проходящей через некоторую точку $\hat{y} \in Z$ параллельно плоскости двух координат (u, v) . Каждую точку $y \in \mathbb{R}^m$ можно представить в виде $y = (u, v, w)$, где u и v — значения выбранных координат, а w — совокупность значений остальных $(m - 2)$ -х координат (мы не будем обращать внимание на порядок координат). Тогда двумерное сечение множества Z плоскостью, проходящей через точку $\hat{y} = (\hat{u}, \hat{v}, \hat{w}) \in Z$, есть множество

$$G_{u,v}(Z, \hat{w}) = \{(u, v) | (u, v, \hat{w}) \in Z\}.$$

Конечно, это двумерное сечение содержит точку $(\hat{u}, \hat{v})$. Итак, сечение заданного множества Z , параллельное плоскости некоторых двух координат (u, v) , определяется значениями остальных координат. Меняя значения w , получаем различные параллельные сечения множества Z . Недоминируемые границы этих сечений дают в совокупности представление о границе Парето множества Z .

Опыт визуализации различных наборов сечений множеств Y и Y_P позволил сделать вывод о том, что для человека наиболее информативным является рассмотрение таких наборов сечений, в которых от сечения к сечению меняется значение лишь одной из координат совокупности w . Набор параллельных двумерных сечений множества Y (множества Y_P), в котором используется именно такое правило изменения координат совокупности w , принято называть *картой решений*.

Для того чтобы можно было визуализировать карты решений в интерактивном режиме (или проводить анимацию карт решений, что оказывается весьма эффективным способом демонстрации паретовой границы), требуется рассчитывать и изображать сотни или даже тысячи сечений изучаемых множеств за время порядка десятка секунд. Именно для этого заранее аппроксимировалось множество Y или Y_P в каком-либо простом виде, удобном для быстрого расчета сечений.

Заметим, однако, что в этой процедуре имеется следующее неприятное свойство множества Y , которое может помешать изучению границы Парето.

Дело в том, что не при всех наборах $\hat{w}$ точки недоминируемой границы двумерного сечения

$$G_{u,v}(Y, \hat{w}) = \{(u, v) | (u, v, \hat{w}) \in Y\}$$

множества Y проходят через границу Парето множества Y . Отдельное сечение само по себе не дает возможности понять проходит ли оно через границу Парето. Для того, чтобы понять, соответствуют ли точки недоминируемой границы сечения точкам границы Парето, требуется изобразить и сравнить большое число параллельных сечений множества Y .

Сечения множества Y_P лишены этого недостатка, их связь с точками границы Парето более проста. Прежде всего, заметим, что ОЭП обладает важным свойством монотонности:

$$G_{u,v}(Y_P, \hat{w}) \subseteq G_{u,v}(Y_P, \bar{w}) \quad \text{при} \quad \hat{w} \geq \bar{w}.$$

Благодаря этому сечения не расширяются с ростом координат точки $\hat{w}$, в частности, границы сечений не пересекаются на картах решений.

Кроме того, имеет место одно важное свойство недоминируемых границ сечений множества Y_P .

Лемма 15.1. *Пусть бинарное отношение Парето обладает НМ-свойством на $P(Y) \subseteq Y$. Тогда, если точка $(\hat{u}, \hat{v})$ принадлежит недоминируемой границе сечения $G_{u,v}(Y, \hat{w})$, то найдется такой набор значений остальных критериев $w' \geq \hat{w}$, что $(\hat{u}, \hat{v}, w') \in P(Y)$.*

Доказательство. Согласно лемме 3.1, каждая точка Y_P либо принадлежит множеству Y , либо доминируется точкой множества Y , которую в силу НМ-свойства отношения Парето на $P(Y) \subseteq Y$ можно считать недоминируемой. Рассмотрим такую точку $(u', v', w') \in P(Y)$, что $u' \geq \hat{u}$, $v' \geq \hat{v}$, $w' \geq \hat{w}$. Поскольку $w' \geq \hat{w}$, то $G_{u,v}(Y_P, w') \subseteq G_{u,v}(Y_P, \hat{w})$. Отсюда $(u', v') \in G_{u,v}(Y_P, \hat{w})$. Так как $u' \geq \hat{u}$, $v' \geq \hat{v}$, а точка $(\hat{u}, \hat{v})$ — недоминируемая точка $G_{u,v}(Y_P, \hat{w})$, то $u' = \hat{u}$, $v' = \hat{v}$. Лемма доказана. $\square$

Таким образом, недоминируемые точки сечения множества Y_P при любом $w = \hat{w}$ порождаются точками $P(Y)$ с $w' \geq \hat{w}$, т.е. точками w , не менее предпочтительными, чем $\hat{w}$. Быть может, более удобно выглядит эквивалентное утверждение о том, что недоминируемые точки границы сечения множества Y_P при $w = \hat{w}$ являются паретовскими в двухкритериальной задаче

$$(u, v) \rightarrow \max, \quad (u, v, w) = \varphi(x), x \in X, w \geq \hat{w},$$

т.е. недоминируемые точки сечения ОЭП характеризуют эффективное замещение между двумя рассматриваемыми критериями, в то время как на остальные наложено ограничение $w \geq \hat{w}$.

