bhv

И. Г. Черноруцкий

МЕТОДЫ ОПТИМИЗАЦИИ

Компьютерные технологии

К оптимизации сводится огромное количество задач компьютерного моделирования из различных предметных областей. Широко распространено мнение о том, что для успешного решения таких задач достаточно иметь стандартное программное обеспечение и современные аппаратные средства. Однако, как показывает практика, прежде всего необходима высокая профессиональная подготовка специалистов в этой весьма сложной области компьютерных вычислений. Слишком много проблем возникает как при проведении реальных вычислений, так и при интерпретации получаемых результатов.

Научное Издание

И. Г. Черноруцкий

МЕТОДЫ ОПТИМИЗАЦИИ
Компьютерные технологии

Рекомендовано Учебно-методическим объединением по университетскому политехническому образованию в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлениям подготовки магистров «Системный анализ и управление» и «Информатика и вычислительная техника»

Санкт-Петербург

«БХВ-Петербург»

2011

УДК 681.3.06
ББК 32.973.26-018.2
Ч-49

Черноруцкий И. Г.

Ч-49 Методы оптимизации. Компьютерные технологии. — СПб.: БХВ-Петербург, 2011. — 384 с.: ил.

ISBN 978-5-9775-0784-4

В книге изложены теория, методы и основные элементы компьютерных технологий оптимизации. Наиболее подробно описаны методы решения конечномерных задач с учетом таких особенностей, как невыпуклость и плохая обусловленность минимизируемых функционалов. Рассмотрены многопараметрические и многокритериальные задачи. В качестве модельной предметной области выбраны задачи управления. Рассматриваемый материал иллюстрируется многочисленными примерами.

Для научных работников, преподавателей и студентов технических вузов

УДК 681.3.06
ББК 32.973.26-018.2

Группа подготовки издания:

Главный редактор	*Екатерина Кондукова*
Зам. главного редактора	*Татьяна Лапина*
Зав. редакцией	*Григорий Добин*
Редактор	*Анна Кузьмина*
Компьютерная верстка	*Натальи Караваевой*
Корректор	*Виктория Пиотровская*
Оформление обложки	*Елены Беляевой*
Зав. производством	*Николай Тверских*

Лицензия ИД № 02429 от 24.07.00. Подписано в печать 29.07.11.
Формат $70\times100^1/_{16}$. Печать офсетная. Усл. печ. л. 30,96.
Тираж 1000 экз. Заказ №
"БХВ-Петербург", 190005, Санкт-Петербург, Измайловский пр., 29.

Санитарно-эпидемиологическое заключение на продукцию
№ 77.99.60.953.Д.005770.05.09 от 26.05.2009 г. выдано Федеральной службой
по надзору в сфере защиты прав потребителей и благополучия человека.

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"
199034, Санкт-Петербург, 9 линия, 12.

ISBN 978-5-9775-0784-4

Оглавление

Предисловие

Эта книга представляет собой расширенную и в значительной степени модифицированную версию более ранних изданий: "Методы оптимизации"[1] и "Методы оптимизации в теории управления"[2].

В последнее время термин "оптимизация" часто употребляется "всуе", например, в области интернет-технологий или даже в парикмахерском деле — "оптимизация прически". Мы будем использовать его в обычном математическом смысле. В книге рассматриваются проблемы построения "наилучших" систем с помощью компьютерного моделирования. Оптимизируемые системы могут принадлежать к различным предметным областям. Например, это могут быть технические, экономические, экологические, программные, физические и другие системы. Важно лишь, чтобы существовали компьютерные модели, позволяющие вычислять "качество" или показатели эффективности системы в зависимости, например, от выбираемых (управляемых) параметров. В последнем случае имеем важный класс задач "параметрической оптимизации". Могут возникать и более сложные ситуации, когда требуется оптимальным образом выбрать не набор чисел (значений параметров), а некоторую функцию или функции. В особо сложных случаях речь не может идти о построении оптимальной системы. Требуется решить более "скромную" задачу — построить механизм последовательного улучшения характеристик оптимизируемой системы.

Соответствующие компьютерные технологии и являются предметом дальнейшего изложения.

В основу книги положен курс лекций по методам оптимизации, который читается автором на протяжении ряда лет на факультете технической кибернетики Санкт-Петербургского государственного политехнического университета. Тем не менее, книга позиционируется как монография, т. к. является "научным трудом, углубленно разрабатывающим ограниченный круг вопросов" (см. Современный словарь иностранных слов. — М.: Рус. Яз., 1992). Обсуждаемые в книге вопросы теории и практики оптимизации, действительно, не содержат многих традиционных разделов. С одной стороны, это вызвано ограниченным объемом, т. к. не ставилась задача создания "энциклопедии" или стандартного всеохватывающего учебника по оп-

[1] Черноруцкий И. Г. Методы оптимизации. — СПб.: Изд-во СПбГТУ, 1998.

[2] Черноруцкий И. Г. Методы оптимизации в теории управления. — СПб.: Питер, 2004.

тимизации. С другой стороны, и это главное, в книге отражены личные пристрастия автора, основанные на более чем тридцатилетнем опыте решения реальных задач. Многие суждения и выводы являются субъективными и основаны все на том же опыте. По-видимому, это нормальная ситуация, тем более что автор не претендует на "истину в последней инстанции" и всегда готов корректировать свою точку зрения.

В частности, в книге не рассмотрены специальные методы выпуклой оптимизации, методы линейного программирования, методы учета ограничений, основанные на операции проектирования на допустимое множество. Соответствующий материал присутствует во многих учебниках и монографиях по оптимизации, и автор здесь счел возможным его опустить. Не рассмотрены и многие другие вопросы, например, методы, основанные на теории двойственности. Так уж случилось, что автору в своей практической деятельности обычно "не везло". Задачи, которые приходилось решать, оказывались нелинейными, невыпуклыми, негладкими, плохо обусловленными *(жесткими)*. Часто встречались многопараметрические задачи с большим числом аргументов минимизируемых функционалов или бесконечномерные задачи. Поэтому как раз эти вопросы в книге обсуждаются достаточно подробно (за исключением бесконечномерных задач), вплоть до конкретных алгоритмов и примеров решения реальных задач. Красной нитью через все изложение проходит проблема *жесткости*, или *овражности*, присущая многим практическим оптимизационным (или вариационным) задачам. Один из видных разработчиков и исследователей *методов случайного поиска* Л. А. Растригин определял овражные оптимизационные задачи, как задачи, в которых вероятность убывания целевой функции при случайном выборе направления спуска чрезвычайно мала. И именно поэтому в данной книге из рассмотрения исключены также алгоритмы случайного поиска и, в частности, генетические алгоритмы. Их относительно невысокая эффективность для жестких или овражных задач подтверждается и практически.

Несмотря на кажущуюся неполноту изложения, на основе представленного материала, вообще говоря, можно решать достаточно разнообразные прикладные оптимизационные задачи, выбирая соответствующую траекторию в том иерархическом арсенале методов и средств, которые все-таки содержатся в книге. Эта траектория сама по себе может и не быть оптимальной, но, как правило, она приводит к результату.

Книга предназначена для многих категорий читателей. Во-первых, это студенты технических вузов, имеющих разделы по теории оптимизации в своих учебных планах. Во-вторых, это компьютерные математики-практики (современные инженеры). В-третьих, это могут быть профессиональные математики, разрабатывающие теорию оптимизации в функциональных пространствах. Дело в том, что в книге представлены некоторые результаты собственной научной деятельности автора и соответствующие новые подходы к решению жестких конечномерных задач. Обобщение этих результатов на бесконечномерный случай может представлять определенный интерес для математиков-теоретиков.

При изложении методов, алгоритмов и сопутствующих вычислительных технологий автор следовал системе взглядов Н. Н. Моисеева, который говорил, например, в предисловии редактора перевода к книге Ж. Сеа [63]: "Утверждение, что сходимость — это главное качество предлагаемого алгоритма, казалось вначале очевидным. Однако очень скоро обнаружилось, что это качество не является ни достаточным, ни необходимым для эффективного окончания вычислений, т. е. главной цели, ради которой создавался алгоритм". И еще: "Понимание того факта, что классическое представление о сходимости алгоритмов, как об основном содержании теории вычислительных процессов, не соответствует требованиям к анализу, которые выдвигает практика, является основным стимулом научных поисков". При этом Н. Н. Моисеев, являясь в первую очередь математиком, ни в коей мере не принижал значение выработанных в математике принципов.

