Л. М. АБРАМОВ, В. Ф. КАПУСТИН

МАТЕМАТИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

ЛЕНИНГРАДСКИЙ ОРДЕНА ЛЕНИНА
И ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ им. А. А. ЖДАНОВА

Л. М. АБРАМОВ, В. Ф. КАПУСТИН

МАТЕМАТИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

Допущено Министерством высшего и среднего
специального образования СССР
в качестве учебного пособия для студентов вузов,
обучающихся по специальности «Экономическая кибернетика»

ИЗДАТЕЛЬСТВО ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА
ЛЕНИНГРАД 1981

Печатается по постановлению
Редакционно-издательского совета
Ленинградского университета

УДК 519.95
Абрамов Л. М., Капустин В. Ф. **Математическое программирование:** Учеб. пособие. Л.: Изд-во Ленингр. ун-та, 1981. 328 с. Ил. — 39, библиогр. — 52 назв.

Книга написана на основе курса «Математические методы исследования операций», который читается на отделении экономической кибернетики экономического факультета Ленинградского университета. В ней содержатся теоретические основы выпуклого и дискретного программирования и алгоритмы решения соответствующих задач. Рассматриваются некоторые экономические ситуации, которые формализуются как задачи математического программирования. Строгость изложения сочетается с возможностью освоения вычислительных алгоритмов без предварительного разбора их обоснования и изучения теоретического материала.

Книга рассчитана на студентов экономических вузов и факультетов, а также на специалистов экономических служб предприятий и ведомств.

Рецензенты: кафедра исследования операций ф-та вычислительной математики и кибернетики Московского университета (зав. кафедрой проф. П. С. Краснощеков), проф. В. А. Залгаллер (Ленинградское отделение математического института им. В. А. Стеклова)

$А\,\dfrac{20204-009}{076(02)-81}\,65-80.1502000000$

© Издательство Ленинградского университета, 1981 г.

ИБ № 1038

Абрамов Леонид Михайлович
Капустин Виталий Федорович

Математическое программирование

Редакторы *А. А. Гранаткина, Г. И. Чередниченко*
Художественный редактор *В. Н. Васильев*
Технический редактор *А. В. Борщева*
Корректор *Л. И. Чулкова*

Сдано в набор 29.02.80. Подписано в печать 11.12.80. М 33560. Гарнитура литературная. Печать высокая. Печ. л. 20,5. Формат 60×90$^1/_{16}$. Бум. тип. № 1. Уч.-изд. л. 19,61. Тираж 10000 экз. Заказ 631. Цена 95 коп.
Издательство ЛГУ им. А. А. Жданова, 199164, Ленинград, В-164, Университетская наб., 7/9.

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.

ОГЛАВЛЕНИЕ

此为试读，需要完整PDF请访问：www.ertongbook.com

ПРЕДИСЛОВИЕ

Эта книга возникла на основе курса лекций, который авторы в течение ряда лет читали на отделении экономической кибернетики экономического факультета Ленинградского университета. В ней содержатся теория выпуклого программирования, алгоритмы решения линейных, дискретных и нелинейных экстремальных задач. Обсуждаются вопросы, связанные с экономико-математическим моделированием.

Уже издано довольно много книг и статей, которые полезно использовать как учебные пособия или справочный материал по отдельным разделам курса математического программирования. Однако некоторые из них (например, [6, 11, 30, 48]) давно стали библиографической редкостью, другие изданы малым тиражом и практически недоступны студентам; наконец, только в [31, 32] обсуждаются все разделы типовой программы курса математического программирования для специальности «Экономическая кибернетика». Все это вызвало необходимость в создании настоящего учебного пособия.

Содержание книги в основном соответствует учебным программам по указанному курсу. Особую роль играет первая глава, включающая элементы выпуклого анализа. Ее материал используется как для обоснования теоретических положений и методов, так и для геометрических интерпретаций, проясняющих основные идеи выпуклого программирования. Мы старались сделать изложение более доступным, но, тем не менее, предупреждаем читателя, что изучение курса потребует от него некоторых усилий, совершенно естественных при работе с математической литературой. Особо отметим, что предполагается знакомство читателя с определенными разделами курсов линейной алгебры и математического анализа.

Форма изложения не является традиционной. Здесь мы продолжаем следовать идеям, положенным в основу [1]. Отметим наиболее важные моменты.

В книге рассмотрена теория выпуклого программирования (глава 3). Теория линейного программирования, в частности теоремы двойственности для задач линейного программирования, следует из общей теории. Однако при изложении вычислительных методов мы вынуждены были отказаться от такого подхода: алгоритмы решения задач линейного программирования предшествуют вычислительным методам решения нелинейных задач, поскольку используются в методах выпуклого, квадратичного и дискретного программирования.

Алгоритмы решения задач следует воспринимать не только как вычислительные и логические предписания, но и как конструкции, которые можно использовать для эффективного доказательства некоторых утверждений.