Приведенные факты объясняют, почему удобно рассматривать визуализацию множества Y_P , а не множества Y . В то же время надо помнить, что есть много прикладных задач, в которых не очень ясно, действительно ли

нужна максимизация какого-то критерия при любых значениях остальных критериев. В этом случае, который формально уже не относится к задаче многокритериальной оптимизации, но может быть интересен на практике, может быть полезной визуализация множества Y .

Итак, для того, чтобы задать двумерное сечение многомерного множества (в нашем случае — множества Y_P), надо выбрать пару критериев (u, v) , достижимые значения которых будут визуализироваться при заданных значениях остальных критериев $w = \hat{w}$. Такие два критерия называют *координатными* критериями. Для остальных (некоординатных) критериев w надо задать их значения $\hat{w}$. При этом для визуализации карты решений требуется рассчитать и изобразить целую совокупность двумерных сечений ОЭП.

15.2. Построение сечений выпуклых множеств

В выпуклом случае, как уже было показано в предыдущем параграфе, осуществляется полиэдральная аппроксимация множеств Y и Y_P , т.е. аппроксимация в виде множества решений системы линейных неравенств в критериальном пространстве. Наличие уже построенной аппроксимации позволяет быстро находить двумерные сечения изучаемых множеств следующим образом. Пусть множество Y (или Y_P) уже аппроксимировано в виде системы линейных неравенств

$$\sum_{k=1}^m g_{lk} y_k \leq g_l, l = 1, \dots, L, \quad (15.1)$$

которая определяет в пространстве $\mathbb{R}^m$ некоторое непустое множество. Для расчета двумерного сечения этого множества нужно фиксировать значения всех координат, кроме двух, скажем, y_i и y_j . Тогда двумерное сечение множества (15.1) в координатах (y_i, y_j) можно записать в виде системы

$$g_{li} y_i + g_{lj} y_j \leq g_l - \sum_{k \neq i, j} g_{lk} \hat{y}_k, l = 1, \dots, L,$$

где $\hat{y}^k$ — заданные значения координат y_k , $k \neq i, j$, обозначенные через $\hat{w}$ в определении сечения, данном в начале параграфа. Таким образом, в том случае, когда множество задано системой неравенств, сечение описывается совокупностью ограничений, получаемой умножением подматрицы, составленной из соответствующих столбцов матрицы системы (15.1) на вектор $\hat{y}_k = (y_k, k \neq i, j)$. После этого, для изображения границ двумерного сечения многогранного множества, аппроксимирующего множество Y_P , требуется рассчитать вершины сечения, а затем соединить их ребрами. Были разработаны алгоритмы, позволяющие находить вершины серии двумерных сечений многогранных множеств на экране персонального компьютера достаточно быстро [11]. Благодаря этому многие сотни сечений могут быть рассчитаны и изображены за секунды, что позволяет не только быстро изображать карты решений по заказу пользователя, но и осуществлять их анимацию. Таким образом, вычислительные аспекты расчета серии сечений в

выпуклом случае были решены. Для решения этой задачи в невыпуклом случае были также разработаны эффективные методы, которые обсуждаются в конце параграфа.

После выбора координатных критериев остается решить два вопроса: как выбрать совокупность наборов значений некоординатных критериев и как расположить получаемые сечения на экране. Существует два подхода к решению этого вопроса: неструктуризованная визуализация и структуризованная визуализация. Рассмотрим эти подходы.

15.3. Неструктуризованная визуализация границы Парето

Начнем с самого простого случая трех критериев. Тогда имеется лишь один некоординатный критерий и требуется задать совокупность значений этого критерия. При первом ознакомлении с множеством значения некоординатного критерия удобно распределить равномерно, что может быть сделано автоматически (на последующих стадиях анализа пользователь может задать и другую совокупность интересных для него значений некоординатного критерия). Рассмотрим вопрос о расположении сечений трехмерного множества на экране. В зависимости от цели исследования, сечения могут быть расположены одно рядом с другим или наложены одно на другое. Как уже говорилось, в последнем случае изображения принято называть картами решений. Сечения могут быть изображены различными цветами либо различными штриховками. Связь между значениями третьего критерия, который в этом случае принято называть цветовым, и его цветом обычно указываются в палитре, расположенной около карты сечений. Наложение сечений в картах решений позволяет легко сравнивать сечения между собой — в силу монотонного расширения сечений ОЭП при ухудшении значения третьего критерия пользователь может легко понять, как изменение некоординатного критерия влияет на границу Парето. Примеры карт решений будут приведены далее в этом параграфе.

Рассмотрим теперь случай четырех критериев. К некоординатным критериям теперь относятся два из них, и для того чтобы изобразить серию сечений, требуется задать совокупности значений обоих некоординатных критериев. Эти значения могут быть заданы сеткой в пространстве критериев, каждому узлу которой соответствует одно двумерное сечение. Выбор значений, так же как и в случае трех критериев, может быть осуществлен как автоматически, так и пользователем. Как и в случае трех критериев, двумерные сечения можно попытаться наложить одно на другое.

