

ОПТИМИЗАЦИЯ
И ИССЛЕДОВАНИЕ
О П Е Р А Ц И Й



В. Ф. ДЕМЬЯНОВ
В. Н. МАЛОЗЕМОВ

Введение в минимакс

ОПТИМИЗАЦИЯ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

Редактор серии

Н. Н. МОИСЕЕВ

ИЗДАТЕЛЬСТВО «НАУКА»

ГЛАВНАЯ РЕДАКЦИЯ

ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ

МОСКВА 1972

В. Ф. ДЕМЬЯНОВ, В. Н. МАЛОЗЕМОВ

ВВЕДЕНИЕ В МИНИМАКС

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1972

Введение в минимакс. В. Ф. Демьянов, В. Н. Малоземов, Главная редакция физико-математической литературы изд-ва «Наука», 1972, стр. 368.

Минимакс (минимизация максимального отклонения) — принцип оптимального выбора параметров.

В первых двух главах книги рассматривается простейшая (и исторически первая) линейная минимаксная задача — построение алгебраического полинома наилучшего приближения.

В остальных четырех главах развивается общая теория нелинейных минимаксных задач. Отдельно рассматриваются дискретный и непрерывный случаи, отсутствие и наличие ограничений на параметры.

Основные вопросы: дифференцируемость функции максимума по направлениям, необходимые условия минимакса, достаточные условия локального минимакса, методы последовательных приближений для нахождения стационарных точек.

Большое количество примеров и рисунков иллюстрируют основные результаты теории.

Книга рассчитана на студентов и аспирантов физико-математических факультетов и широкий круг научных работников и инженеров, интересующихся экстремальными задачами.

Библ. 86 назв. Илл. 37.

Владимир Федорович Демьянов, Василий Николаевич Малоземов

ВВЕДЕНИЕ В МИНИМАКС

М., 1972 г., 368 стр. с илл.

Редакторы *И. К. Даугавет* и *М. М. Горячая*

Техн. редактор *И. Ш. Аксельрод*

Корректоры *О. А. Сигал, Л. Н. Боровина*

Сдано в набор 6/IX 1971 г. Подписано к печати 2/II 1972 г. Бумага 84×108¹/₃₂, тип. № 2. Физ. печ. л. 11,5. Условн. печ. л. 19,32. Уч.-изд. л. 18,26. Тираж 15 000. Т-01537. Цена книги 1 р. 35 к. Заказ № 1319.

Издательство «Наука»

Главная редакция физико-математической литературы.

Москва, В-71, Ленинский проспект, 15.

Ордена Трудового Красного Знамени Ленинградская типография № 2 имени Евгении Соколовой Главполиграфпрома Комитета по печати при Совете Министров СССР. Измайловский проспект, 29.

ОГЛАВЛЕНИЕ

Предисловие	7
Глава I. Дискретная задача наилучшего приближения функций алгебраическими полиномами	13
§ 1. Постановка задачи	13
§ 2. Чебышевская интерполяция	14
§ 3. Общая дискретная задача. Алгоритм Валле-Пуссена	26
§ 4. <i>R</i> -алгоритм	34
§ 5. Сведение к задаче линейного программирования	38
Глава II. Непрерывная задача наилучшего приближения функций алгебраическими полиномами	45
§ 1. Постановка задачи	45
§ 2. Теорема Чебышева. Полиномы Чебышева	46
§ 3. Предельные теоремы	52
§ 4. Метод последовательных чебышевских интерполяций Ремеза	55
§ 5. Метод сеток	60
§ 6*. О поведении коэффициентов полиномов наилучшего приближения	63
Глава III. Дискретная минимаксная задача	68
§ 1. Постановка задачи	68
§ 2. Свойства функции максимума	69
§ 3. Необходимые условия минимакса	77
§ 4. Достаточные условия локального минимакса. Некоторые оценки	86
§ 5. Метод покоординатного спуска. Метод наискорейшего спуска. Отрицательные примеры	94
§ 6. Первый метод последовательных приближений	106
§ 7. ε -стационарные точки. Второй метод последовательных приближений	117
§ 8. <i>D</i> -функция. Третий метод последовательных приближений	126
§ 9. Заключительные замечания	137
Глава IV. Дискретная минимаксная задача с ограничениями на параметры	144
§ 1. Постановка задачи	144
§ 2. Необходимые условия минимакса	145
§ 3. Геометрическая интерпретация необходимых условий	149
§ 4. Достаточные условия локального минимакса при наличии ограничений	156
§ 5. Некоторые оценки	160
§ 6. Метод последовательных приближений для нахождения стационарных точек	164

Глава V. Обобщенная задача нелинейного программирования	172
§ 1. Постановка задачи	172
§ 2. Свойства множеств, определяемых неравенствами	172
§ 3. Необходимые условия минимакса	183
§ 4*. Зависимость направления спуска от способа задания множества Ω	188
§ 5. Множители Лагранжа и теорема Куна — Таккера	194
§ 6. Первый метод последовательных приближений	199
§ 7. Нахождение (ϵ, μ) -квазистационарных точек. Второй метод последовательных приближений	210
§ 8. Метод наискорейшего спуска. Случай линейных ограничений	214
§ 9. Случай нелинейных ограничений. Способы корректировки направлений	220
§ 10. Метод штрафных функций	225
§ 11. Заключительные замечания	229
Глава VI. Непрерывная минимаксная задача	232
§ 1. Постановка задачи	232
§ 2. Основные теоремы	232
§ 3. Геометрическая интерпретация необходимого условия минимакса. Некоторые следствия	241
§ 4. О сходимости сеточного метода	253
§ 5. Частный случай теоремы о минимаксе	267
§ 6*. Разыскание седловых точек на многогранниках	273
§ 7. Наилучшее приближение функций нескольких аргументов обобщенными полиномами	283
§ 8. Наилучшее приближение функций, заданных на отрезке, алгебраическими полиномами	291
Приложение I. Алгебраическая интерполяция	297
§ 1. Разделенные разности	297
§ 2. Интерполяционные полиномы	300
Приложение II. Выпуклые множества и выпуклые функции	304
§ 1. Выпуклые оболочки. Теорема отделимости	304
§ 2. Выпуклые конусы	311
§ 3. Выпуклые функции	319
Приложение III. Непрерывные и непрерывно дифференцируемые функции	322
§ 1. Непрерывные функции	322
§ 2. Некоторые равенства и неравенства для непрерывных функций	323
§ 3. Непрерывно дифференцируемые функции	327
Приложение IV. Нахождение ближайшей к началу координат точки многогранника. Итеративные методы	335
Добавление. О задаче Мандельштама	357
Комментарии и библиография	361
Литература	364

ПРЕДИСЛОВИЕ

1. Минимакс — это один из принципов оптимального выбора параметров.

Сущность этого принципа легко объяснить на примере обработки эмпирических данных. Пусть задана таблица значений некоторой функции:

$$y_k = y(t_k), \quad k \in [0 : N]. \quad (1)$$

Требуется выбрать среди всех алгебраических полиномов

$P_n(A, t) = \sum_{i=0}^n a_i t^i$, где $A = (a_0, \dots, a_n)$, такой алгебраический полином $P_n(A^*, t)$, который достаточно хорошо аппроксимирует таблицу (1). Понятие «достаточно хорошо» требует уточнения. Будем искать полином $P_n(A^*, t)$, у которого максимальное отклонение

$$\max_{k \in [0 : N]} |y_k - P_n(A^*, t_k)|$$

минимально среди максимальных отклонений всех других алгебраических полиномов $P_n(A, t)$:

$$\max_{k \in [0 : N]} |y_k - P_n(A^*, t_k)| = \min_{\{A\}} \max_{k \in [0 : N]} |y_k - P_n(A, t_k)|.$$

Полином $P_n(A^*, t)$ называется полиномом наилучшего приближения таблицы (1) в чебышевском смысле. Иначе можно сказать, что A^* есть решение следующей минимаксной задачи:

$$\max_{k \in [0 : N]} |y_k - P_n(A, t_k)| \rightarrow \min_{\{A\}}.$$

Разумеется, вместо $P_n(A, t)$ можно взять любое другое параметрическое семейство функций $Q(X, t)$, где $X = (x_1, \dots, x_n)$, причем на вектор параметров X может быть наложено ограничение: $X \in \Omega$, где Ω — множество в E_n . Тогда, решая минимаксную задачу

$$\max_{k \in [0 : N]} |y_k - Q(X, t_k)| \rightarrow \min_{X \in \Omega},$$

найдем вектор $X^* \in \Omega$ такой, что $Q(X^*, t)$ наилучшим (в чебышевском смысле) образом аппроксимирует таблицу (1) среди всех других функций $Q(X, t)$, $X \in \Omega$.