Обсуждаемые в книге методы, технологии и алгоритмы, в основном, также получили соответствующее теоретическое обоснование вплоть до доказательства сходимости для некоторых стандартных случаев. Но наибольшую роль здесь играют исследования по обоснованию различных аспектов практического применения методов, отличных от установления самого факта сходимости.

Книга условно состоит из двух частей: *глава 1* и *главы 2—8*.

В *главе 1* рассмотрены общие математические структуры функционального анализа, позволяющие сформулировать основные положения теории оптимизации в банаховых и гильбертовых пространствах. Получены необходимые и достаточные условия экстремума. Обсуждаются с тех же единых позиций некоторые понятия вариационного исчисления. Например, уравнения Эйлера вариационного исчисления получаются как необходимые условия экстремума, вытекающие из равенства нулю градиента соответствующего функционала. Вводится определение минимизирующей последовательности в функциональном пространстве, как основного понятия практической оптимизации. Рассмотрены общие методы построения минимизирующих последовательностей (методы минимизации) в гильбертовых пространствах. При желании материал *главы 1* может быть опущен, если читателя интересуют исключительно прикладные вопросы оптимизации в конечномерных пространствах. Но этот материал оказывается совершенно необходимым в бесконечномерном случае и, в частности, при решении задач оптимального управления. Изложение является замкнутым, элементарным и формально не требует предварительной подготовки в этой области. Однако неявно предполагается, что читатель достаточно хорошо знаком с теорией конечномерных векторных пространств.

Главы 2—8 посвящены конечномерным задачам оптимизации (математическому программированию).

Основное внимание в этих главах уделено практическим вопросам оптимизации, тесно связанным с работой в различных предметных областях. При этом важнейшее значение приобретают элементы сопутствующих вычислительных технологий, связанные с формализацией конкретной прикладной проблемы и разработкой сценария оптимизации. Для изучения соответствующих вопросов, обычно выпадающих из стандартных руководств по теории и методам оптимизации, целесообразно вести изложение "на фоне" некоторой предметной области, а не в абстрактном виде. Это позволяет, с одной стороны, глубже понять возникающие проблемы, а с другой — получить наглядные и конкретные интерпретации применяемых технологий.

В качестве такой "модельной" предметной области в книге выбрана теория управления, изучаемая с различной степенью подробности по многим направлениям подготовки современных специалистов в области компьютерного моделирования систем. Таким образом, эта предметная область уже знакома большинству читателей данной книги. С другой стороны, важность такого выбора объясняется широким практическим применением аппарата современной алгоритмической теории управления при создании прикладных программных систем, в том числе, при разработке встроенных систем управления. Ведь когда мы говорим об управлении, то подразумеваем очень простую по своей сути задачу: необходимо сформировать такое целенаправленное воздействие на объект управления, чтобы перевести его в некоторое "желаемое" состояние. Понятно, что под эту простую схему подпадают многие ситуации, связанные как с нашей повседневной деятельностью, так и возникающие внутри технических, экономических, программных, эко и других систем. Обычно речь идет о некотором наилучшем в определенном смысле, т. е. "оптимальном" управлении, что и определяет необходимость создания теории управления, базирующейся на принципах оптимальности. Помимо обеспечения оптимальности самого управляющего воздействия, алгоритмизация и компьютерная реализация большинства этапов процесса управления сводится к решению целой цепочки различных оптимизационных задач. В этом смысле можно утверждать, что методы оптимизации являются важным алгоритмическим фундаментом современной теории управления.

Специфика формулируемых в теории управления оптимизационных задач отражает общую ситуацию, возникающую, как уже говорилось, при решении практически любых реальных задач. Поэтому и в этом смысле, изучая соответствующие оптимизационные проблемы на примере задач теории управления, можно получать общезначимые выводы и заключения.

Опыт компьютерного моделирования и проведения реальных оптимизационных вычислений, особенно в конечномерных пространствах, показывает, что, как правило, невозможно при создании прикладных программных систем эффективно использовать универсальное алгоритмическое обеспечение "в чистом виде" из-за резкого понижения скорости сходимости предлагаемых оптимизационных процедур. Весьма вероятной оказывается хорошо знакомая специалистам ситуация полной остановки ("заклинивания" или "залипания", что соответствует англоязычному термину jamming) алгоритма минимизации целевого функционала задолго до достижения искомого оптимума. Практика показывает, что один из основных факторов сложности решения реальных оптимизационных задач может быть связан с уже упомянутым специальным случаем явления плохой обусловленности, когда целевой функционал имеет "жесткий" или "овражный" характер, т. е. резко возрастает по одним направлениям и слабо изменяется вдоль других, что и вызывает указанные трудности.

Высокие степени жесткости возникают не в исключительных "паталогических" случаях, а отражают обычную ситуацию при решении практически любой прикладной задачи конечномерной оптимизации. Наиболее остро проблема жесткости стоит при решении задач параметрической идентификации объектов компьютерного моделирования и управления, а также в задачах оптимального параметрического синтеза как реальных, т. е. уже существующих, так и проектируемых систем при наличии ограничений и векторного критерия оптимальности.

В настоящее время уже трудно установить, когда впервые было указано на явление овражности, как на типичную практическую ситуацию, затрудняющую работу классических оптимизирующих процедур. Во всяком случае, уже в начале компьютерной эры в 1959 г. в фундаментальном труде А. А. Фельдбаума "Вычислительные устройства в автоматических системах" [76] были, по существу, рассмотрены все основные признаки овражной ситуации, включая явление "заклинивания" и понятие многомерного дна оврага. Широкую известность проблема овражности приобрела после работ И. М. Гельфанда и М. Л. Цетлина (1961—1962 г.), посвященных методам управления техническими системами и применению поисковых процедур в системах автоматической оптимизации. Эти исследования привели к созданию известного "метода оврагов". В 1967 г. Л. А. Растригин обобщил метод оврагов на общий случай овражной ситуации, когда размерность дна оврага может превышать единицу. Следующий шаг в изучении проблемы был предпринят в монографии Л. А. Растригина "Системы экстремального управления" (1974 г.), где в связи с общими вопросами применения поисковых методов параметрической оптимизации в системах экстремального управления рассмотрены различные аспекты овражной ситуации и, в частности, построены простейшие модели овражных экстремальных объектов [54]. Значительный вклад в теорию плохо обусловленных экстремальных задач внесли работы Ю. В. Ракитского, возглавлявшего ленинградскую школу по изучению проблемы жесткости и указавшего на принципиальную связь явления овражности с концепцией жесткости систем обыкновенных дифференциальных уравнений. Эти исследования были поддержаны будущим академиком РАН Н. С. Бахваловым и нашли отражение в брошюре Ю. В. Ракитского, С. М. Устинова и И. Г. Черноруцкого "Численные методы решения жестких систем обыкновенных дифференциальных уравнений" [56] и монографии тех же авторов "Численные методы решения жестких систем" [57]. В этих работах впервые была разработана общая теория систем, описываемых жесткими дифференциальными моделями, включающая асимптотическую теорию жестких систем и основы теории минимизации жестких функционалов.

Однако, несмотря на важность полученных результатов, проблема ими не исчерпывалась. Оставался нерешенным целый ряд существенных вопросов, связанных с развитием теории жестких (овражных) оптимизационных задач, а также общих принципов построения методов и алгоритмов конечномерной оптимизации по жестким целевым функционалам. Кроме того, представляла практический интерес разработка проблемно-ориентированного алгоритмического и программного обеспечения, учитывавшего специфику отдельных классов задач и структурные особенности применяемых на практике критериев качества. При построении таких методов и алгоритмов должны были учитываться такие естественные для приложений сопутствующие факторы сложности, как невыпуклость минимизируемых функционалов, их негладкость, а в ряде случаев и высокая размерность вектора аргументов. Такие исследования проводились в течение последних десятилетий, и были получены достаточно интересные для теории и практики результаты.