Алгоритмы не доведены до машинной реализации. Они могут служить лишь основой для написания машинных программ. Принципы и методы составления программ не являются предметом изучения настоящего курса.

Предполагается возможным выборочное чтение некоторых глав и отдельных параграфов. Книга написана так, чтобы можно было изолированно изучать теорию и вычислительные методы, знакомиться в первую очередь с вычислительными аспектами и лишь потом — с идеями и точным обоснованием. В какой мере это удалось — судить читателю.

В конце каждой главы приведены упражнения. Некоторые из них сформулированы непосредственно в тексте. Выполнение упражнений мы настоятельно рекомендуем читателю. Для проведения практических занятий можно использовать задачники [14, 17, 24], последний из них согласован с материалом настоящего пособия.

О системе нумерации и ссылок. Пункты параграфа имеют двойную нумерацию: первое число соответствует номеру параграфа в данной главе, второе — номеру пункта в этом параграфе. Ссылки на пункты внутри главы осуществляются естественным образом: например, ссылка «см. п. 3.6» в тексте четвертой главы означает, что мы отсылаем читателя к названному пункту этой главы. При ссылках на пункты других глав перед номером пункта добавляется номер главы. Так, ссылка «см. п. 4.3.6» в тексте шестой главы означает, что мы отсылаем читателя к п. 3.6 гл. 4. Упражнения нумеруются двумя числами: первое указывает номер главы, второе — номер упражнения в этой главе.

Об одной особенности: для j-й компоненты вектора $x \in R^n$ мы ввели обозначение $x[j]$ (в примерах, как правило, используется обычное обозначение x_j). Это позволяет, как нам кажется, упростить изложение и обоснование некоторых алгоритмов.

Конец доказательства, а также окончание алгоритма отмечается знаком ■.

Указатель литературы не претендует на полноту. На большинство работ имеются прямые ссылки в тексте учебного пособия, остальные [4, 5, 7—9, 12, 25, 26, 36—40, 50, 51] также включены в указатель, так как мы считали полезным обратить на них внимание читателя.

Мы благодарим сотрудников кафедры исследования операций Ленинградского университета и участников семинара этой кафедры за полезное обсуждение книги [1], а также С. А. Ашманова, П. С. Краснощекова и В. Ю. Лисицына за ценные советы и рекомендации. Мы глубоко благодарны В. А. Залгаллеру, особенно тщательно изучившему рукопись. Его замечания и советы, несомненно, способствовали улучшению книги.

Глава 1

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

Содержание этой главы несколько разнородно. Первые два параграфа включают обозначения, понятия и термины из линейной алгебры и математического анализа. В этих параграфах сформулированы также некоторые утверждения. Поскольку предполагается, что читатель знаком с названными курсами, утверждения приведены здесь без доказательства. В остальных параграфах достаточно подробно и с полными доказательствами рассматривается теория выпуклого анализа в R^n — теоретическая основа настоящего курса математического программирования. Понимание этой теории требует определенной математической культуры. Поэтому читателям, интересующимся, главным образом, вычислительными проблемами математического программирования, следует начать чтение книги со второй главы, используя первую только как необходимый справочный материал.

§ 1. Некоторые обозначения, понятия и термины линейной алгебры

1.1. Через N обозначается множество всех натуральных чисел, через R, как обычно, — множество вещественных чисел или вещественная прямая, через R^n — n-мерное вещественное векторное (или аффинное) пространство, элементами которого являются упорядоченные n-ки вещественных чисел ($n \in N$). Элементы этого пространства будем называть точками, векторами или планами и обозначать строчными буквами латинского алфавита a, b, ..., x, y, ... Если $x \in R^n$, то $x = (x[1], x[2], \ldots, x[n])$ или $x^{\text{т}} = (x[1], x[2], \ldots, x[n])$ (т — знак транспонирования). Иногда там, где это не может вызвать недоразумений, главным образом в примерах, будем писать $x = (x_1, \ldots, x_n)$. Через e_i, $i \in 1:n$, всегда будем обозначать элементы *канониче-*

ского базиса пространства R^n:

$$e_i[j]=\begin{cases}1, & \text{если} \quad j=i,\\ 0, & \text{если} \quad j\in 1:n\setminus\{i\}.\end{cases}$$

Числа $x[j]$ — координаты вектора x в каноническом базисе — будем называть *компонентами* вектора x.

