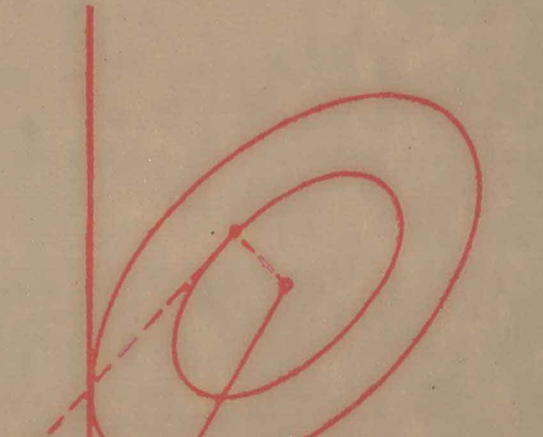
М. Базара, К. Шетти

Нелинейное программирование

Теория и алгоритмы



ИЗДАТЕЛЬСТВО «МИР»

Уважаемый читатель!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, издательство «Мир».

Мохтар Базара, К. Шетти

НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ.
ТЕОРИЯ И АЛГОРИТМЫ.

Науч. ред. И. А. Маховая
Мл. науч. ред. Т. А. Денисова
Художник В. А. Медников
Художественный редактор В. И. Шаповалов
Технический редактор Г. Б. Алюлина
Корректор А. Я. Шехтер

ИБ № 2637

Сдано в набор 03.04.81. Подписано к печати 09.02.82. Формат 60×90/16. Бумага типографская № 1. Гарнитура литературная. Печать высокая. Объем 18,25 бум. л. Усл. печ. л. 36,5. Усл. кр. отт. 36,5. Уч.-изд. л. 34,87. Изд. № 1/1225. Тираж 10000 экз. Зак. 1100. Цена 2 р. 80 к.

ИЗДАТЕЛЬСТВО «МИР»
129820, Москва, И-110, ГСП, 1-й Рижский пер., 2.

Ленинградская типография № 2 головное предприятие ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 198052, г. Ленинград, Л-52 Измайловский проспект, 29.

Нелинейное программирование
Теория и алгоритмы

ИЗДАТЕЛЬСТВО · МИР ·

Nonlinear Programming Theory and Algorithms

Mokhtar S. Bazaraa
C. M. Shetty

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia

John Wiley and Sons
New York Chichester Brisbane Toronto
1979

М. Базара, К. Шетти

Нелинейное программирование

Теория и алгоритмы

Перевод с английского
Т. Д. Березневой и В. А. Березнева

под редакцией Д. Б. Юдина

Издательство «Мир» · Москва 1982

ББК 22.143
17
УДК 51.38 + 519.9

Базара М., Шетти К.
Б 17 Нелинейное программирование. Теория и алгоритмы: Пер. с англ. — М.: Мир, 1982.
583 с.

Относительно простой, но достаточно строгий курс нелинейного программирования. Монография, написанная известными американскими специалистами, поможет подготовить инженеров к совместной с математиками работе по переводу прикладных задач на формальный язык.

Для инженеров и математиков-прикладников, специализирующихся в области нелинейного программирования и оптимизации.

Б $\frac{20204 - 020}{041(01) - 82}$ 20 — 82, ч.1 1502000000 ББК 22.143

Редакция литературы по математическим наукам

Предисловие редактора перевода

Оптимизационный подход к постановке и решению задач синтеза сложных систем является важным резервом повышения качества управления, планирования и проектирования. Выбор целей оптимизации и областей изменения управляющих параметров — задача конкретных общественных, экономических и технических наук. Разработка аппарата оптимизации — предмет математического программирования.

Общеизвестны рост и достижения линейного программирования в повышении эффективности экономического моделирования и оптимизации плановых решений. Более скромны успехи линейного программирования в технических приложениях, управлении и проектировании. Совершенствование технологических процессов и качества управления объектами различной природы, так же как и создание экономных и надежных конструкций, требует, как правило, учета нелинейных эффектов. В последние годы все чаще возникает необходимость в использовании нелинейных моделей и для описания экономических процессов.