Сформулированные задачи составляют настолько важную алгоритмическую проблему при оптимизации современных технических, экономических, программных и других систем и методов управления ими, что назрела необходимость представить их в систематизированном виде. Данная книга, по мнению автора, позволяет частично восполнить этот пробел.

Дадим рекомендации читателям-непрофессионалам, собирающимся использовать книгу в качестве учебного пособия по соответствующим дисциплинам или разделам дисциплин. Есть хорошее замечание кого-то из известных людей: "Если Вы, перечитывая в очередной раз "Фауста" Гете, не находите для себя чего-то нового, то Вы остановились в своем развитии". Автор далек от мысли сравнивать данную книгу с "Фаустом" и, тем более, себя с Гете, однако для многих, изучающих теорию оптимизации, особенно студентов вузов, следует иметь в виду, что предлагаемая вашему вниманию книга также задумывалась как "многослойная". При первом чтении многие вопросы могут казаться неясными, хотя общая картина должна быть понятна. Когда вы лучше подготовитесь и изучите некоторые вопросы по указанной дополнительной литературе, можно приступать ко второму чтению и т. д. В итоге все зависит от вашей предварительной математической подготовки и опыта практической работы. В книге даются ссылки на учебную и научную литературу. Многие из этих публикаций являются бестселлерами и многократно переиздавались и переиздаются. Однако, как правило, ссылки сделаны на наиболее ранние "непереработанные" издания, которые часто оказываются (по мнению автора) и наиболее интересными. При желании читатель без труда обнаружит в каталогах и более поздние версии соответствующих публикаций.

Полное понимание представленных в книге материалов предполагает знакомство со стандартными курсами математики для высших технических учебных заведений (содержащими, в частности, достаточные сведения по линейной и матричной алгебре), численного анализа, теории вероятностей (в частности, корреляционной теории) и основ теории управления. Если вы владеете этим материалом в обычном для технических вузов объеме, то никаких проблем возникнуть не должно.

Для облегчения работы над книгой и согласования применяемых терминов и обозначений в *главе 2* даны необходимые для данного курса общие сведения из теории управления и сформулированы характерные постановки оптимизационных задач.

Представленные в книге элементы теории, методы и алгоритмы, которые можно отнести к разряду новых или, говоря более скромно, нетрадиционных, появились не вчера. В разное время и при разных обстоятельствах они обсуждались с акад. АН СССР (впоследствии РАН) Н. Н. Моисеевым, чл. корр. РАН П. А. Бутыриным, профессорами В. А. Городецким, В. Я. Катковником, А. А. Первозванским, Ю. В. Ракитским, Л. А. Растригиным, С. М. Устиновым и др. Всем им автор благодарен за сделанные замечания и поддержку.

Основные обозначения и терминологические замечания

R	Множество вещественных (действительных) чисел
R^n	n-мерное евклидово пространство
$\forall$, $\exists$	Кванторы всеобщности и существования
$f : X \to Y$	Отображение f множества X во множество Y
$k \in 1 : N$	Число k принимает последовательно все значения из множества натуральных чисел от 1 до N включительно
x^n	Последовательность элементов x^n
a, b , $\langle a, b \rangle$	Скалярное произведение векторов a и b
$A > 0$	Матрица A положительно определена
$\triangleq$	Равно по определению
$D = \ x \in H \mid P \ x$	Подмножество элементов множества H, обладающих свойством $P \ x$
$\varnothing$	Пустое множество
$C^k \ D$	Множество k раз непрерывно дифференцируемых на множестве D функций
$\cdot \ ^+$	Псевдообратная матрица
МНК	Метод наименьших квадратов
$\operatorname{diag} \lambda_i$	Диагональная матрица
$\operatorname{cond} \ A$	Спектральное число обусловленности матрицы A
J', J''	Вектор градиента и матрица вторых производных функционала $J \ x$
$\arg\min J \ x$	Минимизатор функционала $J \ x$
$\operatorname{Arg}\min J \ x$	Множество всех минимизаторов функционала $J \ x$

Сделаем также некоторые терминологические замечания по тексту книги. Согласно общепринятым математическим канонам функционального анализа (см., например, [43]) термин *"функционал"* означает (однозначное) отображение произвольного множества во множество вещественных чисел. В силу этого замечания вещественная функция от вещественного переменного, а также вещественная функция от нескольких вещественных переменных тоже являются функционалами. Иногда, особенно в прикладных технических публикациях, предлагается называть функционалом более частный случай отображения некоторого *множества функций* во множество вещественных чисел. Пример такого функционала дает определенный интеграл от интегрируемой на заданном промежутке функции. Мы, однако, будем придерживаться стандартной математической терминологии.

Под задачами *математического программирования* будем понимать конечномерные задачи оптимизации, т. е. задачи поиска максимума или минимума функционала, определенного на некотором подмножестве *конечномерного* евклидова пространства. В этом случае, например, задачи теории оптимального управления, формулируемые как задачи поиска оптимальных управляющих функций из некоторого подмножества бесконечномерного пространства, уже не будут относиться нами к задачам математического программирования.

Мы будем следовать сложившейся в математике традиции и называть определитель матрицы Гессе *гессианом*, определитель матрицы Якоби — *якобианом*, определитель матрицы Вронского — *вронскианом* и т. д. В некоторых книгах и учебниках по математическому программированию и оптимизации сама матрица Гессе называется гессианом, что является необоснованным и приводит к ненужной путанице.

Введение. Постановка задачи оптимизации

Исследование способов получать те или иные результаты (в широком смысле слова) наилучшим, наивыгоднейшим способом есть задача теории оптимизации.

Математически проблема оптимизации описывается следующим образом. Рассматривается некоторое непустое множество U элементов (вообще говоря, произвольной природы), называемое множеством допустимых элементов. Рассматривается некоторая ограниченная снизу функция (функционал) J, ставящая в соответствие каждому элементу множества допустимых элементов какое-либо действительное (вещественное) число. Требуется найти *минимизатор* — такой элемент $u^* \in U$, которому соответствует минимальное значение J:

$$J(u^*) \le J(u), \quad \forall u \in U. \tag{1}$$

Заданная функция J формализует понятие качества оптимизируемого объекта и часто (не совсем строго) называется критерием. На практике J может также отражать стоимость, время и т. п. Если заменить J на $-J$, то задача минимизации превратится в задачу максимизации и наоборот.

Приведенная постановка задачи выглядит достаточно общей, однако необходимо сделать некоторые уточнения и замечания.

Во-первых, здесь имеется в виду задача оптимизации с одним критерием оптимальности, задаваемым функционалом J. Возможны и более общие *многокритериальные* постановки задачи. Для конечномерного случая такие задачи будут рассмотрены далее достаточно подробно.

Во-вторых, задача (1) в общем случае может и не иметь решения. Искомый элемент u^* *(минимизатор)* может отсутствовать среди допустимых элементов. Например, функция $\exp(x)$, $x \in R$ не имеет минимизаторов на вещественной оси (множестве действительных чисел) R, хотя и ограничена снизу (инфимум равен нулю). Целесообразно поэтому рассматривать более общую задачу — задачу построения минимизирующих последовательностей: требуется найти последовательность элементов u_n множества U, удовлетворяющих условию:

$$\lim_{n \to \infty} J(u_n) = \inf_{x \in U} J(u). \tag{2}$$

Последовательность u_n называется *минимизирующей*. Здесь также возникают определенные трудности. Если минимизатор u^* существует (понятно, что он не обязан быть единственным), то минимизирующая последовательность может к нему не сходиться. Приходим к так называемым *некорректным задачам оптимизации*. И так как большинство численных методов оптимизации являются средством построения именно минимизирующих последовательностей, то необходимо совершенно четко понимать, что задача поиска хороших оценок соответствующих минимизаторов в случае некорректности не может быть решена без дополнительного исследования сходимости. Мы здесь должны различать известные из анализа типы сходимости: *по аргументу* и *по функционалу*. Чаще всего для практических приложений достаточно иметь сходимость по функционалу, т. к. неважно при каких допустимых значениях аргумента u достигается приемлемо малое значение J ("задачи аппроксимации"). И наоборот, иногда сама задача оптимизации ставится лишь для того, чтобы получить механизм определения хороших оценок минимизатора u^* ("задача идентификации"). В последнем случае совершенно необходимо исследовать сходимость по аргументу.