Если $a\in R^n$, $b\in R^n$, то через ab обозначаем *скалярное произведение* векторов a и b: $ab=\sum_{i=1}^n a[i]\,b[i]$; произведение матриц $A=(a[i,j],\ i\in 1:k,\ j\in 1:m)$ и $B=(b[i,j],\ i\in 1:m,\ j\in \in 1:n)$ обозначается через AB; если $k=1$ или $n=1$, будем писать aB или Ab. Знак транспонирования «т» в записи векторов и при умножении векторов и матриц ставится лишь тогда, когда его отсутствие может вызвать недоразумения; в записи скалярного произведения знак транспонирования не ставится. Матрица, обратная к квадратной невырожденной матрице A, обозначается через A^{-1}, единичная матрица — через I или I_n, если возникает необходимость указать ее порядок. Строку матрицы A с номером i будем обозначать через a_i, а ее j-й столбец — через a^j.

1.2. *Размерность подпространства* T пространства R^n обозначается через $\dim T$. Она равна числу векторов базиса пространства T ($\dim\{0\}=0$ по определению).

Пересечение любого множества подпространств пространства R^n есть подпространство в R^n. Если T_1 и T_2 — подпространства в R^n, то $T_1+T_2=\{x\in R^n : x=t_1+t_2,\ t_1\in T_1,\ t_2\in T_2\}$ — *сумма подпространств* T_1 и T_2 — подпространство в R^n.

Всякое подпространство T пространства R^n может быть представлено как множество решений некоторой однородной системы линейных уравнений

$$\begin{array}{l} a[1,1]\,x[1]+\ \dots\ +a[1,n]\,x[n]=0,\\ \dots\dots\dots\dots\dots\dots\\ a[m,1]\,x[1]+\ \dots\ +a[m,n]\,x[n]=0,\end{array} \tag{1}$$

или в матричной записи $Ax=\mathbf{0}$, где $A=(a[i,j],\ i\in 1:m,\ j\in \in 1:n)$, $x=([1],\ \dots,\ x[n])$. Система (1) называется *уравнением подпространства* T. Верно и обратное: множество решений системы (1) всегда есть подпространство в R^n, и его размерность равна $n-r(A)$, где через $r(A)$ обозначен ранг матрицы A. Заметим, что при построении уравнения подпространства T матрицу A можно выбрать такой, чтобы $m=r(A)=n-\dim T$.

Через $\mathscr{L}(M)$ обозначим *линейную оболочку* множества $M\subset R^n$ — наименьшее подпространство, содержащее множество M. Линейная оболочка $\mathscr{L}(M)$ состоит из всех (конечных) линейных комбинаций элементов множества M:

$$\mathscr{L}(M)=\left\{x\in R^n : x=\sum_{i=1}^k \alpha_i a_i,\ a_i\in M,\ \alpha_i\in R,\ k\in N\right\}.$$

Размерность подпространства $\mathscr{L}(M)$ равна максимальному числу линейно независимых векторов в множестве M.

1.3. *Векторы* $a, b \in R^n$ называются *ортогональными*, если их скалярное произведение $ab=0$. *Множества* M_1 и M_2 в R^n называются *ортогональными*, если каждый вектор из M_1 ортогонален каждому вектору из M_2. Множество всех векторов, ортогональных непустому множеству $M \subset R^n$, называется *ортогональным дополнением* множества M и обозначается через $M^\perp$. Ортогональное дополнение непустого подмножества пространства R^n есть подпространство в R^n. Если T — подпространство, то $(T^\perp)^\perp = T$. Если T — подпространство в R^n и $T^\perp$ — его ортогональное дополнение, то $T \cap T^\perp = \{0\}$, $T+T^\perp = R^n$.

Сумма попарно ортогональных подпространств $T_1, T_2, \ldots, T_k$ называется *ортогональной суммой подпространств*. Для обозначения ортогональной суммы применяется знак $\oplus: T_1 \oplus T_2 \oplus \ldots$ $\ldots \oplus T_k = \sum_{i=1}^k \oplus T_i$. Если $T = \sum_{i=1}^k \oplus T_i$, то $\dim T = \sum_{i=1}^k \dim T_i$.

Представление пространства R^n в виде ортогональной суммы его подпространств T_i, $i \in 1:k$, называется *разложением пространства R^n в ортогональную сумму подпространств*. При этом если $R^n = \sum_{i=1}^k \oplus T_i$, то $\sum_{i=1}^k \dim T_i = n$ и любой вектор $a \in R^n$ однозначно представим в виде $a = \sum_{i=1}^k a_i$, $a_i \in T_i$. Вектор a_i называется *ортогональной проекцией* вектора a на подпространство T_i. В частности, $R^n = T \oplus T^\perp$, $\dim T^\perp = n - \dim T$ и любой вектор $a \in R^n$ однозначно представим в виде $a = b + c$, $b \in T$, $c \in T^\perp$.

1.4. Пусть $a \in R^n$, $b \in R^n$, $a \neq b$. Множество $L = \{x \in R^n: x = (1-\alpha)a + \alpha b,\ \alpha \in R\}$ называется *прямой*, проходящей через точки a и b. Ее подмножества $[a, b] = \{x \in R^n:\ x = (1-\alpha)a + \alpha b,\ \alpha \in [0, 1]\}$ и $L^+ = \{x \in R^n: x = (1-\alpha)a + \alpha b, \alpha \geqslant 0\}$ называются соответственно *отрезком*, соединяющим точки a и b, и *лучом*, идущим из a через b. Заметим, что отрезок можно определить и при $a = b$: в этом случае он состоит из единственной точки.