Минимакс — не единственный принцип оптимального выбора параметров. В той же задаче об обработке эмпирических данных наилучшим полиномом можно считать полином $P_n(\tilde{A}, t)$, у которого минимальна сумма квадратов отклонений, т. е.

$$\sum_{k=0}^N (y_k - P_n(\tilde{A}, t_k))^2 = \min_{\{A\}} \sum_{k=0}^N (y_k - P_n(A, t_k))^2.$$

Вопрос о выборе принципа оптимизации зависит от характера задачи и требований, предъявляемых к ее решению. Оставляя этот вопрос в стороне, заметим лишь, что минимаксное решение в некоторых случаях обладает замечательным характеристическим свойством — наличием альтернанса. В задаче об аппроксимации это означает, что существуют $(n+2)$ узла $t_{i_0} < t_{i_1} < \dots < t_{i_{n+1}}$ (их количество на единицу больше, чем размерность вектора параметров A), в которых разность $y_{i_k} - P_n(A^*, t_{i_k})$ достигает максимального по абсолютной величине значения с последовательной переменой знака.

Приведем еще два примера минимаксных задач.

Задача Мандельштама. Пусть $X = (x_1, \dots, x_n)$ и

$$F(X, t) = \left| \sum_{k=1}^n \cos(kt + x_k) \right|.$$

Среди всех $X \in E_n$ требуется найти вектор X^* такой, что

$$\max_{t \in [0, 2\pi]} F(X^*, t) = \min_{\{X\}} \max_{t \in [0, 2\pi]} F(X, t).$$

Эта задача возникает в теории электрических цепей.

Матричные игры. Пусть задана вещественная $m \times n$ -матрица $A = \{a_{ij}\}$, $i \in [1:m]$, $j \in [1:n]$. Пусть, далее,

$$T_n = \left\{ X = (x_1, \dots, x_n) \mid x_i \geq 0, \sum_{i=1}^n x_i = 1 \right\},$$

$$G_m = \left\{ Y = (y_1, \dots, y_m) \mid y_j \geq 0, \sum_{j=1}^m y_j = 1 \right\}.$$

Требуется найти векторы $X^* \in T_n$ и $Y^* \in G_m$, для которых

$$\begin{aligned} \max_{Y \in G_m} (AX^*, Y) &= \min_{X \in T_n} \max_{Y \in G_m} (AX, Y), \\ \min_{X \in T_n} (AX, Y^*) &= \max_{Y \in G_m} \min_{X \in T_n} (AX, Y). \end{aligned}$$

В книге рассматривается минимаксная задача общего вида

$$\max_{Y \in G} F(X, Y) \rightarrow \min_{X \in \Omega} \quad (2)$$

где Ω — выпуклое замкнутое множество в E_n , а G — ограниченное замкнутое множество в E_m .

Если функция $F(X, Y)$ линейна по X при каждом фиксированном $Y \in G$, а множество Ω задается с помощью линейных равенств и неравенств, то исходная задача (2) называется линейной минимаксной задачей. В противном случае задача называется нелинейной. В зависимости от того, будет ли $\Omega \neq E_n$ или $\Omega = E_n$, говорят о минимаксной задаче с ограничениями или без ограничений. Переменная X называется параметром.

Имеются три основные идеи, которые могут быть использованы для решения минимаксных задач.

I. *Поиск экстремального базиса.* Пусть функция $F(X, Y)$ выпукла по X на Ω при каждом фиксированном $Y \in G$. Тогда на G можно найти r точек $Y_1, \dots, Y_r$, где $1 \leq r \leq n+1$, таких, что исходная минимаксная задача (2) равносильна следующей задаче:

$$\max_{Y \in G_r} F(X, Y) \rightarrow \min_{X \in \Omega} \quad (3)$$

где $G_r = \{Y_1, \dots, Y_r\}$. Множество G_r называется экстремальным базисом. Если экстремальный базис известен, то, решая обычно простую задачу (3), получим решение и исходной задачи (2).

II. *Минимизация функции максимума.* Положим

$$\varphi(X) = \max_{Y \in G} F(X, Y).$$

Тогда исходная задача (2) равносильна задаче минимизации функции $\varphi(X)$ на Ω .

III. *Нахождение седловой точки.* Точка $[X^*, Y^*]$ называется седловой точкой функции $F(X, Y)$ на множестве

$\Omega \times G$, если $F(X^*, Y) \leq F(X^*, Y^*) \leq F(X, Y^*)$ для всех $X \in \Omega$ и $Y \in G$. Допустив, что у функции $F(X, Y)$ существует седловая точка $[X^*, Y^*]$ на $\Omega \times G$, будем иметь

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) = F(X^*, Y^*) = \max_{Y \in G} \min_{X \in \Omega} F(X, Y).$$

Таким образом, в рассматриваемом случае минимаксная задача (2), а также двойственная к ней задача

$$\min_{X \in \Omega} F(X, Y) \rightarrow \max_{Y \in G},$$

сводятся к задаче об отыскании седловой точки.

Первая идея принадлежит П. Л. Чебышеву. В настоящее время она наиболее полно разработана для одномерных задач аппроксимации.

Вторая идея очень проста. В последние года она получила значительное развитие благодаря установлению при естественных предположениях дифференцируемости функции максимума по всем возможным направлениям.

Третью идею естественно связывать с именем Дж. фон Неймана. Она используется для решения игровых задач.

2. Переходим к изложению содержания книги по главам.

Первые две главы посвящены задаче наилучшего в чебышевском смысле приближения функций алгебраическими полиномами (глава I — дискретный случай, глава II — непрерывный случай). Изложение строится на основе чебышевской интерполяции. Это дает возможность весьма быстро получить все основные результаты чебышевской теории и попутно — методы для нахождения полиномов наилучшего приближения. Трудности, связанные с переходом от дискретного случая к непрерывному, весьма незначительны. Задача наилучшего равномерного приближения функций является простейшей (и исторически первой) линейной минимаксной задачей.

В третьей главе рассмотрена нелинейная минимаксная задача без ограничений на параметры (дискретный случай). В простейшей ситуации устанавливаются основные результаты минимаксной теории: дифференцируемость функции максимума по направлению, необходимые условия минимакса, достаточные условия локального

минимакса. Много внимания уделяется методам последовательных приближений для нахождения стационарных точек функции максимума $\varphi(X)$.

Те же вопросы, но при наличии выпуклых ограниченных рассмотрены в главах IV и V (дискретный случай).

В шестой главе осуществляется переход от дискретного случая к непрерывному.

Следует отметить, что наиболее принципиальной является третья глава.

В книге имеются четыре приложения, содержащие весь вспомогательный материал.

Ссылки на литературу в основном тексте отсутствуют. В конце книги помещен раздел «Комментарии и библиография», в котором даются исторические справки и литературные ссылки.

3. Авторы пользуются индуктивным методом изложения (вначале рассматриваются частные случаи минимаксных задач, а потом — более общие). На это есть две причины. Во-первых, книга предназначена для первоначального знакомства с предметом. Во-вторых, некоторые частные случаи имеют специфику, которая при изложении требует больше места, чем то, что следует из общих теорем. Так обстоит дело с полиномиальной аппроксимацией и с методами решения дискретных минимаксных задач. Поэтому мы предпочли вначале рассмотреть частные случаи вместе со всей их спецификой, а затем уже перейти к общей теории.

4. Книга состоит из трех практически независимых частей: главы I и II — полиномиальная аппроксимация; главы III — V — дискретные минимаксные задачи и нелинейное программирование; глава VI — основные результаты минимаксной теории в общей формулировке и их приложения.

Звездочкой отмечены те параграфы и пункты в параграфах, которые при первом чтении можно опустить.

Параграфы нумеруются независимо внутри каждой главы. При ссылках на формулы, леммы и теоремы внутри главы номер главы не указывается.

Поясним, наконец, смысл двух обозначений:

$\stackrel{df}{=}$ — равно по определению,

$[0 : N]$ — множество целых чисел от нуля до N включительно.