И наконец, в-третьих. Иногда в качестве множества допустимых элементов U выступает некоторое подмножество основного (базового) множества V, которое также задается. В этом случае имеем дело с т. н. задачей оптимизации с ограничениями, позволяющими выделить множество U из общего пространства V. Обычно ограничения задаются с помощью системы уравнений (например, дифференциальных) и неравенств.

Для математического анализа основных задач (1), (2) приходится наделять объекты, участвующие в постановке задач, определенными математическими структурами. Так множество U или V обычно считаются наделенными структурой банахова, гильбертова или конечномерного евклидова пространства. Это, как правило, оказывается достаточным обобщением структур, с которыми мы имеем дело при разработке и применении теории оптимизации. То же относится и к критериальной функции J, которая часто наделяется такими свойствами, как ограниченность, непрерывность, гладкость, выпуклость и т. д.

Отметим теперь несколько конкретных реализаций приведенных постановок задач. Для определенности будем говорить о задаче (1). Рассмотрим две альтернативные теории: конечномерные задачи оптимизации и бесконечномерные задачи.

Конечномерные задачи. Задачи математического программирования. *Математическим программированием* (МП) называется теория конечномерных оптимизационных задач (1). Минимизатор ищется среди элементов множества $U \subset R^n$. R^n наделяется структурой линейного n-мерного векторного пространства *(см. далее)*. Это серьезное ограничение, т. к. результаты теории математического программирования непосредственно неприменимы, если множество U является множеством объектов, не описываемых набором из n чисел, а является, например, множеством непрерывных на промежутке a, b функций. Множество допустимых элементов U в МП задается различными способами, например, соответствующей системой неравенств (или равенств и неравенств). Для формулировки задачи математического программирования вводится целевой функционал

$$J : R^n \to R$$

и функционалы ограничений

$$G_i : R^n \to R,\ i = 1,\ \ldots,\ m.$$

Множество допустимых элементов U может задаваться системой неравенств, например, вида:

$$U = \left\{ u \,\middle|\, u \in R^n;\ G_i(u) \geq 0,\ i = 1, \ldots, m \right\}.$$

Существуют и иные способы задания множества U.

Замечание. Равенство можно представить в виде двух разнонаправленных неравенств, но обычно целесообразно работать непосредственно с равенствами, если они изначально заданы в явном виде.

Если ограничения отсутствуют, то имеем задачу математического программирования без ограничений. Такие задачи в последующих разделах будут рассмотрены подробно, т. к. они часто оказываются алгоритмическим фундаментом при решении более общих задач.

В случае, если основные функционалы задачи МП являются линейными, то такая задача называется задачей *линейного программирования* (ЛП). Нелинейные задачи МП называются также задачами *нелинейного программирования* (НП). Существуют и другие классы задач МП.

Бесконечномерные задачи оптимизации. Типичными представителями бесконечномерных задач являются хорошо изученные задачи оптимального управления. Рассмотрим характерный пример.

Пусть управляемый процесс описывается системой обыкновенных дифференциальных уравнений (ОДУ)

$$\frac{dx}{dt} = f\left(t,\ x,\ u\right),\quad x\left(t_0\right) = x_0.$$

Здесь $x(t)$ — n-мерный вектор состояний системы; $u(t)$ — r-мерный вектор управлений, принадлежащий заданному множеству U допустимых вектор-функций. Например, U — множество кусочно-непрерывных на промежутке $[t_0,\ t_1]$ вектор-функций. Требуется определить такое управление $u(t) \in U$, чтобы минимизировать функционал

$$J\left(x,\ u\right) = \int_{t_0}^{t_1} F\left(t,\ x,\ u\right) dt.$$

Функция F задана. Естественно, в оптимальном управлении используются функционалы и с иной, более сложной структурой, а также более сложные системы ограничений.

Сформулированную задачу можно интерпретировать по-разному. Например, поиск ведется в пространстве функций $u(t)$. Тогда для вычисления J по заданному аргументу $u(t)$ решается заданная система ОДУ и полученная в результате зависимость $x(t)$ подставляется в выражение для J. Во втором варианте поиск ведется в пространстве (x, u), а система ОДУ выступает, как ограничение-равенство.

В книге даются необходимые элементарные сведения из функционального анализа, позволяющие без обращения к дополнительной литературе ознакомиться с важнейшими для теории оптимизации математическими структурами. Методы функционального анализа оказываются особенно важными при изучении задач бесконечномерной оптимизации. Как будет далее показано, с помощью относительно простых и достаточно общих конструкций функционального анализа можно с единых позиций охватить обширный материал, связанный с классическим вариационным исчислением, необходимыми и достаточными условиями экстремума и т. д. Основные выводы и понятия конечномерной оптимизации оказываются частными случаями полученных общих соотношений.

При рассмотрении элементов функционального анализа ставилась задача максимально разгрузить материал от деталей и в то же время дать логически завершенные математические конструкции, достаточные для формулировки необходимых для наших целей результатов теории оптимизации.

Глава 1

Математические основы. Элементы функционального анализа

1.1. Множества

В 1872 г. Георг Кантор, создатель теории множеств, определил множество как "объединение в одно целое объектов, хорошо различимых нашей интуицией или нашей мыслью". Это были времена наивной теории множеств. Такое положение долго не сохранилось. Интуитивное понимание множества оказалось в некоторых случаях логически порочным из-за антиномий или парадоксов теории множеств.

Антиномия Рассела (1903 г.). Для произвольного множества можно поставить вопрос о том, является ли оно своим собственным элементом. Например, множество студентов студентом не является и потому — не собственный элемент. В то же время множество всех множеств естественно считать содержащим себя как элемент. Рассмотрим теперь множество S всех множеств, не являющихся собственными элементами. Тогда получаем следующий обескураживающий результат: если предположить, что $S \in S$, то по определению S имеем $S \notin S$. С другой стороны, предположение $S \notin S$ приводит к следствию $S \in S$. Получили явное противоречие. Данный парадокс можно разрешить с помощью следующих более аккуратных рассуждений. Действительно, на самом деле вначале мы предположили существование множества S всех множеств, не являющихся собственными элементами. После этого получили, что $S \in S$ в том и только в том случае, если $S \notin S$. Из этого мы должны заключить, что такое множество S не существует. Указанное объяснение справедливо. Однако это мало способствует уменьшению парадоксальности рассмотренной ситуации. Действительно, тот факт, что не может существовать множество с четко очерченными свойствами (не содержащими себя в качестве элемента), столь же настораживает, как и прямое противоречие.

Чтобы исключить подобные антиинтуитивные результаты, в теории множеств был использован аксиоматический метод, позволяющий исправить ситуацию с помощью сознательного сужения (обеднения) интуитивного понятия множества.

Мы далее будем использовать дескриптивную (наглядную) теорию множеств. Нижеследующее введение в теорию множеств не является сколько-нибудь полным. Предполагается, что читатель знаком с элементами теории множеств из общих курсов по математике, и нам остается лишь согласовать обозначения, напомнить основные соотношения и сделать некоторые замечания.

Множества будем обозначать прописными буквами, а их элементы — строчными:

- $a \in A$ — элемент a принадлежит множеству A;
- $a \notin A$ — элемент a не принадлежит множеству A.

Если справедливо $\forall x \in A \Rightarrow x \in B$, то множество B включает множество A: $A \subset B$. (Здесь и далее двойная стрелка означает логическое следование.) $\forall$ — квантор общности. $\forall x$ читается: для любого x.

Равенство множеств: $(A \subset B \ \wedge \ B \subset A) \Rightarrow A = B$. ($\wedge$ — знак логического "И".)

Если X — множество, а P — некоторое свойство, то подмножество X, состоящее из тех x, для которых справедливо $P(x)$, обозначается как $\{x \in X / P(x)\}$. Пустое множество (не содержащее элементов) обозначается как $\varnothing$. Множество всех подмножеств некоторого множества A будем обозначать как $\Im(A)$.