Как показывает опыт, такая картина обычно является очень сложной и мало информативной. Можно, наоборот, расположить сечения в виде двумерной матрицы, каждый элемент которой — окно, изображающее единственное сечение, соответствующее узлу сетки значений некоординатных критериев. Это, однако, также оказывается неудобным, поскольку затрудняет сравнение сечений. Самым удобным расположением сечений является расположение в ряд нескольких карт решений, каждая из которых соответ-

ствуется некоторому фиксированному значению одного из некоординатных критериев (скажем, четвертого).

В случае пяти критериев каждая карта решений соответствует фиксированным значениям двух некоординатных критериев (скажем, четвертому и пятому), поэтому приходится рассматривать матрицу карт решений. Число карт решений в матрице может определяться пользователем, она зависит от его интересов (а также и от качества изображения).

Пользователь может легко изменить разбиение критериев на типы, т.е. выделить другие координатные и цветовой критерии, изменить значения некоординатных критериев, для которых строятся карты решений — компьютер быстро пересчитает сечения. Не представляет сложности и реализация сужения диапазона значений любого из критериев. Все эти возможности связаны с тем, что ОЭП аппроксимировано заранее в виде (15.1).

Заметим, что поскольку карту решений можно изобразить практически мгновенно, то возможно управление значениями некоординатных критериев (кроме цветового) с помощью такого широко распространенного элемента программного обеспечения как прокрутка. Пользователь может задать значение критерия, связанного с прокруткой, передвигая движок прокрутки. Если же совокупность значений такого критерия расположить равномерно, то процесс их перебора можно автоматизировать — значения могут последовательно перебираться компьютером, а соответствующие карты — визуализироваться. Если число заданных значений критерия прокрутки достаточно велико, то последовательная смена рисунков может создать иллюзию движения — анимационный эффект.

Рассмотрим возможности использования анимационного эффекта более подробно. Начнем с того, что его можно использовать уже при трех критериях: двух координатных критериях и одном некоординатном, связанном с прокруткой. Задав большое число равномерно расположенных значений некоординатного критерия, можно последовательно сменять двумерные сечения его ОЭП. Если совокупность значений некоординатного критерия и скорость смены сечений выбрана разумно, то возникнет анимационный эффект, который позволяет пользователю изучить воздействие некоординатного критерия на достижимые значения координатных критериев. Такой способ визуализации совокупности сечений принято называть динамической визуализацией. В процессе динамической визуализации пользователь имеет возможность зафиксировать тот “кадр” (т.е. значение автоматически перебираемого критерия прокрутки), в котором изображено некоторое интересное для него сечение. Выбрав не слишком большое число таких кадров, можно получить ряд двумерных сечений, рассмотренный ранее. При этом выбор кадров уже будет отражать интересы пользователя.

В случае четырех и пяти критериев, используя одну или две прокрутки, можно реализовать динамическую визуализацию карты решений, ряда или матрицы двумерных сечений и т.д. Вопрос состоит только в том, за сколькими сечениями и рисунками способен одновременно следить пользователь.

Направивается сравнение динамической визуализации с видеоманипулятором, с помощью которого исследователь просматривает заранее снятый

фильм об изучаемом множестве, выбирая попутно наиболее понравившиеся ему кадры. Хотя с точки зрения пользователя такая аналогия вполне уместна, рассматриваемая методика отличается тем, что кадры не отрисованы заранее, они быстро рассчитываются в процессе просмотра на основе заранее проведенной предобработки — аппроксимации множества в виде (15.1). Благодаря этому, пользователь имеет свободный доступ к информации: в любой момент он может переключиться с какого-либо определенного “фильма” (скажем, о динамике данного двумерного сечения при изменении некоторого третьего критерия) на изучение связей других переменных, т.е. на другой “фильм”. Более того, ему оказывается легко доступен любой потенциально возможный (виртуальный) “фильм”, который может быть реализован благодаря заранее построенной аппроксимации ОЭП.

Если число критериев больше пяти, то использование прокруток позволяет осуществить динамическую визуализацию карты решений, расположив на прокрутках все критерии, кроме координатных и цветового. Можно также осуществить динамическую визуализацию матрицы карт решений. Для этого выделяют пять критериев, задающих матрицу карт решений, а остальные располагают на прокрутках и используют возможности анимации с помощью прокруток.

15.4. Пример использования неструктуризованной процедуры визуализации границы Парето

Приведем пример практического использования рассматриваемых методов визуализации границы Парето в рамках системы поддержки принятия решений и переговоров, предназначенной для поиска эффективных стратегий улучшения качества воды в бассейне реки Ока¹⁾. Визуализация границы Парето позволяет специалистам осуществить целостный анализ проблемы качества воды в реке и указать на границе Парето цели (предпочтительные сочетания значений критериев), являющиеся достижимыми и недоминируемыми с точки зрения соотношений используемой математической модели. На основе выбранной цели автоматически формируются стратегии использования капиталовложений.