5*. Эта книга является результатом четырехлетней работы авторов. Отметим, что все главы и параграфы писались авторами совместно.

Сейчас, когда работа над книгой закончена, мы благодарим всех, кто оказал нам дружескую поддержку и помощь. Прежде всего мы благодарны Н. Н. Моисееву, который поддержал идею написания книги, проявил исключительное внимание к нашей работе и много сделал для того, чтобы книга быстро увидела свет. Ю. Б. Гермейер, И. К. Даугавет и Б. Н. Пшеничный внимательно прочитали рукопись и сделали ряд тонких и остроумных замечаний, которые были с благодарностью приняты и учтены. Весьма плодотворным было для нас многолетнее сотрудничество с Е. Ф. Войтоном, который предложил ряд практических нелинейных минимаксных задач, возникающих в радиотехнике, и во многом стимулировал работу авторов.

Мы имели счастливую возможность обсудить отдельные результаты с А. М. Вершиком, М. К. Гавуриным, Э. Г. Гильбертом, Д. М. Данскином, В. И. Zubовым, И. П. Мысовских, Л. В. Ньюстадом, Э. Полаком, И. В. Романовским, А. М. Рубиновым и В. А. Якубовичем. Всем им мы выражаем глубокую признательность.

В настоящее время у авторов имеется некоторый практический опыт решения минимаксных задач. В этом большую помощь нам оказали М. С. Альтмарк, В. М. Белых, Т. К. Виноградова, В. А. Даугавет, Б. Ф. Митчелл, А. Б. Певный, А. Л. Фрадков и В. И. Шүшков.

Мы искренне благодарны Э. В. Артемьевой, Л. К. Никитиной, В. И. Савченко и В. С. Филимоновой, помощь которых в оформлении рукописи трудно переоценить.

Основные результаты книги были изложены в лекциях, прочитанных в III Всесоюзной летней математической школе по методам оптимизации, проходившей в августе 1969 года в г. Тирасполе.

ГЛАВА I

ДИСКРЕТНАЯ ЗАДАЧА НАИЛУЧШЕГО ПРИБЛИЖЕНИЯ ФУНКЦИИ АЛГЕБРАИЧЕСКИМИ ПОЛИНОМАМИ

§ 1. Постановка задачи

Пусть задана таблица значений некоторой функции:

$$y_k = y(t_k), \quad k \in [0 : N].$$

t	t_0	t_1	t_2	$\dots$	t_N
y	y_0	y_1	y_2	$\dots$	y_N

Всюду в этой главе предполагается, что значения аргумента (узлы) упорядочены по возрастанию

$$t_0 < t_1 < t_2 < \dots < t_N.$$

Каждый алгебраический полином $P_n(A, t)$ степени не выше n ($n \leq N$)

$$P_n(A, t) = a_0 + a_1 t + \dots + a_n t^n, \quad A = (a_0, a_1, \dots, a_n).$$

имеет по отношению к исходной таблице естественную характеристику — максимальное уклонение:

$$\max_{k \in [0 : N]} |y_k - P_n(A, t_k)|.$$

Положим при фиксированном n

$$\rho = \inf_{\{A\}} \max_{k \in [0 : N]} |y_k - P_n(A, t_k)|.$$

Полином $P_n(A^*, t)$, для которого максимальное уклонение минимально

$$\max_{k \in [0 : N]} |y_k - P_n(A^*, t_k)| = \rho,$$

называется полиномом наилучшего приближения исходной таблицы.

Задача, которая будет рассматриваться в этой главе, заключается в фактическом построении полинома наилучшего приближения.

Решение этой задачи существенно зависит от соотношения между n и N .

Так, если $n = N$ (количество коэффициентов искомого полинома совпадает с количеством узлов в исходной таблице), то полиномом наилучшего приближения будет, очевидно, интерполяционный полином (см. Приложение I, § 2). В этом случае $\rho = 0$.

Первым нетривиальным (и, как будет показано, основным) случаем является случай $N = n + 1$. Он приводит к так называемой чебышевской интерполяции. С помощью чебышевской интерполяции можно строить полином наилучшего приближения и в самом общем случае $N > n + 1$.

§ 2. Чебышевская интерполяция

1. Сохраняя обозначения предыдущего параграфа и полагая $N = n + 1$, сформулируем и докажем следующую теорему.

Теорема 2.1. Пусть

$$\rho = \inf_{\{A\}} \max_{k \in [0 : n+1]} |y_k - P_n(A, t_k)|.$$

В этом случае полином наилучшего приближения существует и единствен. Для того чтобы полином $P_n(A^, t)$ был полиномом наилучшего приближения, необходимо и достаточно, чтобы при некотором h выполнялись соотношения*

$$(-1)^k h + P_n(A^*, t_k) = y_k, \quad k \in [0 : n + 1]. \quad (2.1)$$

Наконец, имеет место равенство $\rho = |h|$.

Доказательство будем проводить в следующей последовательности: достаточность — существование — необходимость — единственность.

Достаточность. Пусть для полинома $P_n(A^*, t)$ при некотором h справедливы соотношения (2.1). Докажем, что $P_n(A^*, t)$ — полином наилучшего приближения.

Положим

$$v(z_1, z_2, \dots, z_n) \stackrel{\text{df}}{=} \begin{vmatrix} 1 & z_1 & \dots & z_1^n \\ 1 & z_2 & \dots & z_2^n \\ \dots & \dots & \dots & \dots \\ 1 & z_n & \dots & z_n^n \end{vmatrix} = \prod_{n \geq k > i \geq 0} (z_k - z_i).$$

Если определитель Δ разложить по элементам первого столбца, то получим

$$\Delta = \sum_{v=0}^{n+1} v(t_0, t_1, \dots, t_{v-1}, t_{v+1}, \dots, t_{n+1}).$$

Поскольку узлы t_k упорядочены по возрастанию, то все слагаемые, входящие в последнюю сумму, положительны. Отсюда следует, что $\Delta > 0$.

Итак, система (2.1') всегда имеет решение и это решение единственно. Вместе с тем положительно решен вопрос о существовании полинома наилучшего приближения.

Необходимость. Прежде всего получим явное выражение для h . Для этого при всех $k \in [0 : n+1]$ положим

$$\alpha_k = \frac{\frac{(-1)^k}{(t_k - t_0) \dots (t_k - t_{n+1})}}{\sum_{v=0}^{n+1} \frac{(-1)^v}{(t_v - t_0) \dots (t_v - t_{n+1})}},$$

где

$$(t_k - t_0) \dots (t_k - t_{n+1}) = \prod_{i=0, i \neq k}^{n+1} (t_k - t_i).$$

Числа α_k обладают следующими свойствами:

а) $\alpha_k > 0$, $k \in [0 : n+1]$;

б) $\sum_{k=0}^{n+1} \alpha_k = 1$;

в) для любого полинома $P_n(A, t)$ степени не выше n

$$\sum_{k=0}^{n+1} (-1)^k \alpha_k P_n(A, t_k) = 0.$$

(Относительно этих свойств см. Приложение I, § 1.)

Теперь запишем очевидное в силу (2.1) равенство

$$\sum_{k=0}^{n+1} (-1)^k \alpha_k [(-1)^k h + P_n(A^*, t_k)] = \sum_{k=0}^{n+1} (-1)^k \alpha_k y_k.$$

Отсюда следует, что

$$h = \sum_{k=0}^{n+1} (-1)^k \alpha_k y_k. \quad (2.3)$$

Переходим к доказательству необходимости. Пусть $P_n(A^*, t)$ — некоторый полином наилучшего приближения.