1.1.1. Операции над множествами и их свойства

Объединение множеств:

$$A \cup B = \{x / (x \in A) \vee (x \in B)\}.$$

Здесь $\vee$ — знак логического "ИЛИ".

Пересечение множеств:

$$A \cap \mathrm{B} = \{x / (x \in A) \wedge (x \in B)\}.$$

Коммутативность операций объединения и пересечения множеств:

$$A \cup B = B \cup A; \ A \cap B = B \cap A.$$

Ассоциативность операций $\cup$ и $\cap$:

$$(A \cup B) \cup C = A \cup (B \cup C);$$

$$(A \cap B) \cap C = A \cap (B \cap C).$$

Дистрибутивность:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C);$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

Разность множеств:

$$A \setminus B = \{x \in A / x \notin B\}.$$

Декартово произведение множеств:

$$A \times B = \{(a, b) / a \in A, \ b \in B\}.$$

Здесь множества A и B могут совпадать.

Подмножество $R \subset A \times B$ называется *бинарным отношением*, или просто *отношением*, заданным на множествах A и B.

Переходим к важнейшим определениям, связанным с понятием функции или отображения.

1.1.2. Функции и отображения

Отношение $F \subset X \times Y$ называется *отображением* X в Y или *функцией*, определенной в X и принимающей значения в Y, если для $\forall x$:

$$(x, y_1) \in F \ \wedge \ (x, y_2) \in F \Rightarrow y_1 = y_2.$$

Если $(x, y) \in F$, то элемент y называется значением F в x и обозначается как $y = F(x)$.

Используются также обозначения:

$$F: X \to Y,$$

$$F: x \to F(x).$$

Принципиально важным свойством отображения является то, что любому значению "аргумента" x ставится в соответствие единственный элемент y. Такое понятие, как "многозначная" функция, здесь не рассматривается. Вполне правомерно, конечно, определить отображение, значениями которого являются подмножества некоторого данного множества, состоящие более чем из одного элемента. Но такое определение практически бесполезно, т. к. не удается разумным образом определить алгебраические операции над значениями таких функций. Например, операция извлечения корня из вещественного числа $\sqrt{z}$ приводит к двум значениям со знаками "плюс" и "минус". Но тогда как понимать равенство: $\sqrt{z} + \sqrt{z} = 2\sqrt{z}$? Левая часть имеет три разных значения, а правая — только два. (Хотя, конечно, существует и понятие Римановой поверхности.)

Сделаем еще одно замечание. Обычно (в "школьной" математике) различают понятия функции и ее графика. В данном выше определении эти понятия совпадают.

В современной математике важнейшую роль играет рассмотрение отображения (функции) как единого объекта (такого же, как точка или число) и проведение ясного различия между отображением F и любым из его значений $F(x)$. Первое есть элемент множества отображений X в Y, обозначаемого как $\Im(X \times Y)$, второе — элемент множества Y, причем

$$F = \{(x, y) \in X \times Y \,/\, y = F(x)\}.$$

Таким образом, отображение F есть некоторое множество упорядоченных пар (x, y).

Пусть $F: X \to Y$. Пусть также $A \subset X$. Тогда множество

$$F(A) = \{ y \in Y / \ \exists x \in A \quad y = F(x) \}$$

называется *образом множества A* при отображении *F*. Здесь $\exists$ — квантор существования. $\exists x$ читается: существует *x*. Прообразом множества $B \subset Y$ при отображении *F* называется множество

$$F^{-1}(B) = \{ x \in X / F(x) \in B \}.$$

Пусть $F: X \to Y$. Пусть $A \subset X$. Тогда множество

$$F \cap (A \times Y) \subset A \times Y$$

называется *сужением отображения F* на множестве *A*.

1.1.3. Виды отображений

Пусть $F: X \to Y$. Тогда:

1. Отображение F называется *сюръективным* (накрытием или *"отображением на"*), если $F(X) = Y$, т. е. для $\forall y \in Y \ \exists x \in X$ такой, что $y = F(x)$.
2. Отображение *F* называется *инъективным* (вложением), если $F(x) = F(x') \Rightarrow x = x'$.
3. Отображение F называется *биективным* (наложением, биекцией), если оно одновременно является сюръективным и инъективным. Будем называть биекцию взаимно-однозначным соответствием.

Отображение произвольного множества во множество действительных (вещественных) чисел называется вещественным *функционалом*. Иногда рассматриваемое множество чисел является множеством комплексных чисел, и мы приходим к понятию комплекснозначного функционала. Далее будут рассматриваться только вещественные функционалы. В этом смысле обычные "школьные" вещественные функции от вещественного аргумента также являются функционалами.

Пусть даны три множества *X*, *Y*, Z и два отображения

$$F: \ X \to Y,$$
$$G: \ Y \to Z.$$

Тогда $x \to G(F(x))$ есть отображение $H: X \to Z$, которое называется *композицией* отображений *G* и *F* (в указанном порядке) или сложным отображением и обозначается символом $\circ$: $H = G \circ F$.

1.1.4. Семейства элементов

Пусть *L* и *X* — два множества. Отображение $L \to X$ иногда называется также *семейством элементов* множества *X*, имеющим *L* множеством индексов и обозначается $\lambda \to x_\lambda$, или $(x_\lambda)_{\lambda \in L}$, или, когда это не может привести к недоразумению, просто (x_λ).

Чаще всего в качестве множества индексов L фигурирует множество целых положительных чисел N или его подмножество (конечное или бесконечное). В этом случае семейство элементов называется *последовательностью* (конечной или бесконечной). Иногда последовательности элементов будут изображаться с применением фигурных скобок. Следует отличать семейство $x_{\lambda}\ {}_{\lambda\in L}$ элементов множества X от подмножества множества X, состоящего из элементов этого семейства.

1.1.5. Счетные множества

Сравнение конечных множеств по числу элементов не вызывает затруднений. Можно сосчитать количество элементов в обоих множествах и сравнить результаты. Можно поступить и по-другому. Именно попытаться установить между сравниваемыми множествами взаимно-однозначное соответствие (биекцию). Ясно, что биекция между двумя конечными множествами может быть установлена тогда и только тогда, когда число элементов в них одинаково. Второй способ сравнения пригоден и для бесконечных множеств.

X называется *равномощным* множеству Y, если существует биективное отображение $X \to Y$. Множество называется *счетным*, если оно равномощно множеству N натуральных чисел. Бесконечное множество, не являющееся счетным, называется *несчетным* множеством. Можно доказать, что если бесконечное множество X счетно, то $\Im\ X$ — несчетно ("имеет мощность континуума").

Счетность множества четных чисел $n \to 2n$ представляет пример того, что бесконечные множества могут быть равномощны своему подмножеству. Всякое бесконечное множество содержит счетное подмножество. Таким образом, среди бесконечных множеств счетные множества являются "самыми маленькими".

Теорема 1.1. Множество действительных чисел из промежутка $0,\ 1$ несчетно.

Доказательство (диагональный процесс Кантора). Допустим противное и предположим, что удалось "перенумеровать" все действительные числа следующим образом:

$$\alpha_1 = 0.a_{11}a_{12}a_{13}\ldots$$
$$\alpha_2 = 0.a_{21}a_{22}a_{23}\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\alpha_n = 0.a_{n1}a_{n2}a_{n3}\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Тогда всегда можно построить действительное число β

$$\beta = 0.b_1 b_2 \ldots b_n \ldots$$
$$a_{nn} = 1 \to b_n = 2$$
$$a_{nn} \neq 1 \to b_n = 1, \quad n = 1,\ 2,\ 3,\ \ldots,$$

не входящее в перенумерованный список действительных чисел. Таким образом, доказано, что множество действительных чисел не может быть счетным. Мощность множества действительных чисел называется мощностью континуума.

> ***Упражнение 1.1***
> Постройте биекцию между множеством действительных чисел из отрезка [0, 1] и множеством точек квадрата со сторонами [0, 1].

1.2. Метрические пространства

В теории метрических пространств развивается геометрический язык, на котором выражаются результаты анализа. Этот язык позволяет придать этим результатам достаточную общность и вместе с тем дать наиболее простые и отражающие суть дела доказательства. Нас будут интересовать топологические аспекты теории метрических пространств, связанные с концепцией предельного перехода, а также алгебраические аспекты при изучении основных операций над элементами метрических пространств. Метрические пространства являются частным видом более общих топологических пространств.