Главная проблема для большинства областей в бассейне Оки — это загрязнение нефтепродуктами. Предположим, что ведутся переговоры между представителями Министерства природных ресурсов, а также Московской и Нижегородской областей, которые являются в экономическом отношении наиболее развитыми областями бассейна. Пусть используются следующие критерии выбора решения:

- 1) максимальные концентрации нефтепродуктов в пределах Московской области z_{r45} (измеряется в ПДК — предельно допустимых концентрациях),
- 2) максимальные концентрации нефтепродуктов в пределах Нижегородской области z_{r75} (измеряется в ПДК),

¹⁾ А.В. Лотов, В.А. Бушенков, О.Л. Черных Структура и опыт использования компьютерной системы поддержки поиска водохозяйственных стратегий // Научно-техническая информация, сер. 2. Информационные процессы и системы, №3, 1998.

- 3) общая стоимость проекта F (измеряется в миллиардах рублей),
 - 4) затраты, осуществляемые на территории Московской области $F4$,
 - 5) затраты, осуществляемые на территории Нижегородской области $F7$.
- Обратим внимание на то, что в данном случае значения всех критериев надо минимизировать.

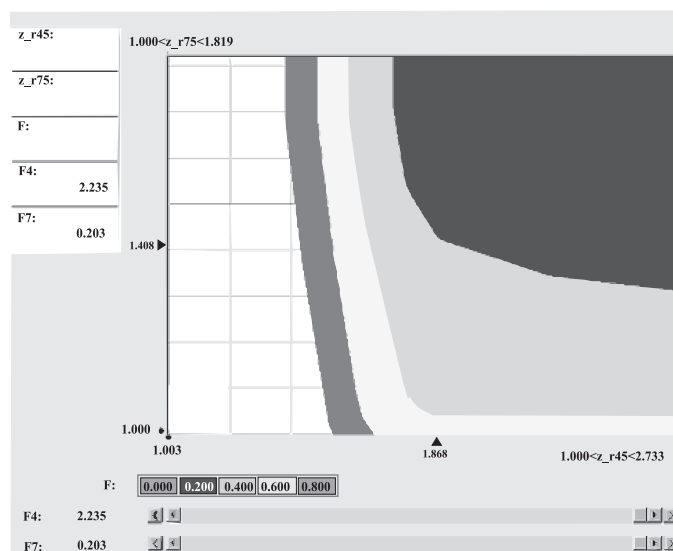
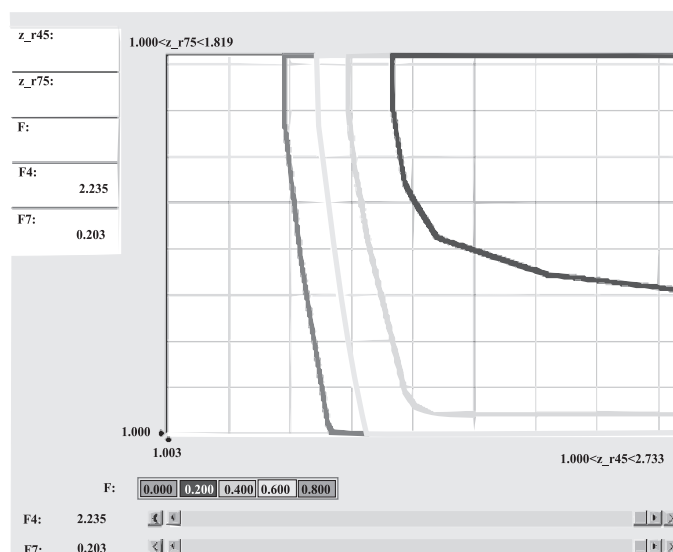


Рис. 15.1.

На рис. 15.1 приведенная черно-белая копия экрана для пяти критериев (поэтому на этой карте решений вместо различных цветов использованы оттенки серого). На осях отложены концентрации нефтепродуктов в Московской области (z_{r45} , по горизонтальной оси) и в Нижегородской области (z_{r75} , по вертикальной оси). Связь между оттенками серого и отображаемой ими общей стоимостью проекта дана в шкале, расположенной под картой решений. Как видно, стоимость проекта меняется от 0 до 0.800 миллиардов рублей с шагом в 200 миллионов рублей. Затраты, осуществляемые на территории Московской и Нижегородской областей, ограничены большими величинами (слайдеры расположены в крайней правой позиции), так что эти ограничения на рис. 15.1 не оказывают влияния на форму карты решений.

Для более точного анализа границы Парето удобно использовать карты решений, на которых изображены только границы сечений. Во всяком случае, инженеры-водохозяйственники предпочитают использовать именно такой вариант карты решений. Для рассматриваемой задачи вариант с изображением границ сечений приведен на рис. 15.2.



Об. 15.2.

Для удобства ориентирования на карте решений, обратим внимание на то, что значения в левом нижнем углу карты для обеих областей равны 1 ПДК, в то время как в правом верхнем эти значения равны текущему загрязнению, то есть 2.73 ПДК для Московской области (показатель z_{r45}) и 1.82 ПДК для Нижегородской области (показатель z_{r75}). Сетка соответствует значениям 1.0, 1.2, и т.д. для горизонтальной оси и 1.0, 1.1, и т.д. для вертикальной оси.