Многие нелинейные задачи оптимизации, встречающиеся в экономике и технике, описываются в естественной постановке выпуклыми или вогнутыми функционалами и выпуклыми областями допустимых значений управляющих параметров. Это определяется как неформальными, так и формальными соображениями. Так, зависимость показателей эффективности технических систем от своих аргументов описывается обычно вогнутой функцией. Чем выше технические характеристики системы, тем труднее добиться приращения величины критерия ее эффективности. Производственные функции, характеризующие экономические системы, определяют вогнутую функциональную зависимость объема выпускаемой продукции от объемов используемых ресурсов. При больших капиталовложениях выпуск продукции на каждый рубль дополнительных вложений увеличивается меньше, чем при малых капиталовложениях. Чем больше объем производства, тем больше средств отвлекается на согласование и организацию взаимодействия между элементами системы.

Формальное описание широкого круга ситуаций, в которых функционалы и множества, определяющие нелинейную задачу математического программирования, оказываются выпуклыми, устанавливается содержательным истолкованием теоремы Кастена, обобщающей известную теорему Ляпунова.

Особое внимание, уделяемое выпуклому программированию, объясняется и тем, что для класса выпуклых задач общего вида и для различных его подклассов можно создать методы, гарантирующие решение любой задачи класса с заданной точностью за время, растущее относительно медленно с увеличением размерности задач. Для нелинейных задач более широких классов таких гарантий дать нельзя.

Предлагаемый вниманию читателя перевод представляет собой относительно элементарный и в то же время достаточно строгий и систематический курс нелинейной оптимизации. Этот курс выгодно отличается от опубликованных в последнее время монографий по различным аспектам теории и методов оптимизации своей практической ориентацией — он адресуется инженеру, ответственному за постановку практических задач. В своем предисловии авторы подробно характеризуют содержание книги, поэтому отметим лишь некоторые методические особенности изложения. Применение введенных в монографии понятий и методов иллюстрируется графическими и численными примерами. Кроме того, в каждой главе содержится большое количество упражнений различной степени трудности. Часть из них непосредственно связана с излагаемым материалом, другие расширяют круг вопросов, рассматриваемых в основном тексте. Имеются достаточно трудные задачи, рассчитанные на подготовленного читателя. В методическом отношении полезны приводимые в конце каждой главы замечания, библиографические справки, обобщения и прикладные и теоретические проблемы, связанные с излагаемым в соответствующей главе материалом. Все это способствует активному усвоению книги. Следует отметить, что работам советских специалистов по математическому программированию в книге достаточного внимания не уделено. Для удобства читателей мы сочли возможным привести дополнительный библиографический список, содержащий работы советских математиков по затронутым в книге вопросам, а также некоторые библиографические комментарии.

Книга М. Базара и К. М. Шетти заинтересует многочисленных специалистов-практиков в области исследования операций, управления, планирования и проектирования, а также будет полезна студентам соответствующих специальностей.

Д. Б. Юдин

Посвящается нашим родителям

Предисловие

Математическое программирование имеет дело с задачей оптимизации значений некоторой целевой функции при ограничениях типа равенств и неравенств. Задача, в которой все фигурирующие при ее описании функции линейны, называется *задачей линейного программирования*. В противном случае имеет место задача *нелинейного программирования*. Разработка симплекс-метода и появление быстродействующих вычислительных машин сделали линейное программирование важным инструментом решения многих проблем, возникающих в самых различных областях. Однако большинство реальных задач не может быть адекватно описано с помощью моделей линейного программирования из-за нелинейности целевой функции или некоторых ограничений. В последние два десятилетия значительного прогресса достигли исследования нелинейных задач. Настоящая книга представляет собой достаточно полное изложение этих достижений.