Нужно доказать, что он удовлетворяет соотношениям (2.1), где h определяется из (2.3). Поскольку $\rho = |h|$, то

$$\max_{k \in [0: n+1]} |y_k - P_n(A^*, t_k)| = |h|. \quad (2.4)$$

Покажем, что на самом деле

$$y_k - P_n(A^*, t_k) = (-1)^k h, \quad k \in [0: n+1]. \quad (2.5)$$

Доказательство следует проводить отдельно для случаев $h=0$, $h>0$, $h<0$. Если $h=0$, то выполнение соотношений (2.5) очевидно в силу (2.4). Пусть $h<0$ (случай $h>0$ рассматривается аналогично), и пусть нарушено хотя бы одно из равенств (2.5). Тогда, во-первых, в силу (2.4)

$$(-1)^k [y_k - P_n(A^*, t_k)] \geq h, \quad k \in [0: n+1],$$

и, во-вторых, при некотором k_0

$$(-1)^{k_0} [y_{k_0} - P_n(A^*, t_{k_0})] > h.$$

Пользуясь равенством (2.3) и свойствами а) — в) чисел α_k , получим

$$\begin{aligned} h &= \sum_{k=0}^{n+1} (-1)^k \alpha_k y_k = \sum_{k=0}^{n+1} (-1)^k \alpha_k [y_k - P_n(A^*, t_k)] = \\ &= \sum_{k=0}^{n+1} \alpha_k \{(-1)^k [y_k - P_n(A^*, t_k)]\} > h \sum_{k=0}^{n+1} \alpha_k = h, \end{aligned}$$

что невозможно. Необходимость доказана.

Единственность полинома наилучшего приближения следует теперь из однозначной разрешимости системы (2.1) — (2.1'). Теорема доказана полностью.

2. Определение. Построение полинома, удовлетворяющего соотношениям (2.1), называется *чебышевской интерполяцией*; полином, удовлетворяющий соотношениям (2.1), называется *интерполяционным по Чебышеву полиномом*.

В предыдущем пункте по существу доказано, что построение полинома наилучшего приближения в случае $N = n + 1$ сводится к чебышевской интерполяции. При этом значение h уже найдено (2.3).

Нам понадобится другое выражение для h . Для того чтобы получить его, зададим в узлах t_k функцию χ следующим образом:

$$\chi_k = \chi(t_k) = (-1)^k, \quad k \in [0 : n + 1].$$

Тогда, вспоминая формулу для разделенных разностей (Приложение I, § 1) и определение чисел α_k , перепишем равенство (2.3) в виде

$$h = \frac{y[t_0, t_1, \dots, t_{n+1}]}{\chi[t_0, t_1, \dots, t_{n+1}]} \quad (2.6)$$

Это и есть требуемое выражение для h .

Переходя к построению полинома наилучшего приближения, введем вначале интерполяционный полином $L_{n+1}(t)$ степени не выше $n + 1$, который однозначно определяется соотношениями

$$L_{n+1}(t_k) = w_k; \quad w_k = y_k - h\chi_k, \quad k \in [0 : n + 1]. \quad (2.1'')$$

Представим $L_{n+1}(t)$ в форме Ньютона

$$L_{n+1}(t) = w_0 + \sum_{k=1}^{n+1} w[t_0, t_1, \dots, t_k] (t - t_0)(t - t_1) \dots (t - t_{k-1})$$

и обратим внимание на последний коэффициент:

$$w[t_0, t_1, \dots, t_{n+1}] = y[t_0, \dots, t_{n+1}] - h\chi[t_0, \dots, t_{n+1}].$$

В силу (2.6) $\omega[t_0, t_1, \dots, t_{n+1}] = 0$, так что на самом деле $L_{n+1}(t)$ является полиномом степени не выше n , который, кроме того, удовлетворяет соотношениям (2.1) — (2.1''). Значит, $L_{n+1}(t) \equiv P_n(A^*, t)$, и мы получаем окончательную формулу для полинома наилучшего приближения

$$P_n(A^*, t) = y_0 - h + \sum_{k=1}^n \{y[t_0, t_1, \dots, t_k] - h\chi[t_0, t_1, \dots, t_k]\} (t - t_0)(t - t_1) \dots (t - t_{k-1}). \quad (2.7)$$

Практически для пользования этой формулой следует:

I. Построить таблицы разделенных разностей функций y и χ .

II. Найти значение h по формуле (2.6).

III. Вычислить $y_0 - h$ и

$$y[t_0, t_1, \dots, t_k] - h\chi[t_0, t_1, \dots, t_k], \quad k \in [1 : n],$$

и подставить их в выражение (2.7) для полинома наилучшего приближения $P_n(A^*, t)$.

IV. Проверить выполнение соотношений

$$y_k - P_n(A^*, t_k) = (-1)^k h, \quad k \in [0 : n + 1].$$

Пример. Рассмотрим функцию $y(t) = |t|$. На отрезке $[-1, 1]$ зададим две системы узлов:

$$1) \quad t_0 = -1, \quad t_1 = -\frac{1}{2}, \quad t_2 = \frac{1}{2}, \quad t_3 = 1;$$

$$2) \quad t_0 = -1, \quad t_1 = -\frac{1}{2}, \quad t_2 = 0, \quad t_3 = \frac{1}{2}, \quad t_4 = 1.$$

Построим полиномы наилучшего приближения функции $|t|$ для этих систем узлов, считая, что $N = n + 1$; в первом случае это будет полином степени не выше 2, во втором — степени не выше 3.

Дальнейшие вычисления проведены согласно рекомендациям I — IV.

1)

t	y	Разделенные разности		
		I	II	III
-1	1	-1	$\frac{2}{3}$	0
$-\frac{1}{2}$	$\frac{1}{2}$			
$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{2}{3}$	
1	1	1		

В данном случае не надо составлять таблицу разностей для функции χ , ибо $y[t_0, t_1, t_2, t_3] = 0$ и, значит, $h = 0$. Сразу выписываем выражение для полинома наилучшего приближения:

$$P_2(t) = 1 - (t + 1) + \frac{2}{3}(t + 1)\left(t + \frac{1}{2}\right) = \frac{1}{3}(2t^2 + 1).$$

Нетрудно проверить справедливость равенств

$$|t_k| - P_2(t_k) = 0, \quad k \in [0 : 3].$$

2)

t	y	Разделенные разности			
		I	II	III	IV
-1	1	-1	0	$\frac{4}{3}$	$-\frac{4}{3}$
$-\frac{1}{2}$	$\frac{1}{2}$				
0	0	-1	2	$-\frac{4}{3}$	
$\frac{1}{2}$	$\frac{1}{2}$	1	0		
1	1	1			

t	χ	Разделенные разности			
		I	II	III	IV
-1	1	-4			
$-\frac{1}{2}$	-1	4	8	$-\frac{32}{3}$	
0	1	-4	-8	$\frac{32}{3}$	$\frac{32}{3}$
$\frac{1}{2}$	-1	4	8		
1	1				

$$h = \frac{-4/3}{32/3} = -\frac{1}{8};$$

$$P_3(t) = \frac{9}{8} - \frac{3}{2}(t+1) + (t+1)\left(t + \frac{1}{2}\right) + 0(t+1)\left(t + \frac{1}{2}\right)t = t^2 + \frac{1}{8}.$$

Нетрудно проверить справедливость равенств

$$|t_k| - P_3(t_k) = (-1)^{k+1} \cdot \frac{1}{8}, \quad k \in [0:4].$$

3*. В заключение этого параграфа рассмотрим один частный случай расположения узлов:

$$t_k = \cos \frac{(n+1-k)\pi}{n+1}, \quad k \in [0:n+1];$$

$$-1 = t_0 < t_1 < \dots < t_{n+1} = 1.$$

В этом случае удастся получить более простые, чем (2.6) и (2.7), формулы для h и для полинома наилучшего приближения $P_n(A^*, t)$.

Теорема 2.2. Пусть $t_k = \cos \frac{(n+1-k)\pi}{n+1}$, $y_k = y(t_k)$, $k \in [0 : n+1]$; тогда справедливы равенства

$$h = \frac{1}{2n+2} \left\{ y_0 + (-1)^{n+1} y_{n+1} + 2 \sum_{s=1}^n (-1)^s y_s \right\} \quad (2.8)$$

и

$$P_n(A^*, t) = \frac{1}{n+1} \sum_{s=0}^{n+1} y_s \tau_s(t), \quad (2.9)$$

где

$$\tau_s(t) = \begin{cases} D_n(x - \pi) = \frac{1}{2} + \sum_{v=1}^n (-1)^v T_v(t) & \text{при } s=0; \\ D_n(x + x_s) + D_n(x - x_s) = \\ = 1 + 2 \sum_{v=1}^n (-1)^v \cos \frac{vs\pi}{n+1} T_v(t) & \text{при } s \in [1 : n]; \\ D_n(x) = \frac{1}{2} + \sum_{v=1}^n T_v(t) & \text{при } s = n+1. \end{cases}$$

Здесь $D_n(x) = \frac{1}{2} + \cos x + \cos 2x + \dots + \cos nx$ — ядро Дирихле;

$$x = \arccos t; \quad x_s = \frac{(n+1-s)\pi}{n+1}, \quad s \in [1 : n];$$

$$T_v(t) = \cos v \arccos t, \quad v = 0, 1, 2, \dots$$

Доказательство. Достаточно проверить справедливость соотношений

$$(-1)^k h + P_n(A^*, t_k) = y_k, \quad k \in [0 : n+1].$$

Предварительно докажем две леммы.