Как видно, форма эффективной границы сильно зависит от стоимости проекта. Рассмотрим границу сечения, связанную с ограничением на общую стоимость проекта в 200 миллионов рублей (самое темное сечение рис. 15.1). Для этой суммы недоминируемые сочетания загрязнений нефтепродуктами в Московской (показатель z_{r45}) и Нижегородской (показатель z_{r75}) областях образуют границу, при движении вдоль которой значения этих величин меняются. Загрязнение в Московской области может быть изменено от максимального (текущего) в 2.73 ПДК до минимально возможного при этих затратах, которое составляет примерно 1.75 ПДК. Если выбрать минимально возможное (при данных затратах) загрязнение в Московской области, равное 1.75 ПДК, то загрязнение в Нижегородской области не может быть менее чем 1.7 ПДК. Таким образом, использование 200 миллионов рублей в интересах Москвы может привести к тому, что загрязнение в Нижегородской области попутно упадет с максимального в примерно 1.8 ПДК до примерно 1.7 ПДК, но не до минимально возможного при таких затратах в 1.3 ПДК. На карте решений ясно видны не только минимальные значения критериев, но и вся недоминируемая граница для величин загрязнения нефтепродуктами в Московской и Нижегородской областях при рассматриваемых суммарных затратах не более 200 миллионов рублей. Часто та-

кую границу называют кривой эффективного замещения (efficient tradeoff curve), поскольку она показывает, как уменьшение значения одного из критериев приводит к увеличению значения другого для парето-эффективных решений.

Недоминируемая граница для сечения, соответствующего 400 миллионам рублей, изображена более светлой линией. Как видно, эта сумма позволяет практически решить проблему загрязнения нефтепродуктами для Нижегородской области (до 1.05 ПДК), если, конечно, капиталовложения будут делаться в ее интересах. В то же время, минимальное загрязнение для Московской области при этих затратах (1.6 ПДК) достигается при значительной величине загрязнения в Нижегородской области (1.7 ПДК). К счастью для нижегородцев, кривая замещения между критериями крута в области минимального загрязнения для Московской области (увеличение загрязнения в Московской области на 0.2 ПДК приводит к уменьшению в загрязнении в Нижегородской области на 0.65 ПДК) — сравните точку с координатами ($z_{r45} = 1.6$, $z_{r75} = 1.7$) с точкой излома ($z_{r45} = 1.8$, $z_{r75} = 1.05$). В связи с этим, нижегородцы могут надеяться настоять на выборе на этой кривой зоны излома, характеризуемой малым загрязнением в Нижегородской области, за счет уступок Москве по каким-либо другим вопросам.

Дальнейшее увеличение суммарных затрат на проект дает значительно меньший эффект — следующая кривая замещения (белый цвет на рис. 15.1 вместо желтого на цветной карте решений) отстоит от предыдущей всего примерно на 0.1-0.2 ПДК. Можно рассмотреть и более точные карты решений, уменьшив диапазон рассматриваемых величин. Можно оценить влияние ограничений, накладываемых на затраты, осуществляемые на территории Московской или Нижегородской областей, рассмотрев матрицу карт решений или применив компьютерную анимацию. Можно также выбрать другие координатные критерии и рассмотреть проблему с иной точки зрения.

Всесторонне ознакомившись с состоянием проблемы, пользователь (ЛПР) может указать предпочтительное сочетание значений двух критериев ($\hat{u}$, $\hat{v}$) прямо на предпочтительной недоминируемой границе одной из подходящих карт решений. Компьютер рассматривает вектор ($\hat{u}$, $\hat{v}$, $\hat{w}$) в качестве целевой точки, после чего находит по ней парето-эффективное решение проблемы, приводящее к значениям критериев, близких к ($\hat{u}$, $\hat{v}$, $\hat{w}$). Поскольку целевая точка ($\hat{u}$, $\hat{v}$, $\hat{w}$) является достижимой, такой метод выбор целевой точки получил название метода *достижимых целей*.

Итак, при визуализации паретовой границы пользователь имел полную свободу выбора двумерных сечений ОЭП, карт решений, их матриц и других средств визуализации. Далее описывается другой подход к визуализации паретовой границы, в котором пользователь получает только простые вопросы о сравнении альтернатив и направляется компьютером при поиске наиболее предпочтительной достижимой цели на паретовской границе.

15.5. Структуризованная процедура визуализации границы Парето

Структуризованная процедура визуализации границы Парето, получившая название метода Pareto Step (“Шаг по границе Парето”), предназначена для решения задачи многокритериальной оптимизации в случае выпуклой ОЭП множества достижимых критериальных векторов. Для таких задач МКО возможна предварительная аппроксимация ОЭП многогранниками, что и используется в данном методе — будем считать, что ОЭП аппроксимирована многогранным множеством, заданным системой линейных неравенств (15.1).