Множество $H \subset R^n$ называется *аффинным*, если $H \neq \varnothing$ и вместе с любыми двумя точками содержит прямую, проходящую через эти точки. Всякое подпространство в R^n является аффинным множеством (множество, состоящее из одной точки, — аффинное).

Если H — аффинное множество и $a \in H$, то множество $H - a = \{x \in R^n: x = h - a,\ \ h \in H\}$ — подпространство в R^n. Обратно, если T — подпространство в R^n и $a \in R^n$, то $T + a$ — аффинное множество в R^n. Таким образом, класс аффинных множеств состоит из всех подмножеств пространства R^n, получаемых сдвигами подпространств пространства R^n (на векторы из R^n).

Если H — аффинное множество, то существует единственное подпространство T, сдвигом которого получается H: $T = H - a$,

$a \in H$. *Размерность аффинного множества H* по определению равна размерности подпространства T, сдвигом которого получается H. Всякая прямая есть аффинное множество размерности единица.

Непустое пересечение любого множества аффинных множеств есть аффинное множество. Сумма $H_1 + H_2 = \{x \in R^n : x = x_1 + x_2,\ x_1 \in H_1,\ x_2 \in H_2\}$ аффинных множеств H_1 и H_2 есть аффинное множество. — **Упражнение 1.2.**

1.5. *Аффинной оболочкой* непустого множества $M \subset R^n$ называется наименьшее аффинное множество, содержащее M, т. е. аффинная оболочка множества M есть аффинное множество, содержащее M и содержащееся в любом аффинном множестве, содержащем M. Поэтому аффинная оболочка множества M может быть определена как пересечение всех аффинных множеств, содержащих множество M. Аффинную оболочку множества M обозначим через $\operatorname{aff} M$.

Если $M \subset R^n$ и $a \in M$, то $\operatorname{aff} M = a + \mathscr{L}(M - a)$. В частности, *если множество M конечно*: $M = \{a_0, a_1, \ldots, a_k\}$, то

$$\begin{aligned}\operatorname{aff} M &= a_0 + \mathscr{L}(M - a_0) = \\ &= \left\{x \in R^n : x = a_0 + \sum\nolimits_{i=1}^{k} \alpha_i (a_i - a_0),\ \alpha_i \in R\right\} = \\ &= \left\{x \in R^n : x = \sum\nolimits_{i=0}^{k} \alpha_i a_i,\ \alpha_i \in R,\ \sum\nolimits_{i=0}^{k} \alpha_i = 1\right\}.\end{aligned}$$

Линейная комбинация $\sum_{i=1}^{m} \alpha_i a_i$ векторов $a_1, \ldots, a_m$ называется *аффинной линейной комбинацией*, если $\sum_{i=1}^{m} \alpha_i = 1$. Таким образом, аффинная оболочка множества M совпадает с множеством всех аффинных линейных комбинаций элементов M. Заметим, что прямая $L = \{(1 - \alpha) a + \alpha b, \alpha \in R\}$ есть аффинная оболочка множества $\{a, b\}$.

1.6. Всякое аффинное множество пространства R^n может быть представлено как множество решений некоторой системы

$$\begin{array}{c} a[1,\ 1]\, x[1] + \ldots + a[1,\ n]\, x[n] = b[1], \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a[m,\ 1]\, x[1] + \ldots + a[m,\ n]\, x[n] = b[m] \end{array} \tag{2}$$

линейных уравнений, или в матричной записи $Ax = b$, где $A = (a[i,\ j],\ i \in 1:m,\ j \in 1:n)$, $b^{\mathrm{T}} = (b[1], \ldots, b[m])$, $x = (x[1], \ldots, x[n])$. Система (2) называется *уравнением аффинного множества H*. Верно и обратное: если система (2) совместна, то множество ее решений есть аффинное множество в R^n размерности $n - r(A)$, где $r(A)$ — ранг матрицы A. Систему (2) можно построить так, чтобы $m = r(A) = n - \dim H$.

1.7. Аффинное множество $H \subset R^n$, $\dim H = n - 1$, называется *гиперплоскостью* в R^n. Всякая гиперплоскость есть аффинная оболочка такой системы $\{a_1, a_2, \ldots, a_n\}$ из n векторов, что векторы $a_2 - a_1, \ldots, a_n - a_1$ линейно независимы (эти век-

此为试读，需要完整PDF请访问：www.ertongbook.com

торы образуют базис подпространства, сдвигом которого получается H).