Пусть E — некоторое множество. *Расстояние* в E есть отображение (функционал) d произведения $E \times E$ во множество действительных чисел R:

$$d : E \times E \to R .$$

Предполагается, что функционал d обладает следующими свойствами:

1. $\forall\ x, y :\ d(x, y) \geq 0.$
2. $d(x, y) = 0 \Leftrightarrow x = y.$
3. $\forall x, y :\ d(x, y) = d(y, x).$
4. $\forall x, y, z :\ d(x, z) \leq d(x, y) + d(y, z),\quad x, y, z \in E.$

Свойство 4 называется неравенством треугольника.

Множество E с заданным в нем расстоянием d называется *метрическим пространством*. Обычно это пространство, т. е. пару (E, d), обозначают одной буквой E.

Примеры метрических пространств.

1. Множество $C[a, b]$ всех непрерывных действительных функций, определенных на отрезке $[a, b]$ с расстоянием

$$d(f, g) = \max_{t \in [a, b]} \left| f(t) - g(t) \right|.$$

2. В этом же множестве функций можно ввести расстояние

$$d(f, g) = \left(\int_a^b \left(f(t) - g(t) \right)^2 dt \right)^{1/2}.$$

В результате получим пространство *непрерывных функций с квадратичной метрикой* $C^2[a, b]$.

Обычно та или иная метрика для построения метрического пространства выбирается в соответствии с особенностями решаемой задачи. Первая метрика в приведенном примере отражает достаточно жесткое требование к близости функций. Ее применяют, например, при решении задачи равномерного приближения функций (равномерная аппроксимация), когда требуется гарантировать, чтобы на всем отрезке $[a, b]$ отклонение функций f и g было меньше некоторой заданной величины. Вторая метрика отражает менее жесткое требование. Оно заключается в том, что для "подавляющего" большинства (но не для всех) значений аргумента t из $[a, b]$ абсолютная величина $|f-g|$ также мала. Во многих случаях, например, при обработке результатов наблюдений квадратичное приближение является наиболее приемлемым, т. к. оно позволяет сглаживать отдельные локальные выбросы аппроксимируемой (экспериментальной) функции с помощью некоторой теоретической расчетной зависимости. В итоге можно получить достаточно точное общее представление о характере протекающего процесса даже при наличии ошибок при измерении экспериментальных зависимостей.

1.2.1. Изометрия

Пусть E, E' — два метрических пространства с расстояниями, соответственно, d, d'. Тогда биективное отображение f: E, E' называется *изометрией*, если для $\forall (x, y) \in E\times E$ имеем

$$d'(f(x), f(y)) = d(x, y).$$

Если f: E, E' — биективное отображение, то в соответствии с определением биективности отношение $y = f(x)$ функционально не только по y, но и по x, т. е. это отношение определяет не только функцию E, E', но и $E' \to E$. Последнее отображение называется *обратным* к f и обозначается $x = f^{-1}(y)$. Оно само биективно. Если f — изометрия $E \to E'$, то тогда f^{-1} будет изометрией $E' \to E$. Важность понятия изометрии заключается в том, что любая теорема, доказанная в E и в формулировке которой участвуют только расстояния между элементами E, немедленно дает соответствующую теорему в любом изометрическом пространстве E'.

Если E, E' — два множества, причем E — метрическое пространство с расстоянием d, то при наличии биекции f: $E \to E'$ множество E' также становится метрическим пространством с индуцированным расстоянием d'. Говорят, что расстояние d' перенесено в E' при помощи отображения f.

В теории метрических пространств удобен геометрический язык. В частности, элементы метрического пространства часто называются точками.

1.2.2. Шары, сферы, диаметр, окрестности

Пусть заданы: метрическое пространство (E, d), $a \in E$ и действительное число $r > 0$. Тогда множества

$$B(a; r) = \{x \in E \,/\, d(a, x) < r\},$$

$$B'(a; r) = \{x \in E \,/\, d(a, x) \le r\},$$

$$S(a; r) = \{x \in E \,/\, d(a, x) = r\}$$

называются, соответственно, *открытым шаром* (с центром a и радиусом r), *замкнутым шаром* и *сферой*.

Пусть E — метрическое пространство и $A \subset E$, $B \subset E$, $A \ne \varnothing$, $B \ne \varnothing$. Тогда *расстоянием* от A до B называется число

$$d(A, B) = \inf_{\substack{x \in A \\ y \in B}} d(x, y).$$

Диаметром множества A называется число

$$\delta(A) = \sup_{\substack{x \in A \\ y \in A}} d(x, y).$$

Диаметр может быть и бесконечным.

Ограниченным множеством в E называется непустое множество, диаметр которого конечен.

Открытым множеством в метрическом пространстве (E, d) называется множество $A \subset E$ такое, что

$$\forall x \in A \ \exists r > 0 \Rightarrow B(x; r) \subset A.$$

Таким образом, любая точка открытого множества A входит в это множество вместе с некоторым, содержащим ее открытым шаром.

Пусть $A \subset E$, $A \ne \varnothing$. Тогда *открытой окрестностью* множества A называется любое открытое множество, содержащее A, а *окрестностью* множества A — любое множество, содержащее открытую окрестность A.

Если множество A состоит из одной точки $A = \{x\}$, то мы говорим об *окрестностях точки* x, а не множества $A = \{x\}$.

Фундаментальной системой окрестностей множества A называется семейство $\{U_\lambda\}$ окрестностей A, обладающее тем свойством, что любая окрестность A содержит хотя бы одно из множеств семейства $\{U_\lambda\}$.

Точка $x \in A$ называется *внутренней точкой* множества A, если A является окрестностью точки x. Множество всех внутренних точек множества A называется *внутренностью* множества A и обозначается $\overset{\circ}{A}$.

Внутренняя точка множества $E \setminus A$ называется *внешней точкой* множества A.

Точкой прикосновения множества A называется такая точка $x \in E$, любая окрестность которой имеет с A непустое пересечение, т. е. содержит хотя бы одну точку из A. Множество всех точек прикосновения множества A называется *замыканием* A и обозначается символом $\overline{A}$. По определению, *замкнутое множество* в E есть дополнение открытого множества, т. е. B замкнуто, если $E \setminus B$ открыто. $A \subset E$ замкнуто, если $A = \overline{A}$. Например, замкнутый шар есть замкнутое множество. $\overline{A}$ — наименьшее замкнутое множество, содержащее A.

Точка $x \in E$ называется *предельной точкой* множества $A \subset E$, если любая ее окрестность содержит бесконечно много точек из A. Сама предельная точка не обязана принадлежать множеству A. Можно дать определение замкнутого множества на языке предельных точек. Именно множество A замкнуто, если оно содержит все свои предельные точки.

Точка $x \in E$ называется *граничной точкой* множества A, если она является точкой прикосновения как A, так и $E \setminus A$. Граничная точка x множества A характеризуется тем свойством, что в любой ее окрестности содержится по крайней мере одна точка множества A и по крайней мере одна точка множества $E \setminus A$. Множество $Fr\ A$ всех граничных точек множества A называется *границей* множества A.

1.2.3. Сепарабельные пространства, подпространства, непрерывные отображения

Множество A в метрическом пространстве E называется *плотным* относительно множества B (или в B), если любая точка $x \in B$ есть точка прикосновения множества A, т. е. если $B \subset \overline{A}$. В этом случае любая окрестность любой точки $x \in B$ содержит точку множества A.

Множество A, плотное в E, называется *всюду плотным*. В этом случае $\overline{A} = E$. Метрическое пространство E называется *сепарабельным*, если в E существует не более чем счетное (т. е. конечное или счетное) всюду плотное множество. Например, множество рациональных чисел всюду плотно в множестве действительных чисел R, поэтому пространство действительных чисел сепарабельно. Точно так же сепарабельны пространства $C\ a, b$, $C^2\ a, b$ из приведенных ранее примеров.