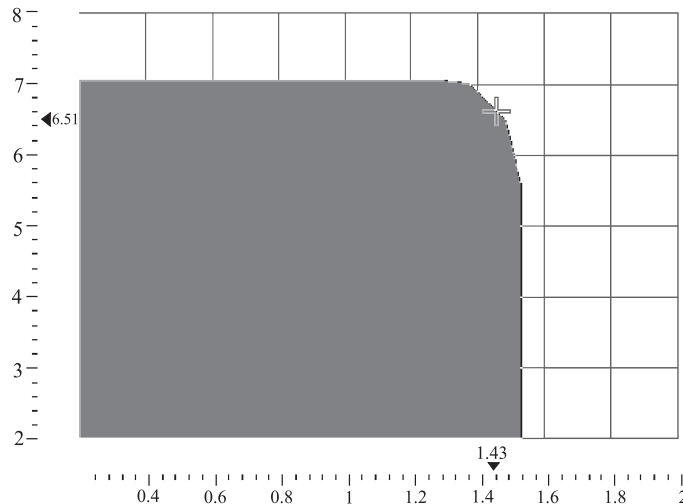
Рассмотрим сначала вариант метода, основанный на изображении единственного двухкритериального сечения. Метод Pareto Step, в отличие от неструктуризованной процедуры, описанной в ранее, состоит из заранее сформулированных итераций, в которых определены действия ЛПР и компьютера. Перед $(k + 1)$ -й итерацией требуется, чтобы был задан текущий вектор $y^{(k)} \in P(Y)$.

(k+1)-я итерация

шаг 1. ЛПР получает (в виде строки или на диаграмме) текущий недоминируемый вектор

$$y^{(k)} = (y_1^{(k)}, \dots, y_i^{(k)}, \dots, y_j^{(k)}, \dots, y_m^{(k)})$$

и указывает критерий, значение которого он хочет увеличить в первую очередь, а также критерий, которым он готов пожертвовать. Пусть это критерии y_i и y_j соответственно.



Об. 15.3.

шаг 2. ЛПР получает на экране компьютера сечение ОЭП в координатах y_i и y_j , проходящее через точку $y^{(k)}$. На паретовой границе Парето этого сечения высвечивается текущая точка $(y_i^{(k)}, y_j^{(k)})$ (см. 15.3). ЛПР предлагается указать на границе Парето более предпочтительные значения y_i и y_j , которые обозначим через $y_l^+, l = i, j$. В результате осуществления этой операции оказывается задана новая критериальная точка

$$\tilde{y}_l^{(k+1)} = \begin{cases} y_l^{(k)}, l \neq i, j \\ y_l^+, l = i, j \end{cases}.$$

шаг 3. По указанной точке $\tilde{y}^{(k+1)}$ находится недоминируемая точка $y^{(k+1)} \geq \tilde{y}^{(k+1)}$, на чем итерация завершается.

Отметим, что шаг 3 необходим из-за того, что ЛПР указывает точку на *аппроксимации* границы, так что эта точка может не лежать на точной границе Парето.

Этот метод достаточно прост для человека, который должен отвечать на простые вопросы и выбирать новую точку, изменяя значения только двух критериев. Ясно, что ЛПР имеет право на ошибку. Доказана сходимость метода при правильном поведении ЛПР и гладкости паретовой границы. Недостатком метода является его медленная сходимость.

Для ускорения сходимости метода предлагается его немного усложнить — изображать на картинке не одно сечение, а несколько, т.е. карту решений. При этом на карте решений изображаются несколько двумерных сечений, отличающихся значением какого-либо третьего критерия, скажем, y_s . Тогда $(k+1)$ -я итерация метода принимает следующий вид:

шаг 1. ЛПР получает текущий недоминируемый вектор

$$y^{(k)} = (y_1^{(k)}, \dots, y_i^{(k)}, \dots, y_j^{(k)}, \dots, y_m^{(k)})$$

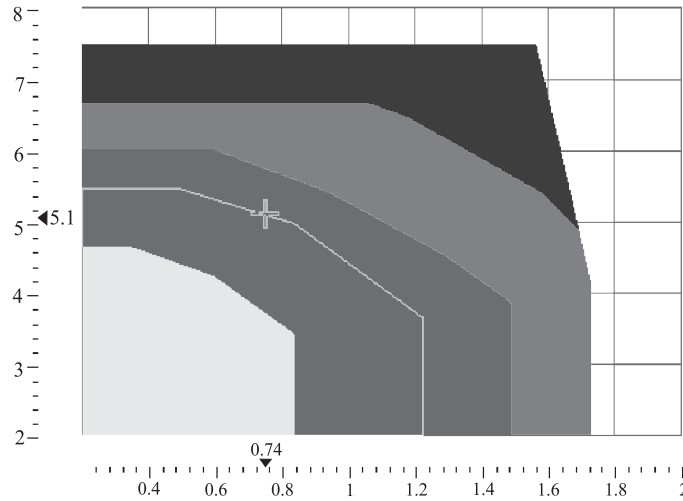
и выбирает три критерия, значения которых он хочет изменить. Пусть это y_i, y_j, y_s .