Книга разделена на три основные части: выпуклый анализ; условия оптимальности и двойственности; вычислительные методы. Конечной целью при изучении оптимизационных задач является построение эффективных вычислительных схем для решения практических задач. Выпуклый анализ, включающий в себя теории выпуклых множеств и выпуклых функций, играет важную роль при изучении области определения задач оптимизации. Условия оптимальности и теория двойственности могут использоваться как для получения критериев, отличающих оптимальное решение, так и для обоснования вычислительных методов.

При подготовке монографии особое внимание уделялось полноте и замкнутости изложения, дающим возможность использовать ее также в качестве учебного пособия либо справочника. Чтобы облегчить усвоение вводимых понятий и методов, в каждой главе приводятся детально разобранные примеры и иллюстрации. Кроме того, каждая глава содержит упражнения. Они включают в себя (1) простые числовые задачи, предназначенные для закрепления обсуждаемых вопросов; (2) задачи, дополняющие изложенный в главе основной материал;

(3) теоретические упражнения для более подготовленных читателей. Каждая глава завершается библиографическими сведениями и комментариями, которые будут полезны читателю при более глубоком изучении предмета. В конце книги приводится обширная библиография по математическому программированию.

Первая глава содержит примеры задач из разных областей инженерных дисциплин, которые могут быть формально сведены к задачам математического программирования. В частности, обсуждаются модели математического программирования, к которым сводятся задачи дискретного и непрерывного оптимального управления. Изложение иллюстрируется примерами управления производством и запасами. Рассматривается также пример, связанный со строительством шоссейных дорог. Приводятся примеры проектирования конструкций в строительной механике и машиностроении. Устойчивые состояния электрической цепи интерпретируются как оптимальные решения задачи квадратичного программирования. Исследуются нелинейные задачи, возникающие при управлении водными ресурсами. Наконец, обсуждаются нелинейные модели, возникающие в стохастическом программировании.

Остальные главы книги разделены на три части. Первая часть, содержащая гл. 2 и 3, посвящена анализу выпуклых множеств и выпуклых функций. Во второй главе обсуждаются топологические свойства выпуклых множеств, понятия опорной и разделяющей гиперплоскостей, многогранные множества и их свойства. Кроме того, здесь приводится краткий очерк линейного программирования. В третьей главе рассмотрены свойства выпуклых функций, и в частности субдифференцируемость и условия существования минимума и максимума на выпуклом множестве. Здесь же описываются обобщения понятия выпуклой функции и исследуется их взаимосвязь. Алгоритмы нелинейного программирования, построенные для выпуклых функций, могут быть использованы для решения задач более широкого класса, содержащих псевдовыпуклые и квазивыпуклые функции.

Вторая часть (гл. 4—6) посвящена условиям оптимальности и теории двойственности. В гл. 4 рассматриваются классические условия оптимальности Ф. Джона и Куна — Таккера как для задач с ограничениями-равенствами, так и для задач с ограничениями-неравенствами. В гл. 5 изучаются различные типы условий регулярности. В гл. 6 с помощью функции Лагранжа вводится понятие двойственности и исследуются условия оптимальности, связанные с седловыми точками. Обсуждаются теоремы двойственности, свойства двойственных функций и методы решения двойственных задач. Существуют также и другие определения двойственности. Наиболее обнадеживающим с точки зрения создания алгоритмов решения нелинейных задач представляется определение двойственности по Поэтомуку. Результаты, которые могут быть получены при разных подходах к двойственности, оказываются сравнимыми. Поэтому, а также в силу ограниченности объема монографии мы предпочли обсуждать в книге двойственность по Лагранжу, а остальные подходы ввели только в упражнениях.