Лемма 2.1. Функция $T_v(t) = \cos v \arccos t$, $-1 \leq t \leq 1$, является алгебраическим полиномом степени v со старшим коэффициентом 2^{v-1} , $v = 1, 2, \dots$

Доказательство. Прежде всего заметим, что

$$T_0(t) \equiv 1, \quad T_1(t) = t.$$

Далее, поскольку

$$\cos(v+2)x + \cos vx = 2 \cos(v+1)x \cos x,$$

то, полагая в этом тождестве $x = \arccos t$, получим для функций $T_v(t)$ рекуррентное соотношение

$$T_{v+2}(t) = 2tT_{v+1}(t) - T_v(t), \quad v = 0, 1, 2, \dots,$$

из которого очевидным образом следует утверждение леммы. В частности,

$$T_2(t) = 2t^2 - 1, \quad T_3(t) = 4t^3 - 3t \quad \text{и т. д.}$$

Полиномы $T_v(t)$ называются полиномами Чебышева.

Следствие. Функции $\tau_s(t)$, $s \in [0 : n+1]$, определенные выше, являются алгебраическими полиномами степени n .

Таким образом, и все выражение, стоящее в правой части равенства (2.9), является алгебраическим полиномом степени $\leq n$.

Лемма 2.2. Для ядра Дирихле $D_n(x)$ при $n = 0, 1, 2, \dots$ справедливы равенства

$$D_n\left(\frac{s\pi}{n+1}\right) = \begin{cases} n + \frac{1}{2}, & \text{если } s = 0; \\ \frac{1}{2}(-1)^{s+1}, & \text{если } s = \pm 1, \dots, \pm(2n+1). \end{cases}$$

Доказательство. По определению ядра Дирихле $D_n(0) = n + \frac{1}{2}$. Для того чтобы найти значения $D_n(x)$ в точках $x = \frac{s\pi}{n+1}$, $s = \pm 1, \dots, \pm(2n+1)$, заметим,

что

$$\frac{1}{2} + \cos x + \cos 2x + \dots + \cos nx = \frac{\sin \frac{2n+1}{2} x}{2 \sin \frac{x}{2}}.$$

Поэтому

$$\begin{aligned} D_n\left(\frac{s\pi}{n+1}\right) &= \frac{\sin \frac{2n+1}{2} \left(\frac{s\pi}{n+1}\right)}{2 \sin \frac{s\pi}{2n+2}} = \frac{\sin \left(s\pi - \frac{s\pi}{2n+2}\right)}{2 \sin \frac{s\pi}{2n+2}} = \\ &= \frac{(-1)^{s+1} \sin \frac{s\pi}{2n+2}}{2 \sin \frac{s\pi}{2n+2}} = \frac{1}{2} (-1)^{s+1}, \quad s = \pm 1, \dots, \pm(2n+1). \end{aligned}$$

Лемма доказана.

Возвращаясь к доказательству теоремы, выпишем выражение для $P_n(A^*, t)$ в узлах t_k , $k \in [0: n+1]$:

$$\begin{aligned} P_n(A^*, t_k) &= \frac{1}{n+1} \sum_{s=0}^{n+1} y_s \tau_s(t_k) = \\ &= \frac{1}{n+1} \left\{ y_0 D_n\left(\frac{k\pi}{n+1}\right) + y_{n+1} D_n\left(\frac{(n+1-k)\pi}{n+1}\right) + \right. \\ &\quad \left. + \sum_{s=1}^n y_s \left[D_n\left(\frac{(s+k)\pi}{n+1}\right) + D_n\left(\frac{(k-s)\pi}{n+1}\right) \right] \right\}. \end{aligned}$$

Пусть $k=0$. Тогда, учитывая, что $(-1)^v = (-1)^{-v}$ при любом целом v , получим

$$\begin{aligned} P_n(A^*, t_0) &= \frac{1}{n+1} \left\{ y_0 \left(n + \frac{1}{2}\right) + y_{n+1} \frac{1}{2} (-1)^n + \right. \\ &\quad \left. + \sum_{s=1}^n y_s \frac{1}{2} [(-1)^{s+1} + (-1)^{s+1}] \right\} = \\ &= y_0 - \frac{1}{2n+2} \left\{ y_0 + (-1)^{n+1} y_{n+1} + 2 \sum_{s=1}^n (-1)^s y_s \right\} = y_0 - h. \end{aligned}$$

Если $k = n + 1$, то

$$\begin{aligned} P_n(A^*, t_{n+1}) &= \\ &= \frac{1}{n+1} \left\{ y_0 \frac{1}{2} (-1)^n + y_{n+1} \left(n + \frac{1}{2} \right) + \sum_{s=1}^n (-1)^{n+s} y_s \right\} = \\ &= y_{n+1} - \frac{(-1)^{n+1}}{2n+2} \left\{ y_0 + (-1)^{n+1} y_{n+1} + 2 \sum_{s=1}^n (-1)^s y_s \right\} = \\ &= y_{n+1} - (-1)^{n+1} h. \end{aligned}$$

Пусть, наконец, $k \in [1 : n]$. В этом случае

$$\begin{aligned} P_n(A^*, t_k) &= \frac{1}{n+1} \left\{ y_0 \frac{1}{2} (-1)^{k+1} + y_{n+1} \frac{1}{2} (-1)^{n+k} + \right. \\ &\quad \left. + y_k n + \sum_{\substack{s=1 \\ s \neq k}}^n y_s (-1)^{s+k+1} \right\} = \\ &= y_k - \frac{(-1)^k}{2n+2} \left\{ y_0 + (-1)^{n+1} y_{n+1} + 2 \sum_{s=1}^n (-1)^s y_s \right\} = \\ &= y_k - (-1)^k h. \end{aligned}$$

Теорема доказана.

Следствие. Пусть $t_k = \cos \frac{(n+1-k)\pi}{n+1}$, $k \in [0 : n+1]$. Тогда для любого алгебраического полинома $Q_n(t)$ степени $\leq n$ справедливо равенство

$$Q_n(t_0) + (-1)^{n+1} Q_n(t_{n+1}) + 2 \sum_{s=1}^n (-1)^s Q_n(t_s) = 0. \quad (2.10)$$

Доказательство. Положим $y_k = Q_n(t_k)$, $k \in [0 : n+1]$. Осуществляя чебышевскую интерполяцию в этом случае, получим, согласно (2.6) и свойству 4 разделенных разностей (см. Приложение I, § 1):

$$h = \frac{y[t_0, t_1, \dots, t_{n+1}]}{\chi[t_0, t_1, \dots, t_{n+1}]} = 0.$$

Формула (2.10) следует теперь из (2.8).

§ 3. Общая дискретная задача. Алгоритм Валле-Пуссена

1. Рассмотрим общий случай дискретной задачи $N > n + 1$. Положим

$$y_k = y(t_k), \quad k \in [0 : N]; \quad t_0 < t_1 < \dots < t_N;$$

$$\rho = \inf_{\{A\}} \max_{k \in [0 : N]} |y_k - P_n(A, t_k)|.$$

Требуется построить полином наилучшего приближения $P_n(A^*, t)$:

$$\max_{k \in [0 : N]} |y_k - P_n(A^*, t_k)| = \rho.$$

В этом параграфе будет показано, что общая дискретная задача допускает точное решение. Оно может быть получено с помощью конечного числа чебышевских интерполяций.

2. Введем следующее

Определение. *Базисом σ* называется любая $(n + 2)$ -точечная подсистема узлов

$$\sigma = \{t_{i_0} < t_{i_1} < \dots < t_{i_{n+1}}\}. \quad (3.1)$$

Множество базисов $\{\sigma\}$ будем обозначать через Ξ .