шаг 2. ЛПР получает на экране компьютера карту решений, причем критерии y_i, y_j расположены на осях, а значения критерия y_s задают сечения ОЭП (см. 15.4). Кроме неподвижных сечений, соответствующих заданным значениям y_s , изображается подвижная граница, положение которой определяется значением y_s , указанным на прокрутке. Текущая точка высвечивается на границе подвижного сечения. ЛПР предлагается передвинуть границу в желаемое положение и выбрать точку $y_l^+, l = i, j, s$, на ней. В результате этого получается новая точка:

$$\tilde{y}_l^{(k+1)} = \begin{cases} y_l^{(k)}, l \neq i, j, s \\ y_l^+, l = i, j, s \end{cases}.$$

шаг 3. По этой точке находится недоминируемая точка $y^{(k+1)} \geq \tilde{y}^{(k+1)}$.

Этот метод сходится быстрее, чем метод с единственным сечением, однако сходимость в случае пяти-семи критериев все же недостаточно быстрая. Заметим, что описанный метод является развитием метода STEM (в модификации Лаукса).



Об. 15.4.

15.6. Визуализация ОЭП в невыпуклом случае

Рассмотрим теперь вопрос о визуализации аппроксимации ОЭП в невыпуклом случае, которая, как было показано в предыдущем параграфе, строится в виде совокупности конусов с вершинами в точках, близких к границе Парето. Внутренняя аппроксимация ОЭП в задаче многокритериальной максимизации имеет вид

$$T^* = \cup \{y + (-\mathbb{R}_+^m) : y \in T\}.$$

Двумерное сечение одного конуса рассчитывается элементарно, а затем все сечения изображаются одновременно на экране компьютера одним цветом. В результате пользователь получает двумерное сечение объединения конусов. Для построения карты решений, как и в выпуклом случае, двумерные сечения, соответствующие различным значениям третьего критерия, окрашиваются в различные цвета и накладываются одно на другое. Как и в выпуклом случае, рядом с картой решений изображается шкала соответствия цвета и значения третьего критерия. В отличие от выпуклого случая, при визуализации ОЭП обычно изображаются “толстые сечения”, представляющие собой объединения сечений для небольшого диапазона значений, так что в шкале указываются границы диапазонов значений третьего сечения.

На рис. 15.5 изображена черно-белая копия цветного экрана, демонстрирующего границу Парето в пятикритериальной задаче МКО. Под картой

решений на экране компьютера изображены прокрутки, позволяющие с помощью слайдеров менять диапазоны значений четвертого и пятого критериев.

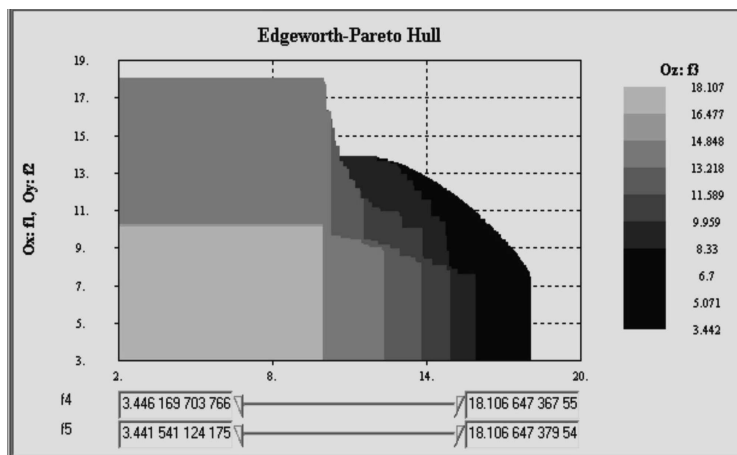


Рис. 15.5.

Как видно на рисунке, при большом числе точек можно увидеть критериальные замещения и даже норму критериального замещения в тех точках границы Парето, где она близка к гладкой. Как и в выпуклом случае, ЛПР имеет возможность указать предпочтительную целевую точку. После этого компьютер находит доминирующую ее точку в списке критериальных точек и соответствующее решение.

На рис. 15.6 приведена черно-белая копия экрана в прикладной четырехкритериальной задаче, связанной с оптимизацией охлаждения стали в процессе ее непрерывной разливки²⁾. В этой задаче требовалось уменьшать значения всех критериев. Как видно на рисунке, малым значениям третьего критерия (от 0 до 0.01) соответствует область, не включающая малые значения по другим показателям (заметим, что значение критерия, расположенного на прокрутке, выбрано равным нулю). Таким образом, идеальная точка, равная в этой задаче нулю по всем критериям, не достижима. Видно, что при увеличении значения третьего критерия появляются небольшие дополнительные возможности уменьшения значений критериев, отложенных на осях. Это, однако, не позволяет найти критериальную точку с нулевым значением одновременно для двух критериев, отложенных на осях. На карте решений ясно видны замещения между критериями, которые в данном случае имеют совсем не простой вид.