Пусть $A \neq \varnothing$, $A \subset E$. Тогда сужение на $A \times A$ отображения d: $E \times E \to R$ является расстоянием в A и называется расстоянием, индуцированным в A расстоянием d пространства E. Метрическое пространство, определяемое этим индуцированным расстоянием, называется *подпространством* A метрического пространства E.

Пусть E и E' — два метрических пространства, d и d' — расстояния в них. Отображение f: $E \to E'$ называется *непрерывным в точке* $x_0 \in E$, если для $\forall \varepsilon > 0$ $\exists \delta > 0$ такое, что

$$d\ x, x_0\ < \delta \Rightarrow d'\ f\ x_0\ , f\ x\ < \varepsilon .$$

Можно сказать нагляднее. Отображение f: $E \to E'$ называется непрерывным в точке $x_0 \in E$, если точка $f(x)$ сколь угодно близка к точке $f(x_0)$, как только точка x достаточно близка к x_0.

Отображение f называется *непрерывным в пространстве* E (или просто непрерывным), если оно непрерывно в каждой точке пространства E.

Интересен следующий факт. Сужение отображения f: $E \to E'$ на подпространстве F метрического пространства E может оказаться непрерывным и в случае, когда f не является непрерывным ни в одной точке $x \in E$. Примером служит отображение f: $R \to R$, равное нулю на множестве Q рациональных чисел и единице на его дополнении (функция Дирихле). Сужение отображения f на Q постоянно и потому непрерывно.

Отображение f: $E \to E'$ называется *равномерно непрерывным*, если для $\forall \varepsilon > 0$ $\exists \delta > 0$ такое, что

$$\forall x, y: d(x, y) < \delta \Rightarrow d'(f(x), f(y)) < \varepsilon .$$

(Здесь важно, что ε и δ не зависят от x и y.) Ясно, что из равномерной непрерывности следует непрерывность. Обратное, вообще говоря, неверно. Например, отображение $x \to x^2$ непрерывно на R, но не равномерно непрерывно.

Упражнение 1.2

Докажите последнее утверждение.

1.2.4. Гомеоморфизмы, пределы, полные пространства

Отображение f: $E \to E'$ называется *гомеоморфизмом*, если f — биекция, а f и f^{-1} непрерывны. Два метрических пространства E и E' называются гомеоморфными, если существует гомеоморфизм $E \to E'$.

Упражнение 1.3

Докажите, что изометрия всегда будет гомеоморфизмом, но не наоборот.

Понятие гомеоморфизма, в частности, оказывается полезным в том смысле, что можно оперировать не с исходным метрическим пространством, а с гомеоморфным ему, а затем перенести результаты на основное пространство.

Пусть $\{x_n\}$ — последовательность точек в метрическом пространстве E. Тогда $a \in E$ называется *пределом последовательности* $\{x_n\}$, если для $\forall \varepsilon > 0$ $\exists n_0 \in N$ такое, что $n \ge n_0 \Rightarrow d(a, x_n) < \varepsilon$.

Используется запись

$$a = \lim_{n\to\infty} x_n \text{, или } \lim_{n\to\infty} d\left(a,\ x_n\right) = 0 .$$

$\{x_n\}$ называется сходящейся последовательностью.

1.2.5. Последовательности Коши, полные пространства

Последовательностью Коши в метрическом пространстве E называется последовательность $\{x_n\}$, удовлетворяющая условиям: для $\forall \varepsilon > 0\ \ \exists n_0$ такое, что

$$p \geq n_0,\ q \geq n_0\ \Rightarrow d\left(x_p,\ x_q\right) < \varepsilon .$$

Иногда последовательности Коши называются фундаментальными последовательностями.

Теорема 1.2. Любая сходящаяся последовательность есть последовательность Коши.

Доказательство. Пусть $a = \lim\limits_{n\to\infty} x_n$, тогда для $\forall \varepsilon > 0\ \ \exists n_0 \in N$ такое, что

$$n \geq n_0 \Rightarrow d\left(a,\ x_n\right) < \varepsilon/2 .$$

В силу неравенства треугольника имеем: если $p \geq n_0$, $q \geq n_0$, то

$$d\left(x_p,\ x_q\right) \leq d\left(a,\ x_p\right) + d\left(a,\ x_q\right) < \varepsilon .$$

Что и требовалось доказать.

Метрическое пространство E называется *полным*, если любая последовательность Коши сходится в E (разумеется, к точке пространства E).

Упражнение 1.4

Докажите, что промежуток (a, b), пространство $C^2[a,\ b]$ — неполные пространства, а пространство $C[a,\ b]$ — полное.

Фундаментальная важность полных пространств основывается на том, что для доказательства сходимости некоторой последовательности в таком пространстве достаточно установить, что она является последовательностью Коши (т. е. удовлетворяет критерию Коши). Главное различие между проверкой этого критерия и непосредственной проверкой определения сходящейся последовательности состоит в том, что в критерии Коши не нужно заранее знать значение предела. Сходимость оказывается внутренним свойством последовательности.

Теорема 1.3. В полном метрическом пространстве E любое замкнутое множество $F \subset E$ является полным подпространством.

Доказательство. Так как E полно, то последовательность Коши $\{x_n\}$ точек множества F сходится к $a \in E$, и так как $x_n \in F$, то a является точкой прикосновения множества F, т. е. $a \in \bar{F}$. Из замкнутости F следует, что $F = \bar{F}$, т. е. $a \in F$. Что и требовалось доказать.

В приведенном рассуждении заключается важность понятия замкнутости.

1.2.6. Принцип сжимающих отображений

Данный принцип используется при доказательстве теорем существования и единственности в различных разделах функционального анализа.

Дадим необходимые предварительные определения и утверждения.

1. Пусть (E, d) — метрическое пространство и f: $E \to E$ — отображение E в себя.

   Тогда отображение f называется *сжимающим*, если существует такое число α: $0 < \alpha < 1$, что для $\forall x,\ y \in E$ выполняется неравенство

   $$d(f(x), f(y)) \le \alpha d(x, y).$$

   В этом случае отображение f называется также *сжатием*.

2. Элемент $x \in E$ называется *неподвижной точкой* отображения f, если $f(x) = x$, т. е. неподвижные точки — это решения уравнения $f(x) = x$.

3. Любое сжимающее отображение непрерывно. Действительно, для $\forall x_0 \in E$ имеем

   $$d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \alpha\delta = \varepsilon,$$

   т. е. достаточно для $\forall \varepsilon > 0$ положить $\delta = \varepsilon/\alpha$.

Теорема 1.4 (принцип сжимающих отображений). Всякое сжимающее отображение, определенное в полном метрическом пространстве E, имеет одну и только одну неподвижную точку.

Доказательство.

1. Существование.

   Пусть x_0 — произвольный элемент E. Положим $x_{n+1} = f(x_n)$, $n = 0, 1, \ldots$ Докажем, что $\{x_n\}$ — последовательность Коши. Так как f — сжатие, имеем

   $$d(x_2, x_1) = d(f(x_1), f(x_0)) \le \alpha d(x_1, x_0)$$

   $$d(x_3, x_2) \le \alpha d(x_2, x_1) \le \alpha^2 d(x_1, x_0)$$

   $$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

   $$d(x_{n+1}, x_n) \le \alpha^n d(x_1, x_0).$$

Применяя многократно неравенство треугольника, получим

$$
\begin{aligned}
&d(x_{n+p}, x_n) \le d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \ldots + \\
&+ d(x_{n+1}, x_n) \le \alpha^{n+p-1} d(x_1, x_0) + \alpha^{n+p-2} d(x_1, x_0) + \ldots + \\
&+ \alpha^n d(x_1, x_0) = \alpha^n d(x_1, x_0)\left(\alpha^{p-1} + \alpha^{p-2} + \ldots + 1\right) \le \\
&\le \alpha^n d(x_1, x_0) \frac{1}{1-\alpha} = \frac{\alpha^n}{1-\alpha} d(x_1, x_0) \underset{n\to\infty}{\to} 0.
\end{aligned}
$$

Таким образом, $\{x_n\}$ — последовательность Коши. Так как E — полное пространство, то последовательность $x_n \underset{n\to\infty}{\to} a \in E$. Из непрерывности f следует, что последовательность $f(x_n)$ сходится к $f(a)$. Переходя к пределу в равенстве $x_{n+1} = f(x_n)$, получим $f(a) = a$. Существование доказано.