На каждом базисе σ вида (3.1) можно осуществить чебышевскую интерполяцию, т. е. построить алгебраический полином $P_n(A(\sigma), t)$, удовлетворяющий соотношениям

$$(-1)^k h(\sigma) + P_n(A(\sigma), t_{i_k}) = y_{i_k}, \quad k \in [0 : n + 1],$$

где

$$h(\sigma) = \sum_{k=0}^{n+1} (-1)^k \alpha_{i_k}(\sigma) y_{i_k} \quad (3.2)$$

и

$$\alpha_{i_k}(\sigma) = \frac{(-1)^k}{\sum_{v=0}^{n+1} \frac{(t_{i_k} - t_{i_0}) \dots (t_{i_k} - t_{i_{n+1}})}{(-1)^v (t_{i_v} - t_{i_0}) \dots (t_{i_v} - t_{i_{n+1}})}},$$

$$k \in [0 : n + 1].$$

Пусть $\rho(\sigma)$ есть наилучшее приближение на базисе σ :

$$\rho(\sigma) = \inf_{\{A\}} \max_{k \in [0:n+1]} |y_{t_k} - P_n(A, t_k)|.$$

Напомним, что тогда справедливы следующие равенства:

$$\rho(\sigma) = \max_{k \in [0:n+1]} |y_{t_k} - P_n(A(\sigma), t_k)| = |h(\sigma)|.$$

Положим, наконец,

$$\varphi(A(\sigma)) = \max_{k \in [0:N]} |y_k - P_n(A(\sigma), t_k)|,$$

так что $\varphi(A(\sigma))$ есть максимальное отклонение полинома $P_n(A(\sigma), t)$ уже на всей системе узлов.

Очевидно, что для любого базиса $\sigma \in \Xi$

$$\varphi(A(\sigma)) \geq \rho(\sigma).$$

Лемма 3.1. Если для некоторого базиса σ^*

$$\sigma^* = \{t_{i_0}^* < t_{i_1}^* < \dots < t_{i_{n+1}}^*\}$$

оказалось, что $\varphi(A(\sigma^*)) = \rho(\sigma^*)$, то

$$\varphi(A(\sigma^*)) \stackrel{\text{def}}{=} \max_{k \in [0:N]} |y_k - P_n(A(\sigma^*), t_k)| = \rho,$$

т. е. полином $P_n(A(\sigma^*), t)$ является искомым полиномом наилучшего приближения.

Доказательство. В силу определения $\rho(\sigma^*)$ имеем для любого полинома $P_n(A, t)$

$$\max_{k \in [0:n+1]} |y_{t_k}^* - P_n(A, t_{i_k}^*)| \geq \rho(\sigma^*).$$

Тем более

$$\max_{k \in [0:N]} |y_k - P_n(A, t_k)| \geq \rho(\sigma^*). \quad (3.3)$$

Поскольку по условию леммы

$$\rho(\sigma^*) = \varphi(A(\sigma^*)) = \max_{k \in [0:N]} |y_k - P_n(A(\sigma^*), t_k)|, \quad (3.4)$$

то в силу (3.3) и (3.4) получим для любого полинома $P_n(A, t)$:

$$\max_{k \in [0:N]} |y_k - P_n(A, t_k)| \geq \max_{k \in [0:N]} |y_k - P_n(A(\sigma^*), t_k)|.$$

Последнее неравенство по определению и означает, что $P_n(A(\sigma^*), t)$ является полиномом наилучшего приближения на всей системе узлов.

Лемма доказана.

3. Итак, если для некоторого базиса σ оказалось, что $\varphi(A(\sigma)) = \rho(\sigma)$, то $P_n(A(\sigma), t)$ является полиномом наилучшего приближения в общей дискретной задаче.

Теперь обратимся к тем базисам $\sigma \in \Xi$ вида (3.1), для которых

$$\varphi(A(\sigma)) > \rho(\sigma). \quad (3.5)$$

Опишем некоторое стандартное преобразование S , которое переводит всякий базис σ , удовлетворяющий неравенству (3.5), в новый базис $\sigma_1 = S\sigma$

$$\sigma_1 = \{t_{i_0}^{(1)} < t_{i_1}^{(1)} < \dots < t_{i_{n+1}}^{(1)}\},$$

обладающий следующим свойством:

$$\rho(\sigma_1) > \rho(\sigma). \quad (3.6)$$

Преобразования такого рода в дальнейшем будут играть основную роль.

Итак, переходим к описанию преобразования S . Пусть

$$\Delta_\sigma(t_k) = y_k - P_n(A(\sigma), t_k), \quad k \in [0 : N],$$

и t_{k_0} — узел, в котором

$$|\Delta_\sigma(t_{k_0})| = \varphi(A(\sigma))$$

(если таких узлов несколько, берем любой из них).

Возможны три случая:

$$1) t_{i_0} < t_{k_0} < t_{i_{n+1}};$$

$$2) t_{k_0} < t_{i_0};$$

$$3) t_{k_0} > t_{i_{n+1}}.$$

В каждом из этих случаев преобразование S определяется по-своему.

Первый случай. Характерная ситуация изображена на рис. 1.

Пусть целое число ν таково, что

$$t_{i_\nu} < t_{k_0} < t_{i_{\nu+1}}.$$

Тогда сразу полагаем

$$t_{i_l}^{(1)} = t_{i_l}, \quad l \neq v, v+1; \quad l \in [0 : n+1].$$

Если $\text{sign } \Delta_\sigma(t_{k_0}) = \text{sign } \Delta_\sigma(t_{i_v})$, то

$$t_{i_v}^{(1)} = t_{k_0}, \quad t_{i_{v+1}}^{(1)} = t_{i_{v+1}}.$$

Если же последнее условие не выполняется, то

$$t_{i_v}^{(1)} = t_{i_v}, \quad t_{i_{v+1}}^{(1)} = t_{k_0}.$$

Таким образом, базис $\sigma_1 = S\sigma$ в рассматриваемом случае полностью определен. В ситуации, изображенной

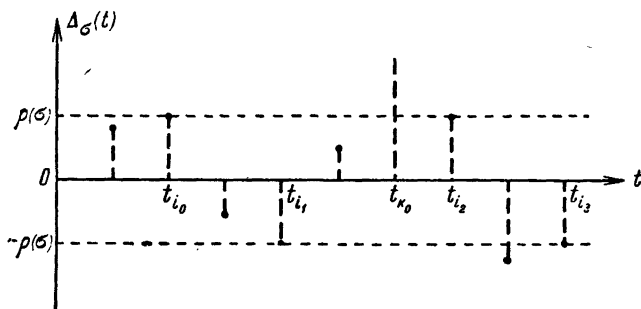


Рис. 1.

на рис. 1, базис σ_1 получается из базиса σ заменой узла t_{i_2} на t_{k_0} .

Второй случай. Если $\text{sign } \Delta_\sigma(t_{k_0}) = \text{sign } \Delta_\sigma(t_{i_0})$, то полагаем

$$t_{i_0}^{(1)} = t_{k_0}; \quad t_{i_l}^{(1)} = t_{i_l}, \quad l \in [1 : n+1].$$

Если же последнее условие не выполняется, то

$$t_{i_0}^{(1)} = t_{k_0}; \quad t_{i_l}^{(1)} = t_{i_{l-1}}, \quad l \in [1 : n+1]$$

(из базиса σ выбрасывается узел $t_{i_{n+1}}$).

Третий случай. Если $\text{sign } \Delta_\sigma(t_{k_0}) = \text{sign } \Delta_\sigma(t_{i_{n+1}})$, то

$$t_{i_l}^{(1)} = t_{i_l}, \quad l \in [0 : n]; \quad t_{i_{n+1}}^{(1)} = t_{k_0}.$$

В противном случае

$$t_{l_l}^{(1)} = t_{l_{l+1}}, \quad l \in [0 : n]; \quad t_{l_{n+1}}^{(1)} = t_{k_0}$$

(из базиса σ выбрасывается узел t_{l_0}).

Теперь преобразование S описано полностью.