Если H — гиперплоскость в R^n, то существуют такой вектор $a \in R^n$, $a \neq 0$, и число β, что $H = \{x \in R^n : ax = \beta\}$. Уравнение $ax = \beta \left(\sum_{j=1}^{n} a[j]\, x[j] = \beta\right)$ называется *уравнением гиперплоскости*, *a — нормальным вектором гиперплоскости H*; вектор a ортогонален подпространству $T = \{x \in R^n : ax = 0\}$, сдвигом которого получается H. Другими словами, каждая гиперплоскость в R^n есть множество уровня соответствующей линейной функции $f(x) = ax$.

Верно и обратное, всякое уравнение $ax = \beta$, $a \neq 0$, $\beta \in R$, определяет гиперплоскость в R^n, причем равносильные уравнения определяют одну и ту же гиперплоскость.

С каждой гиперплоскостью $H = \{x \in R^n : ax = \beta\}$ связывают два множества: $\Pi^+ = \{x \in R^n : ax > \beta\}$ и $\Pi^- = \{x \in R^n : ax < \beta\}$. Эти множества называются *полупространствами*. Говорят, что гиперплоскость *H делит* пространство R^n на два полупространства Π^+ и Π^-: если точки c и d принадлежат одному полупространству (Π^+ или Π^-), то $[c, d] \cap H = \varnothing$; если точки c и d принадлежат разным полупространствам, то $[c, d] \cap H \neq \varnothing$. Очевидно также, что $\Pi^+ \cap \Pi^- = \varnothing$, $\Pi^+ \cup \Pi^- \cup H = R^n$.

Множества $\bar{\Pi}^+ = \Pi^+ \cup H$ и $\bar{\Pi}^- = \Pi^- \cup H$ называются *замкнутыми полупространствами*, порожденными гиперплоскостью H.

1.8. Симметричная квадратная матрица $A = (a[i, j], i, j \in 1:n)$ называется *неотрицательно определенной*, если для любого $x \in R^n$

$$x^{\mathrm{T}} A x = \sum_{i=1}^{n} \sum_{j=1}^{n} a[i, j]\, x[i]\, x[j] \geqslant 0,$$

и *положительно определенной*, если $x^{\mathrm{T}} A x > 0$ для любого $x \in R^n$, $x \neq 0$. Аналогично симметричная матрица A называется *неположительно определенной*, если $x^{\mathrm{T}} A x \leqslant 0$ для любого $x \in R^n$, и *отрицательно определенной*, если $x^{\mathrm{T}} A x < 0$ для любого $x \in R^n$, $x \neq 0$.

Матрица A положительно (неотрицательно) определена в том и только том случае, если матрица $-A$ отрицательно (неположительно) определена. Известны эффективные критерии положительной (отрицательной) и неотрицательной (неположительной) определенности матрицы.

Пусть A — квадратная симметричная матрица порядка n. Определители

$$\Delta_1 = a[1, 1], \quad \Delta_2 = \begin{vmatrix} a[1, 1] & a[1, 2] \\ a[2, 1] & a[2, 2] \end{vmatrix}, \ \ldots$$

называются главными минорами матрицы A. Следующее условие положительной (отрицательной) определенности матрицы называется *условием Сильвестра*:

для положительной определенности матрицы A необходимо и достаточно выполнение неравенств $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_n > 0$;

для отрицательной определенности матрицы A необходимо и достаточно выполнение неравенств $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, ...

Условия $\Delta_1 \geqslant 0$, $\Delta_2 \geqslant 0$, ..., $\Delta_n \geqslant 0$ ($\Delta_1 \leqslant 0$, $\Delta_2 \geqslant 0$, ...) являются необходимыми, но не достаточными для неотрицательной (неположительной) определенности матрицы. Вычислительные процедуры проверки неотрицательной (неположительной) определенности матрицы обычно более трудоемки, чем процедуры проверки положительной (отрицательной) определенности. Например, матрица является неотрицательно определенной (неположительно определенной), если все корни ее характеристического многочлена (они все вещественные и их число с учетом кратностей равно n) неотрицательны (неположительны). Есть и другие способы проверки неотрицательной (неположительной) определенности матрицы.

§ 2. Некоторые обозначения, понятия и термины математического анализа

2.1. В R^n вводится следующее *отношение порядка* ($\Leftrightarrow$ — знак равносильности):

$$a < b \Leftrightarrow a[i] < b[i] \quad \textit{для любого} \quad i \in 1:n;$$

$$a \leqslant b \Leftrightarrow a[i] \leqslant b[i] \quad \textit{для любого} \quad i \in 1:n \quad \textit{и} \quad a \neq b;$$

$$a \leqq b \Leftrightarrow a[i] \leqslant b[i] \quad \textit{для любого} \quad i \in 1:n.$$

Пусть $J \subset 1:n$ (J — некоторое подмножество индексов из $1:n$). Обозначим

$$R_J^{n+} = \{x \in R^n : x[j] \geqslant 0,\ j \in J\}, \qquad R^{n+} = \{x \in R^n : x \geqq 0\},$$

в частном случае $R^+ = \{x \in R : x \geqslant 0\}$.