2. Единственность неподвижной точки.

   Пусть a и b — две неподвижные точки. Тогда

$$d(f(a), f(b)) = d(a, b) \le \alpha d(a, b) \Rightarrow d(a, b) = 0.$$

   Действительно, если $d(a, b) \ne 0$, имеем $d(a, b) \le \alpha d(a, b) \Rightarrow 1 \le \alpha$, что невозможно по условию. Теорема доказана полностью.

Пример. Теорема существования и единственности для дифференциальных уравнений.

Пусть дано дифференциальное уравнение с начальными условиями (задача Коши):

$$\frac{dx}{dt} = f(x), \quad x(t_0) = x_0. \tag{1.1}$$

Везде здесь имеются в виду обычные вещественные функции от вещественной переменной.

Предполагается, что функция f непрерывна в области G, содержащей точку x_0, и удовлетворяет условию Липшица:

$$\left| f(x_1) - f(x_2) \right| \le M \left| x_1 - x_2 \right|.$$

Докажем, что тогда на некотором сегменте $|t - t_0| \le \delta$ существует и только одно решение $x = \varphi(t)$ сформулированной задачи Коши, которое может быть найдено последовательными приближениями (теорема Пикара).

Из (1.1) следует

$$x(t) = x_0 + \int_{t_0}^{t} f(x(t))\, dt. \tag{1.2}$$

Выберем $\delta > 0$ из условия $M\delta < 1$. Обозначим через C пространство непрерывных на сегменте $|t - t_0| \le \delta$ функций с метрикой (расстоянием)

$$d(\varphi_1, \varphi_2) = \max_t |\varphi_1(t) - \varphi_2(t)|.$$

Такое пространство уже рассматривалось ранее. Можно доказать, что это пространство полно. Рассмотрим отображение F: $C \to C$, $\psi = F(\varphi)$, определяемое формулой

$$\psi(t) = x_0 + \int_{t_0}^{t} f(\varphi(t))\,dt, \quad |t - t_0| \le \delta.$$

Докажем, что F — сжатие.

Имеем:

$$\forall t: \ |\psi_1(t) - \psi_2(t)| = \left| \int_{t_0}^{t} f(\varphi_1)\,dt - \int_{t_0}^{t} f(\varphi_2)\,dt \right| \le \int_{t_0}^{t} |f(\varphi_1) - f(\varphi_2)|\,dt \le$$

$$\le M \int_{t_0}^{t} |\varphi_1 - \varphi_2|\,dt \le M\delta \max_t |\varphi_1 - \varphi_2| = M\delta d(\varphi_1, \varphi_2).$$

Отсюда следует, что и

$$d(\psi_1, \psi_2) = \max_t |\psi_1 - \psi_2| \le M\delta d(\varphi_1, \varphi_2) = \alpha d(\varphi_1, \varphi_2), \alpha = M\delta < 1.$$

Таким образом, уравнение (1.2) вида $x = F(x)$ имеет одно и только одно решение в пространстве C.

Теорема доказана.

1.2.7. Компактные пространства

Метрическое пространство E называется *вполне ограниченным*, если для $\forall \varepsilon > 0$ существует конечное множество $F \subset E$ такое, что $\forall x \in E: d(x, F) < \varepsilon$.

В теории метрических пространств это понятие выражает свойство метрического пространства быть *"приблизительно конечным"*. Ясно, что вполне ограниченное метрическое пространство является ограниченным множеством. Обратное, вообще говоря, неверно.

Метрическое пространство E называется *компактным*, если оно является полным и вполне ограниченным.

Пусть в метрическом пространстве E задана последовательность $\{x_n\}$. Тогда точка $b \in E$ называется *предельной точкой* последовательности $\{x_n\}$, если для любой окрестности V точки b и любого m $\exists n \ge m$ такое, что $x_n \in V$.

Ясно, что в любой окрестности предельной точки последовательности содержится бесконечно много точек последовательности (это утверждение может рассматриваться как альтернативное определение предельной точки).

Если последовательность x_n имеет предел, то он является единственной предельной точкой этой последовательности. Обратное, вообще говоря, неверно. Единственная предельная точка не обязана быть пределом последовательности.

Пример. Последовательность x_n действительных чисел, где $x_{2n} = 1/n$, $x_{2n+1} = n$, имеет нуль своей единственной предельной точкой, но не сходится к нулю.

Можно доказать, что если у последовательности x_n в компактном метрическом пространстве есть только одна предельная точка, то она является пределом этой последовательности. (Здесь видна роль компактности. Множество действительных чисел не является компактным.) Если E — компактное метрическое пространство, то любая бесконечная последовательность в E имеет по крайней мере одну предельную точку.

Компактным множеством в метрическом пространстве E называется такое множество A, для которого подпространство A пространства E компактно. Можно показать, что любое компактное множество в метрическом пространстве замкнуто.

Кроме того, в компактном пространстве E всякое замкнутое множество компактно. Действительно, такое множество, очевидно, вполне ограничено и как замкнутое подмножество полного пространства полно, т. е. компактно.

Относительно компактным множеством в метрическом пространстве E называется множество $A \subset E$, замыкание $\overline{A}$ которого компактно. Для того чтобы множество A в метрическом пространстве E было относительно компактно, необходимо и достаточно, чтобы любая последовательность точек множества A имела предельную точку в E.

> ***Упражнение 1.5***
> Покажите, что на множестве действительных чисел относительная компактность эквивалентна ограниченности.

Изучая метрические пространства, мы изучаем множества, между элементами которых введено расстояние. По существу, мы изучаем свойства понятия расстояния между объектами произвольной природы.

1.3. Линейные пространства

Часто приходится встречаться с объектами, над которыми производятся операции сложения и умножения на числа. Например, векторы в геометрии в трехмерном пространстве умножаются на числа и складываются. Вещественные функции вещественных аргументов умножаются на числа и складываются и т. п. Одни и те же операции производятся над совершенно разными объектами. Для того чтобы изучить все такие примеры с единой точки зрения, вводится понятие линейного (или векторного) пространства.

Пусть на множестве L элементов $x,\ y,\ z,\ \ldots$ заданы два отображения:

$$L \times L \to L,$$
$$L \times R \to L,$$

где R — множество действительных чисел ("вещественная прямая"). Обозначим эти отображения как:

$$x,\ y \quad \to x + y \in L,$$
$$x,\ \alpha \quad \to \alpha x \in L$$

соответственно.

Тогда множество L называется *действительным линейным пространством*, если для введенных отображений выполнены следующие требования:

1. $x + y = y + x$ (коммутативность).
2. $x + y \quad + z = x + \quad y + z$ (ассоциативность).
3. $\exists \theta \in L:\ \forall x \in L \Rightarrow x + \theta = x$ (существование нуля).
4. Для $\forall x \in L\ \exists\ -x\ \ \in L \Rightarrow x +\ \ -x\ \ = \theta$ (существование противоположного элемента).
5. $1 \cdot x = x \quad 1 \in R$.
6. $\alpha\ \ \beta x\ \ = \alpha\beta\ \ x$.
7. $\alpha + \beta\ \ x = \alpha x + \beta x$.
8. $\alpha\ \ x + y\ \ = \alpha x + \alpha y$.

Часто говорят о линейном *векторном* пространстве. Сами элементы L также называются векторами.

Если вместо множества R действительных чисел используется множество C комплексных чисел, то получим *комплексное линейное пространство*.

Непустое подмножество линейного пространства L называется *подпространством*, если оно само является линейным пространством по отношению к определенным в L операциям.

Пусть x_α — произвольное непустое *семейство элементов* линейного пространства L (счетность множества x_α не предполагается). Определение понятия семейства элементов было дано выше. Рассмотрим все подпространства линейного пространства L, содержащие заданную систему векторов x_α . Пересечение этих подпространств, очевидно, тоже будет подпространством. Это так называемое *наименьшее* подпространство, содержащее x_α . Оно называется подпространством, *порожденным* множеством x_α , или *линейной оболочкой* семейства элементов x_α .