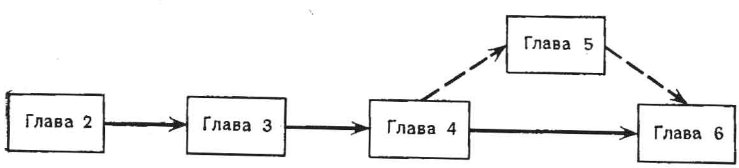
В третьей части, содержащей гл. 7—11, обсуждаются алгоритмы условной и безусловной оптимизации решения задач нелинейного программирования. Глава 7 имеет дело исключительно с теоремами сходимости. При этом алгоритмы рассматриваются как точечно-множественные отображения. Эти теоремы используются для доказательства сходимости методов, изложенных в последующих главах книги. Приведено также краткое описание критериев, которые могут быть использованы для оценки качества алгоритмов. Глава 8 посвящена безусловной оптимизации. Здесь обсуждаются различные методы одномерного поиска и методы минимизации функции нескольких переменных. Отдельно излагаются методы, использующие и не использующие производные. Рассмотрены методы, основанные на понятии сопряженности векторов. В гл. 8—11 доказана сходимость всех описанных методов. Из-за ограниченности места вопросы, связанные с порядком сходимости, кратко рассмотренные в гл. 7, более не обсуждаются. В гл. 9 приводятся методы барьерных и штрафных функций для нелинейного программирования. При использовании этих методов задача сводится к последовательности задач безусловной оптимизации. В десятой главе рассмотрены методы возможных направлений, заключающиеся в том, что сначала ищется некоторое направление спуска из допустимой точки, и затем новая допустимая точка находится минимизацией целевой функции вдоль этого направления. Обсуждается оригинальный метод, предложенный Зойтендейком и модифицированный затем для обеспечения сходимости Топкисом и Вейноттом. В гл. 10 представлены также такие разновидности метода возможных направлений, как метод проекции градиента Розена, метод приведенного градиента Вулфа и выпуклый симплексный метод Зангвилла. В гл. 11 рассматриваются специальные задачи с линейными ограничениями, которые могут быть решены при помощи несильно модифицированного симплекс-метода. В частности, здесь приводятся задачи квадратичного, сепарабельного и дробно-линейного программирования. Для решения задач квадратичного программирования используется описанный в этой главе алгоритм дополнительного ведущего преобразования, принадлежащий Лемке.

Предлагаемую вниманию читателя книгу можно использовать и как справочник по нелинейному программированию, и как руководство для занимающихся исследованием операций, управлением, прикладной математикой и инженерными дисциплинами, связанными с численными методами оптимизации. Материал книги требует некоторой математической зрелости и знания линейной алгебры и методов вычислений. Для удобства читателей в приложении А собраны основные математические определения и утверждения, часто используемые в книге.

Книга может быть использована для подготовки курса «Основы оптимизации» или «Вычислительные методы нелинейного программирования» по указанным ниже схемам. По материалам книг можно также построить двухсеместровый курс, охватывающий обе темы.

1. *Основы оптимизации*

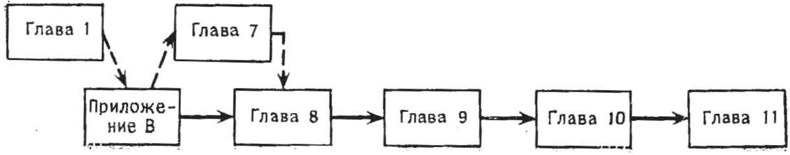
Этот курс предназначается для студентов, специализирующихся по прикладной математике, или для аспирантов факультетов, где прикладная математика не является профилирующей дисциплиной. Предполагаемое содержание курса схематически приведено ниже. Курс может быть прочитан за один семестр. Главу 5 об условиях регулярности можно опустить, не нарушая непрерывности изложения. Кроме того, те, кто знаком с линейным программированием, могут пропустить § 2.6.



2. *Вычислительные методы нелинейного программирования*

Этот курс предназначен для студентов и аспирантов, интересующихся алгоритмами решения задач нелинейного программирования. Схема курса приведена ниже. Он может быть прочитан за один семестр. Лица, заинтересованные главным образом в приложениях, могут опустить гл. 7 и обсуждение сходимости, проводимое в гл. 8—11.

Минимальные сведения из выпуклого анализа и условий оптимальности, необходимые для изучения гл. 8—11, собраны для удобства в приложении В. Первая глава, содержащая много примеров качественных задач, которые могут быть формально описаны задачами нелинейного программирования, будет хорошим введением к этому курсу. Однако если ее опустить, то непрерывность изложения не нарушится.