2.2. *Нормой*, или *длиной*, вектора $a \in R^n$ называется число $\|a\| = \sqrt{aa}$. *Расстояние* $\|a - b\|$ между двумя точками a и b из R^n обозначим через $\rho(a, b)$. *Шар* $\{x \in R^n : \rho(a, x) < r\}$ с центром в точке $a \in R^n$ и радиусом $r > 0$ обозначается через $\mathscr{D}(a, r)$, *замкнутый шар* $\{x \in R^n : \rho(a, x) \leqslant r\}$ — через $\overline{\mathscr{D}}(a, r)$, *сфера* $\{x \in R^n : \rho(a, x) = r\}$ — через $S(a, r)$.

Пусть $a \in R^n$, $b \in R^n$ и $a < b$. Множество $P(a, b) = \{x \in R^n : a < x < b\}$ называется *прямоугольником*, множество $P[a, b] = \{x \in R^n : a \leqq x \leqq b\}$ — *замкнутым прямоугольником* (замкнутый прямоугольник может быть определен и при $a \leqslant b$, $a \leqq b$). Если $a \in R^n$, $d \in R^n$, $d > 0$, то $P(a - d, a + d)$ — *прямоугольник с центром в точке a*; $P[a - d, a + d]$ — *замкнутый прямоугольник с центром в точке a*, если $d \geqq 0$.

2.3. Обычным образом определяется топология в R^n, согласованная с его метрической структурой.

Пространство R^n и пустое множество $\varnothing$ являются одновременно замкнутыми и открытыми множествами в R^n. Других подмножеств в R^n, отличных от R^n и $\varnothing$, являющихся одновременно открытыми и замкнутыми, не существует.

Шары и прямоугольники являются открытыми множествами. Замкнутые шары и замкнутые прямоугольники — замкнутые множества. Границей шара $\mathscr{D}(a, r)$ является сфера $S(a, r)$. Аффинные множества в R^n (в частности, подпространства пространства R^n) — замкнутые множества. Полупространства в R^n — открытые множества, замкнутые полупространства — замкнутые множества. Границей полупространства (замкнутого полупространства) является гиперплоскость, которая порождает это полупространство. Всякое множество, состоящее из конечного числа точек, замкнуто.

Множество всех внутренних точек множества M — его *открытое ядро* — обозначается через M^0 или $\overset{\circ}{M}$. Граница множества M обозначается через ∂M. *Замыкание* множества M обозначается через $\bar{M}$.

2.4. Отображение множества N всех натуральных чисел в R^n называется *последовательностью* в R^n. Последовательности будем обозначать через $\{x_k\}_{k=1}^{\infty}$, или x_1, x_2, ... Если $M \subset R^n$ и $x_k \in M$ при любом $k \in N$, то будем говорить, что $\{x_k\}_{k=1}^{\infty}$ — *последовательность элементов множества* M.

Если $\{x_k\}_{k=1}^{\infty}$ — последовательность в R^n (или в M) и K — бесконечное подмножество множества N, то ограничение отображения $\{x_k\}_{k=1}^{\infty}$ на множестве K называется *подпоследовательностью* (или *частичной последовательностью*) последовательности $\{x_k\}_{k=1}^{\infty}$ и обозначается через $\{x_k\}_{k \in K}$. Если последовательность $\{x_k\}_{k=1}^{\infty}$ имеет предел x_0 (сходится к x_0), то будем писать $x_k \to x_0$ при $k \to \infty$, или $\lim\limits_{k \to \infty} x_k = x_0$. Если x_0 есть предел подпоследовательности $\{x_k\}_{k \in K}$, то будем писать $x_k \to x_0$ при $k \to \infty$, $k \in K$.

2.5. Пусть A и B — непустые множества в R^n. *Расстоянием между множествами* A и B называется число $\rho(A, B) = \inf\{\rho(a, b), a \in A, b \in B\}$. Если существуют такие точки $\bar{a} \in A$, $\bar{b} \in B$, что $\rho(\bar{a}, \bar{b}) = \rho(A, B)$, то говорят, что расстояние между множествами A и B *осуществляется*. Некоторые достаточные условия осуществления расстояния будут приведены в п. 4.5.

2.6. Множество $M \subset R^n$ называется *компактным*, если из любой последовательности точек этого множества можно выделить подпоследовательность, сходящуюся к точке этого множества.

Необходимым и достаточным условием компактности множества $M \subset R^n$ является его ограниченность и замкнутость.

Пересечение любого множества компактных множеств есть множество компактное. Пересечение компактного и замкнутого множеств является множеством компактным. Объединение конечного числа компактных множеств — множество компактное. В частности, всякое конечное множество компактно.