Из построения и из неравенства (3.5) вытекают следующие свойства базиса $\sigma_1 = S\sigma$ в случае, если $\rho(\sigma) > 0$:

$$\text{I. } \operatorname{sign} \Delta_{\sigma_1}(t_{i_k}^{(1)}) = -\operatorname{sign} \Delta_{\sigma}(t_{i_{k+1}}^{(1)}), \quad k \in [0 : n].$$

II. Если обозначить через $t_{i_s}^{(1)}$ узел базиса σ_1 , соответствующий t_{k_0} , то

$$|\Delta_{\sigma}(t_{i_k}^{(1)})| = \rho(\sigma) \quad \text{для всех } k \neq s,$$

$$|\Delta_{\sigma}(t_{i_s}^{(1)})| > \rho(\sigma).$$

Теперь мы имеем возможность доказать основное свойство базиса σ_1 :

$$\rho(\sigma_1) > \rho(\sigma).$$

Доказательство. В силу (3.2) и свойства 4 разделенных разностей (см. Приложение I, § 1)

$$\begin{aligned} \rho(\sigma_1) &= \left| \sum_{k=0}^{n+1} (-1)^k \alpha_{i_k}(\sigma_1) y_{i_k}^{(1)} \right| = \\ &= \left| \sum_{k=0}^{n+1} \alpha_{i_k}(\sigma_1) \{(-1)^k [y_{i_k}^{(1)} - P_n(A(\sigma), t_{i_k}^{(1)})]\} \right| = \\ &= \left| \sum_{k=0}^{n+1} \alpha_{i_k}(\sigma_1) \{(-1)^k \Delta_{\sigma}(t_{i_k}^{(1)})\} \right|. \end{aligned}$$

Пусть вначале $\rho(\sigma) > 0$.

Так как $\alpha_{i_k}(\sigma_1) > 0$, $k \in [0 : n+1]$, и $\sum_{k=0}^{n+1} \alpha_{i_k}(\sigma_1) = 1$, то в силу I и II получим

$$\rho(\sigma_1) = \sum_{k=0}^{n+1} \alpha_{i_k}(\sigma_1) |\Delta_{\sigma}(t_{i_k}^{(1)})| > \rho(\sigma) \sum_{k=0}^{n+1} \alpha_{i_k}(\sigma_1) = \rho(\sigma).$$

Если же $\rho(\sigma) = 0$, то

$$\rho(\sigma_1) = \alpha_{i_s}(\sigma_1) |\Delta_{\sigma}(t_{i_s}^{(1)})| > 0.$$

Итак, во всех случаях $\rho(\sigma_1) > \rho(\sigma)$.

Утверждение доказано.

Полученный результат подытожим в следующей лемме.

Лемма 3.2. *Если для базиса $\sigma \in \Xi$ выполняется неравенство $\varphi(A(\sigma)) > \rho(\sigma)$, то найдется новый базис $\sigma_1 = S\sigma$, для которого $\rho(\sigma_1) > \rho(\sigma)$.*

4. Положим

$$\bar{\rho} = \max_{\sigma \in \Xi} \rho(\sigma).$$

Определение. Базис σ^* , для которого

$$\rho(\sigma^*) = \bar{\rho},$$

называется *экстремальным базисом*.

Поскольку множество базисов Ξ конечно, то, очевидно, экстремальный базис существует, хотя в общем случае он может быть не единственным.

Теорема 3.1. *В общей дискретной задаче полином наилучшего приближения существует и единствен. Для того чтобы полином $P_n(A^*, t)$ был полиномом наилучшего приближения, необходимо и достаточно, чтобы он осуществлял чебышевскую интерполяцию на некотором экстремальном базисе σ^* , т. е. чтобы*

$$P_n(A^*, t) \equiv P_n(A(\sigma^*), t).$$

Наконец, $\rho = \bar{\rho}$, так что

$$\inf_{\{A\}} \max_{k \in [0:N]} |y_k - P_n(A, t_k)| = \max_{\sigma \in \Xi} \rho(\sigma). \quad (3.7)$$

Доказательство. **Достаточность.** Пусть σ^* — некоторый экстремальный базис. Это значит, что для любого $\sigma \in \Xi$

$$\rho(\sigma^*) \geq \rho(\sigma). \quad (3.8)$$

Нужно доказать, что полином $P_n(A(\sigma^*), t)$ является полиномом наилучшего приближения, т. е. что

$$\varphi(A(\sigma^*)) \stackrel{df}{=} \max_{k \in [0:N]} |y_k - P_n(A(\sigma^*), t_k)| = \rho.$$

В силу леммы 3.1 для этого достаточно проверить справедливость равенства $\varphi(A(\sigma^*)) = \rho(\sigma^*)$.

Предположим, что последнее равенство не выполняется. Тогда $\varphi(A(\sigma^*)) > \rho(\sigma^*)$. В силу леммы 3.2 найдется новый базис $\sigma_1 = S\sigma^*$, для которого

$$\rho(\sigma_1) > \rho(\sigma^*),$$

что противоречит (3.8). Достаточность доказана.

Вместе с тем доказана справедливость равенства (3.7), поскольку

$$\rho = \varphi(A(\sigma^*)) = \rho(\sigma^*) = \bar{\rho}. \quad (3.9)$$

Существование полинома наилучшего приближения следует из существования экстремального базиса.

Необходимость. Зафиксируем экстремальный базис σ^* :

$$\sigma^* = \{t_{i_0}^* < t_{i_1}^* < \dots < t_{i_{n+1}}^*\},$$

и пусть $P_n(A^*, t)$ — некоторый полином наилучшего приближения. Поскольку

$$\begin{aligned} \rho(\sigma^*) &\leq \max_{k \in [0:n+1]} |y_{i_k}^* - P_n(A^*, t_{i_k}^*)| \leq \\ &\leq \max_{i \in [0:N]} |y_i - P_n(A^*, t_i)| = \rho = \bar{\rho} = \rho(\sigma^*), \end{aligned}$$

то

$$\max_{k \in [0:n+1]} |y_{i_k}^* - P_n(A^*, t_{i_k}^*)| = \rho(\sigma^*).$$

Значит, $P_n(A^*, t)$ есть полином наилучшего приближения на базисе σ^* . Учитывая, что полином наилучшего приближения на базисе σ^* единствен, заключаем, что

$$P_n(A^*, t) \equiv P_n(A(\sigma^*), t).$$

Отсюда следует и необходимость и единственность.

Теорема доказана.

Следствие. Для того чтобы базис

$$\sigma = \{t_{i_0} < t_{i_1} < \dots < t_{i_{n+1}}\}$$

был экстремальным, необходимо и достаточно, чтобы выполнялось равенство

$$\varphi(A(\sigma)) = \rho(\sigma),$$

или, что то же самое,

$$\max_{i \in [0: N]} |\Delta_{\sigma}(t_i)| = \max_{k \in [0: n+1]} |\Delta_{\sigma}(t_{i_k})|.$$

5. Основная теорема дает возможность описать некоторый метод для построения полинома наилучшего приближения в общей дискретной задаче. Метод заключается в том, что нужно перебрать все базисы и остановиться на том базисе σ^* , для которого величина $\rho(\sigma^*)$ максимальна. Интерполяционный по Чебышеву полином $P_n(A(\sigma^*), t)$ и будет искомым полиномом наилучшего приближения.

Трудность состоит в том, чтобы организовать перебор базисов наиболее разумным образом. Это особенно важно, если число N велико.

Поскольку мы ищем базис с максимальной величиной наилучшего приближения, то следует подчинить перебор следующему естественному требованию: переход от базиса σ_k к базису σ_{k+1} должен осуществляться таким образом, чтобы выполнялось неравенство

$$\rho(\sigma_{k+1}) > \rho(\sigma_k). \quad (3.10)$$

Выполнение последнего неравенства на каждом шаге гарантирует вдобавок то, что один базис при переборе не может встретиться дважды.

Опишем алгоритм Валле-Пуссена для построения полинома наилучшего приближения, который реализует идею поиска экстремального базиса. В этом алгоритме используется преобразование базисов S , которое уже встречалось в п. 3.

Выбираем произвольным образом начальный базис σ_1 и строим полином $P_n(A(\sigma_1), t)$. Если

$$\varphi(A(\sigma_1)) = \rho(\sigma_1),$$

то в силу леммы 3.1 полином $P_n(A(\sigma_1), t)$ есть полином наилучшего приближения. Если же

$$\varphi(A(\sigma_1)) > \rho(\sigma_1),$$

то переходим к базису $\sigma_2 = S\sigma_1$. В этом случае, согласно лемме 3.2,

$$\rho(\sigma_2) > \rho(\sigma_1).$$

Пусть уже построен базис σ_k . Находим полином $P_n(A(\sigma_k), t)$. Если

$$\varphi(A(\sigma_k)) = \rho(\sigma_k),$$

то $P_n(A(\sigma_k), t)$ является искомым полиномом наилучшего приближения. Если же

$$\varphi(A(\sigma_k)) > \rho(\sigma_k),$$

то переходим к базису $\sigma_{k+1} = S\sigma_k$.