Авторы благодарят доктора Роберта Н. Лехрера (Robert N. Lehrer, Director of the School of Industrial and Systems Engineering) из Технологического института (шт. Джорджия) за поддержку при подготовке монографии. Авторы активно обсуждали различные вопросы с доктором Дж. Дж. Гудом (Jamie J. Goode of the School of Mathematics) из Технологического института (шт. Джорджия). Ему принадлежит пример 7.3.3 замкнутости сложных отображений. Авторы глубоко признательны ему за дружбу и сотрудничество. Наконец, авторы благодарят г-жу Кэролин Пиерсма, г-жу Джоан Оуэн и г-жу Кай Уоткинс за успешную расшифровку и перепечатку черновиков монографии, которые иногда выглядели так, будто бы они были написаны по-арабски или на языке хинди.

Атланта, Джорджия
1 января 1979 г.

М. Базара, К. Шетти

1

Введение

Инженеры и специалисты по исследованию операций часто сталкиваются с необходимостью решения оптимизационных задач. На практике встречаются разнообразные в содержательном смысле задачи оптимизации. Это могут быть задачи проектирования, задачи распределения ограниченных ресурсов, задачи расчета траектории полета ракеты и т. п. В недалеком прошлом в задачах такого рода вполне приемлемыми считались решения из довольно широкой области. В инженерном проектировании, например, обычно задавался большой коэффициент надежности. Однако с развитием производства все больше ужесточаются требования, предъявляемые к приемлемому проекту. В других областях деятельности, таких, как проектирование космических кораблей, условия функционирования проектируемых систем предъявляют экстремальные требования к характеристикам проекта. Таким образом, существует очевидная потребность в решении следующих важных вопросов. Каково наиболее эффективное использование имеющихся ресурсов? Можно ли получить более экономный в том или ином смысле проект? В каких пределах можно считать риск допустимым? Важность и актуальность этих проблем вызвали в последние три десятилетия интенсивные разработки моделей и методов оптимизации. Этому способствовало также увеличение количества и быстродействия вычислительных машин.

С другой стороны, развитие моделей и методов оптимизации стимулировалось значительным увеличением размерности и сложности оптимизационных задач, вызванным существенным технологическим подъемом после второй мировой войны. Инженеры и руководители производства оказались вынужденными учитывать все существенные факторы и их взаимосвязи, влияющие на качество принимаемых решений. Некоторые из этих связей не всегда даже могли быть поняты и нуждались, таким образом, в дополнительном анализе и проверяемых гипотезах. Успехи в методах измерений и статистических методах проверки гипотез существенно помогли в установлении взаимодействия между различными компонентами исследуемых систем.

В настоящее время методы исследования операций довольно широко применяются инженерами, экономистами, специалистами, ответственными за принятие решений в военной, культурной, административной и других областях. Это объясняется, по-видимому, уже накопленным положительным опытом практического использования моделей и методов оптимизации. На раннем послевоенном этапе применения исследования операций в планировании и управлении производством, как правило, использовались методы линейного программирования и статистического анализа. Теперь хорошо известны эффективные методы и машинные программы для решения такого рода задач. В этой книге рассматриваются нелинейные задачи, устанавливаются важнейшие свойства их оптимальных решений и обсуждаются различные вычислительные методы нелинейного программирования.

В настоящей главе приводится постановка общей задачи нелинейного программирования, а также рассматриваются некоторые простые практические задачи, описываемые моделями нелинейного программирования. Авторы стремились в этой главе прежде всего показать читателю многочисленные источники прикладных задач нелинейного программирования, не утруждая его глубоким исследованием каждой конкретной задачи.