2.7. Пусть $\{x_k\}_{k=1}^{\infty}$ — последовательность векторов из R^n и $a_1, a_2, \ldots, a_n$ — базис пространства R^n. Каждый вектор x_k представим в виде линейной комбинации векторов a_i, $i \in 1:n$:

$$x_k = \sum\nolimits_{i=1}^{n} \alpha_k^{(i)} a_i.$$

Рассмотрим n числовых последовательностей $\{\alpha_k^{(i)}\}_{k=1}^{\infty}$, $i \in 1:n$. Пусть $x_0 \in R^n$, $x_0 = \sum\nolimits_{i=1}^{n} \alpha_0^{(i)} a_i$.

Последовательность $\{x_k\}_{k=1}^{\infty}$ в том и только в том случае сходится к вектору x_0 при $k \to \infty$, если для любого $i \in 1:n$ числовая последовательность $\{\alpha_k^{(i)}\}_{k=1}^{\infty}$ сходится к $\alpha_0^{(i)}$ при $k \to \infty$. Иными словами, *сходимость в R^n эквивалентна покоординатной сходимости.*

2.8. Функция (отображение) $f: R^n \to R^m$ называется *непрерывной в точке* $x_0 \in R^n$, если полный прообраз любой окрестности точки $f(x_0) \in R^m$ содержит некоторую окрестность точки x_0. Функция f *непрерывна на множестве* $M \subset R^n$, если она непрерывна в каждой точке этого множества.

Пусть $f: R^n \to R^m$, $M \subset R^n$ и $x_0 \in M$. Функция f *непрерывна в точке $x_0 \in M$ относительно множества M* (или *относительно непрерывна в точке x_0 множества M*), если полный прообраз любой окрестности точки $f(x_0)$ содержит пересечение некоторой окрестности точки x_0 с множеством M. Функция f *относительно непрерывна на множестве M*, если f относительно непрерывна в каждой точке множества M.

Аналогично определяется непрерывность (относительная непрерывность) в точке $x_0 \in M$ и на множестве M функции, определенной не на всем пространстве R^n, а лишь на множестве $M' \supset M$.

Функция $f: R^n \to R^m$ в том и только в том случае непрерывна в точке x_0 (относительно непрерывна в точке $x_0 \in M$), если для любой последовательности $\{x_k\}_{k=1}^{\infty}$, сходящейся к точке x_0 (любой последовательности $\{x_k\}_{k=1}^{\infty}$ точек множества M, сходящейся к точке $x_0 \in M$), последовательность $\{f(x_k)\}_{k=1}^{\infty}$ сходится к $f(x_0)$.

Некоторые примеры. *Аффинное отображение* $f: R^n \to R^m$, $f(x) = Ax + b$, где $A = (a[i,j])$ — $m \times n$ матрица и $b \in R^m$, непрерывно на R^n. В частности, непрерывна на R^n всякая *аффинная функция* $f: R^n \to R$, $f(x) = ax + \beta$, $a \in R^n$, $\beta \in R$.

Если $f: R^n \to R^m$ — непрерывное на R^n отображение, то $\|f\|: R^n \to R$, $\|f\|(x) = \|f(x)\|$ (здесь $\|\cdot\|$ — норма в R^m), — непрерывная числовая функция на R^n. Функция $\rho: R^n \times R^n \to R$, $\rho(x, y) = \|x - y\|$, — непрерывная функция на $R^n \times R^n = R^{2n}$.

Если M — компактное множество в R^n и f непрерывна (относительно непрерывна) на M, то $f(M)$ — компактное множество в R^m. В частности, числовая функция f, непрерывная на компактном множестве $M \subset R^n$, принимает на этом множестве наибольшее и наименьшее значения, т. е. в M существуют такие точки $\bar{x}$ и $\underline{x}$, что $f(x) \leqslant f(\bar{x})$ и $f(x) \geqslant f(\underline{x})$ для любого $x \in M$ (теорема Вейерштрасса). Наибольшее значение функции f на множестве M обозначается $\max\limits_{x \in M} f(x)$, а наименьшее $\min\limits_{x \in M} f(x)$.

2.9. Пусть M — открытое множество в R^n и f — числовая функция на M. Функция f называется *дифференцируемой в точке* $x \in M$, если существует такой вектор $x^* \in R^n$, что для любого $\varepsilon > 0$ найдется такое $r > 0$, что шар $\mathscr{D}(x, r) \subset M$ и

$$|f(x+s) - f(x) - x^* s| \leqslant \varepsilon \|s\|$$

для любого $s \in \mathscr{D}(0, r)$. Если функция f дифференцируема в точке x, то вектор x^* определяется однозначно и называется *градиентом* (*вектором градиентом*) функции f в точке x. Градиент функции f в точке x обозначается $\nabla f(x)$, или $\operatorname{grad} f(x)$ (мы будем пользоваться первым обозначением). При этом