²⁾ A. Lotov, V. Berezkin, G. Kamenev, K. Miettinen. Optimal Control of Cooling Process in Continuous Casting of Steel Using a Visualization-Based Multi-Criteria Approach // Applied Mathematical Modelling, 29(7), 2005, 653-672.

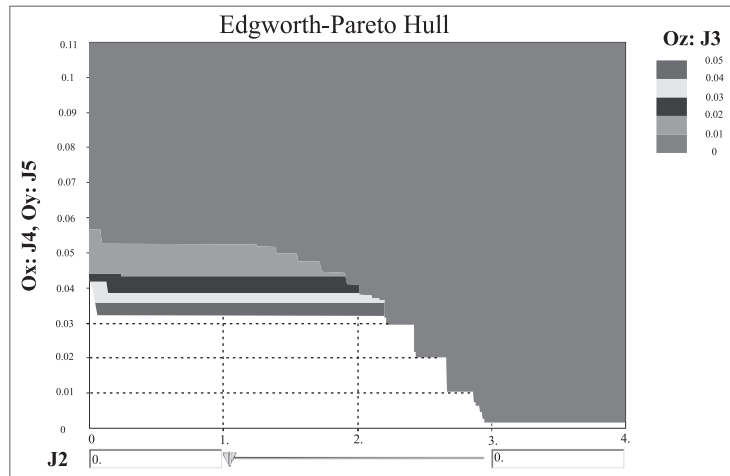


Рис. 15.6.

15.7. Визуализация границы Парето в сети Интернет

Принципиально важно, что визуализация ОЭП может быть легко реализована в сети Интернет. Поскольку с аппроксимацией ОЭП связано примерно 99% затрат вычислительных ресурсов и она отделена от работы пользователя по изучению недоминируемого множества, диалоговая визуализация ОЭП легко реализуется в структуре «сервер-клиент». Аппроксимация ОЭП при этом может быть осуществлена на достаточно мощном сетевом сервере, в то время как визуализация — на персональном компьютере пользователя средствами языка Java с использованием обычных браузеров сети.

Такая структура визуализации ОЭП была впервые реализована в демонстрационном ресурсе сети Интернет, предназначенном для независимого выбора решения в одной условной экологической проблеме. ОЭП была аппроксимирована заранее, после запроса пользователь получал по сети аппроксимацию этого множества и Java-апплет для ее визуализации. В процессе анализа ОЭП пользователь мог выбрать предпочтительную целевую точку, которая передавалась на сервер. Рассчитанная сервером стратегия решения экологической проблемы возвращалась пользователю через несколько минут, причем время ожидания определялось исключительно качеством работы сети. Этот ресурс, который также использовался в лабораторных работах, может служить прообразом средств сети Интернет, предназначенных для информирования неспециалистов о стратегиях решения экологических проблем. Применение сетевых средств визуализации недоминируемого множества дает любому человеку, умеющему хотя бы в минимальной степени пользоваться средствами сети Интернет, возможность

независимого сознательного выбора стратегии решения экологической проблемы. Такая возможность превращает Интернет в важное средство объективного информирования людей, стремящихся участвовать в процессе принятия решений.

Возможно также использование визуализации границы Парето с помощью средств сети Интернет для решения различных финансовых, производственных и других проблем. Эти вопросы, однако, выходят за рамки нашей книги.

Литература

- [1] *С.А. Ашманов* Введение в математическую экономику. М.: Наука, 1984.
- [2] *Ф.П. Васильев* Методы оптимизации, часть 1. М.: изд. МЦНМО, 2011.
- [3] *Ю.Б. Гермейер* Введение в теорию исследования операций. М.: Наука, 1971.
- [4] *В.А. Ильин, В.А. Садовничий, Бл.Х. Сендов* Математический анализ. М.: Проспект, 2004.
- [5] *М. Интрилигатор* Математические методы оптимизации и экономическая теория. М.: Айрис-пресс, 2002.
- [6] *Р.Л. Кини, Х. Райфа* Принятие решений при многих критериях: предпочтения и замещения. М.: Радио и связь, 1981.
- [7] *А.Н. Колмогоров, С.В. Фомин* Элементы теории функций и функционального анализа. М.: Физматлит, 2004.
- [8] *О.И. Ларичев* Объективные модели и субъективные решения. М.: Наука, 1987.
- [9] *О.И. Ларичев* Теория и методы принятия решений. М.: Логос, 2002.
- [10] *А.В. Лотов* Введение в экономико-математическое моделирование. М.: Наука, 1984.
- [11] *А.В. Лотов, В.А. Бушенков, Г.К. Каменев, О.Л. Черных* Компьютер и поиск компромисса. Метод достижимых целей. М.: Наука, 1997.
- [12] *В.В. Подиновский, В.Д. Ногин* Парето-оптимальные решения многокритериальных задач. М.: ФИЗМАТЛИТ, 2007.
- [13] *А.Г. Сухареv, А.В. Тимохов, В.В. Федоров* Курс методов оптимизации. М.: Наука, 1986.
- [14] *Р. Штойер* Многокритериальная оптимизация: теория, вычисления и приложения. М.: Радио и связь, 1992.