1.1. Постановка задачи и основные определения

Общая задача *нелинейного программирования* имеет вид

$$\begin{array}{ll} \text{минимизировать} & f(\mathbf{x}) \\ \text{при условиях} & g_i(\mathbf{x}) \leqslant 0, \quad i = 1, \ldots, m, \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, l, \\ & \mathbf{x} \in X. \end{array}$$

Здесь $f, g_1, \ldots, g_m, h_1, \ldots, h_l$ — определенные на E_n функции, X — множество из E_n, $\mathbf{x}$ — вектор с компонентами $x_1, \ldots, x_n$. Задача заключается в нахождении переменных $x_1, \ldots, x_n$, удовлетворяющих ограничениям и отвечающих при этом минимальному значению функции $f(\mathbf{x})$.

Функцию f обычно называют *целевой функцией*, или *критерием оптимальности*. Каждое условие $g_i(\mathbf{x}) \leqslant 0$, $i = 1, \ldots, m$, называют *ограничением-неравенством* или *ограничением в форме неравенства*, а условие вида $h_i(\mathbf{x}) = 0$, $i = 1, \ldots, l$, — *ограничением-равенством* или *ограничением в форме равенства*. Вектор $\mathbf{x} \in X$, удовлетворяющий всем ограничениям, называют *допустимым решением*, или *допустимой точкой*. Совокупность всех допустимых точек образует *допустимую область*. Таким образом, задача нелинейного программирования заключается в нахождении такой допустимой точки $\bar{\mathbf{x}}$. для которой $f(\mathbf{x}) \geqslant f(\bar{\mathbf{x}})$ при всех допустимых решениях $\mathbf{x}$. Точка $\bar{\mathbf{x}}$ называется *оптимальным решением* или просто *решением* задачи.

Ясно, что задача нелинейного программирования может быть сформулирована как задача максимизации $f(\mathbf{x})$, а ограничения-неравенства записаны в виде $g_i(\mathbf{x}) \geqslant 0$, $i = 1, \ldots, m$. В специальном случае, когда целевая функция $f(\mathbf{x})$ линейна и все ограничения, включая соотношения, описывающие множество X, могут быть представлены в виде линейных равенств и/или неравенств, сформулированная выше задача называется *задачей линейного программирования.*

В качестве примера рассмотрим следующую задачу:

$$\begin{aligned} &\text{минимизировать} && (x_1 - 3)^2 + (x_2 - 2)^2 \\ &\text{при условиях} && x_1^2 - x_2 - 3 \leqslant 0, \\ &&& x_2 - 1 \leqslant 0, \\ &&& -x_1 \leqslant 0. \end{aligned}$$

Целевая функция и три функции ограничений имеют вид

$$\begin{aligned} f(\mathbf{x}) &= (x_1 - 3)^2 + (x_2 - 2)^2, \\ g_1(\mathbf{x}) &= x_1^2 - x_2 - 3, \\ g_2(\mathbf{x}) &= x_2 - 1, \\ g_3(\mathbf{x}) &= -x_1. \end{aligned}$$

На рис. 1.1 изображена допустимая область. Задача заключается в нахождении такой точки из допустимой области, для которой $(x_1 - 3)^2 + (x_2 - 2)^2$ имеет наименьшее возможное значение. Заметим, что точки (x_1, x_2), удовлетворяющие равенству $(x_1 - 3)^2 + (x_2 - 2)^2 = c$, лежат на окружности радиуса $\sqrt{c}$ с центром в точке (3, 2). Для каждого неотрицательного c такая окружность называется *линией уровня* целевой функции, отвечающей заданному значению c. Таким образом, задача заключается в нахождении минимального c, при котором хотя бы одна точка окружности принадлежит допустимой области. Иными словами, нужно найти окружность наименьшего радиуса, которая пересекала бы допустимую область. Как видно из рис. 1.1, такая окружность наименьшего радиуса соответствует $c = 2$ и пересекает допустимую область в единственной точке (2, 1). Поэтому (2, 1) — оптимальное решение, и значение целевой функции в этой точке равно 2.

Использованный выше подход состоит в нахождении оптимального решения с помощью построения линии уровня целевой функции, отвечающей наименьшему значению, при котором