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x[1]}, \frac{\partial f(x)}{\partial x[2]}, \ldots, \frac{\partial f(x)}{\partial x[n]} \right),$$

где $\partial f(x)/\partial x[j]$ — частная производная функции f по j-й координате, $j \in 1:n$. Линейная функция $g_{f, x}: R^n \to R$, $g_{f, x}(s) = \nabla f(x) s$, где $\nabla f(x) s$ — скалярное произведение векторов $\nabla f(x)$ и s, называется *дифференциалом функции f в точке x*. Функция f называется *дифференцируемой на множестве M*, если она дифференцируема в каждой точке этого множества. Отображение $x \mapsto g_{f, x}$ множества M в множество $(R^n)^*$ линейных функций на R^n (или отображение $x \mapsto \nabla f(x)$ множества M в R^n) называется *производным отображением* функции f. Функция f называется *непрерывно дифференцируемой на множестве M*, если ее производное отображение непрерывно на этом множестве. Если $M = R^n$, то говорят, что f *дифференцируема* (*непрерывно дифференцируема*).

Необходимым и достаточным условием непрерывной дифференцируемости функции f на M является существование и непрерывность на M частных производных $\partial f/\partial x[j]$, $j \in 1:n$, функции f.

Функция $f: M \mapsto R$, где M — открытое множество в R^n, является *дважды дифференцируемой на M* (*дважды непрерывно дифференцируемой на M*), если f непрерывно дифференцируема на M и каждая из функций $\partial f/\partial x[j]$, $j \in 1:n$, дифференцируема (непрерывно дифференцируема на M). Функция f дважды непрерывно дифференцируема на M в том и только том случае, если существуют и непрерывны на M все вторые частные производные $\partial^2 f/\partial x[i]\partial x[j]$, $i \in 1:n$, $j \in 1:n$, функции f.

С каждой дважды непрерывно дифференцируемой на M функцией f связывается симметричная квадратная матрица $G(x) = (g(x)[i, j])$ — *матрица Гессе*, строками которой являются градиенты частных производных функции f, т. е. $g(x)[i, j] = \partial^2 f(x)/\partial x[i]\partial x[j]$.

2.10. Пусть M — открытое множество в R^n и $f: M \to R$. При фиксированных $x \in M$, $s \in R^n$ рассмотрим функцию $\varphi(\lambda) = f(x + \lambda s)$, определенную на множестве тех $\lambda \in R$, для которых $x + \lambda s \in M$. Если f непрерывно дифференцируема на M, то для любого $s \in R^n$ существует предел

$$f'(x, s) = \lim_{\lambda \downarrow 0} \frac{1}{\lambda}(f(x + \lambda s) - f(x))$$

(т. е. предел при $\lambda \to 0$ по множеству положительных чисел λ). Этот предел называется *производной функции f в точке x по направлению s*. Кроме того,

$$f'(x, s) = \varphi'(0) = \nabla f(x) s = \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x[j]} s[j]. \qquad (1)$$

Обратное неверно: функция f может иметь в точке x производную по любому направлению s и не быть ни дифференцируемой, ни даже непрерывной в этой точке.

Пусть теперь f дважды непрерывно дифференцируема на M, $x \in M$ и $s \in R^n$. Для таких $\lambda \in R$, что $x + \lambda s \in M$, положим $\psi(\lambda) = \nabla f(x + \lambda s) s = \varphi'(\lambda)$. Тогда

$$\psi'(\lambda) = s^{\mathrm{T}} G(x + \lambda s) s = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x + \lambda s)}{\partial x[i]\partial x[j]} s[i] s[j]. \qquad (2)$$

2.11. Точка $\bar{x} \in R^n$ называется *точкой локального максимума* (*локального минимума*) функции $f: R^n \to R$, если существует такая окрестность $\mathscr{D}(\bar{x}, r)$ точки $\bar{x}$, что $f(x) \leqslant f(\bar{x})$ $(f(x) \geqslant f(\bar{x}))$ для любого $x \in \mathscr{D}(\bar{x}, r)$. Точки локального максимума и локального минимума функции f называются *точками экстремума* функции f.

Если f — непрерывно дифференцируемая функция и $\bar{x}$ — точка экстремума функции f, то $\nabla f(\bar{x}) = \mathbf{0}$. Таким образом, точки экстремума непрерывно дифференцируемой функции f находятся в множестве решений системы уравнений

$$\partial f(x)/\partial x[j] = 0, \quad j \in 1:n. \qquad (3)$$

Векторы из R^n, являющиеся решениями системы (3), называются *стационарными точками* функции f. Стационарные точки функции f могут не быть ее точками экстремума.

Пусть f — дважды непрерывно дифференцируемая функция. Для того чтобы стационарная точка x_0 функции f была точкой ее

此为试读，需要完整PDF请访问：www.ertongbook.com