Описанный процесс за конечное число шагов приводит к экстремальному базису и к полиному наилучшего приближения. Конечность алгоритма следует из того, что конечно число базисов и при переборе, в силу (3.10), один базис не может встретиться дважды.

§ 4. R-алгоритм

1. В алгоритме Валле-Пуссена построения алгебраического полинома наилучшего приближения используется S -преобразование базисов. Это простейшее преобразование приводит по существу к смене одной точки старого базиса σ . При этом в новый базис $\sigma_1 = S\sigma$ привлекается узел, в котором $|\Delta_\sigma(t_k)|$ достигает наибольшего значения.

При использовании S -преобразования приходится на каждом шаге вычислять все $\Delta_\sigma(t_k)$, $k \in [0 : N]$. Алгоритм Валле-Пуссена имеет тот недостаток, что информация от вычисления всех отклонений $\Delta_\sigma(t_k)$ используется минимальным образом, лишь для нахождения наибольшего по абсолютной величине отклонения.

В этом параграфе будет описано новое преобразование базисов R , которое приводит, вообще говоря, к смене нескольких (и, может быть, всех) узлов старого базиса и к большему возрастанию по сравнению с $\rho(S\sigma)$ величины $\rho(R\sigma)$.

2. R -преобразование определяется только для тех базисов σ , у которых $\rho(\sigma) > 0$.

Положим

$$\sigma = \{t_{i_0} < t_{i_1} < \dots < t_{i_{n+1}}\};$$

$$\Delta_\sigma(t_k) = y_k - P_n(A(\sigma), t_k), \quad k \in [0 : N].$$

Будем строить базис $\sigma_1 = R\sigma$:

$$\sigma_1 = \{t_{i_0}^{(1)} < t_{i_1}^{(1)} < \dots < t_{i_{n+1}}^{(1)}\},$$

заботясь о том, чтобы выполнялись следующие условия:

$$\text{I. } \operatorname{sign} \Delta_\sigma(t_{i_k}^{(1)}) = -\operatorname{sign} \Delta_\sigma(t_{i_{k+1}}^{(1)}), \quad k \in [0 : n].$$

$$\text{II. } |\Delta_\sigma(t_{i_k}^{(1)})| \geq \rho(\sigma), \quad k \in [0 : n + 1].$$

III. В базис σ_1 включается хотя бы один узел, в котором достигается $\max_{k \in [0 : N]} |\Delta_\sigma(t_k)| \stackrel{\text{def}}{=} \varphi(A(\sigma))$.

Так же, как и для S -преобразования, выполнение этих условий гарантирует неравенство

$$\rho(\sigma_1) > \rho(\sigma) \quad (4.1)$$

в случае, если $\varphi(A(\sigma)) > \rho(\sigma)$.

Преобразование базисов R осуществляется в $n + 1$ шаг.

Первый шаг. Рассмотрим узлы, содержащиеся в отрезке $[t_0, t_{i_1}]$. Выберем из них два узла t_{v_1} и t_{w_1} так, чтобы $t_{v_1} < t_{w_1}$ и в одном из этих узлов достигалось бы наибольшее, а в другом наименьшее значение $\Delta_\sigma(t_k)$, $k \in [0 : i_1]$.

Положим

$$t_{i_0}^{(1)} = t_{v_1}, \quad t_{i_1}^{(1)} = t_{w_1}.$$

Второй шаг. Рассмотрим узлы из отрезка $[t_{i_1}, t_{i_2}]$. Выберем два узла t_{v_2} и t_{w_2} из условий: $t_{v_2} < t_{w_2}$; в одном из них достигается наибольшее, а в другом наименьшее значение $\Delta_\sigma(t_k)$, $k \in [i_1 : i_2]$.

Пусть

$$\operatorname{sign} \Delta_\sigma(t_{i_1}^{(1)}) = \operatorname{sign} \Delta_\sigma(t_{v_2}).$$

Тогда полагаем $t_{i_2}^{(1)} = t_{w_2}$. Если к тому же $|\Delta_\sigma(t_{i_1}^{(1)})| < |\Delta_\sigma(t_{v_2})|$, то переопределяем $t_{i_1}^{(1)}$: $t_{i_1}^{(1)} \stackrel{\text{def}}{=} t_{v_2}$. Запоминаем $z_2 = 0$.

Пусть $\operatorname{sign} \Delta_\sigma(t_{i_1}^{(1)}) = -\operatorname{sign} \Delta_\sigma(t_{v_2})$.

Тогда полагаем $t_{i_2}^{(1)} = t_{v_2}$. Запоминаем $t_{z_2} = t_{w_2}$ и $z_2 = \Delta_\sigma(t_{z_2})$.

($v + 1$)-й шаг. Пусть уже построены узлы

$$t_{i_0}^{(1)} < t_{i_1}^{(1)} < \dots < t_{i_v}^{(1)}.$$

При этом от последнего шага осталось значение z_v . Кроме того, если $z_v \neq 0$, то имеется еще узел t_{zv} .

Рассмотрим узлы на отрезке $[t_v, t_{v+1}]$. Выберем два узла $t_{v, v+1}$ и $t_{w, v+1}$ из условий: $t_{v, v+1} < t_{w, v+1}$; в одном из них достигается наибольшее, а в другом наименьшее значение $\Delta_\sigma(t_k)$, $k \in [i_v : i_{v+1}]$.

Пусть $\text{sign } \Delta_\sigma(t_{i_v}^{(1)}) = \text{sign } \Delta_\sigma(t_{v, v+1})$.

Тогда, во-первых, запоминаем $z_{v+1} = 0$. Если $|\Delta_\sigma(t_{i_v}^{(1)})| \geq |\Delta_\sigma(t_{v, v+1})|$, то полагаем

$$t_{i_{v+1}}^{(1)} = \begin{cases} t_{w, v+1}, & \text{если } |\Delta_\sigma(t_{w, v+1})| \geq |z_v|; \\ t_{zv} & \text{в противном случае.} \end{cases}$$

Если $|\Delta_\sigma(t_{i_v}^{(1)})| < |\Delta_\sigma(t_{v, v+1})|$, то полагаем $t_{i_{v+1}}^{(1)} = t_{w, v+1}$.

Далее переопределяем $t_{i_v}^{(1)} : t_{i_v}^{(1)} \stackrel{df}{=} t_{v, v+1}$. Наконец, если $|z_v| > |\Delta_\sigma(t_{i_{v-1}}^{(1)})|$, то переопределяем еще и $t_{i_{v-1}}^{(1)} : t_{i_{v-1}}^{(1)} \stackrel{df}{=} t_{zv}$.

Пусть теперь $\text{sign } \Delta_\sigma(t_{i_v}^{(1)}) = -\text{sign } \Delta_\sigma(t_{v, v+1})$. Тогда полагаем

$$t_{i_{v+1}}^{(1)} = \begin{cases} t_{v, v+1}, & \text{если } |\Delta_\sigma(t_{v, v+1})| \geq |z_v|, \\ t_{zv} & \text{в противном случае.} \end{cases}$$

Запоминаем $t_{z, v+1} = t_{w, v+1}$, $z_{v+1} = \Delta_\sigma(t_{z, v+1})$.

Сделаем два замечания, после которых преобразование базисов R будет полностью описано.

Во-первых, на последнем $(n+1)$ -м шаге надо рассматривать узлы не из отрезка $[t_n, t_{n+1}]$, а из отрезка $[t_n, t_N]$.

Во-вторых, на последнем шаге появится z_{n+1} , и если $z_{n+1} \neq 0$, еще $t_{z, n+1}$. Если окажется, что $|z_{n+1}| > |\Delta_\sigma(t_{i_0}^{(1)})|$, то производим сдвиг: $t_{i_k}^{(1)}$ присваивается значение $t_{i_{k+1}}^{(1)}$, $k \in [0 : n]$; $t_{i_{n+1}}^{(1)} \stackrel{df}{=} t_{z, n+1}$. В противном случае никаких изменений не производим.

Теперь R -преобразование базисов описано полностью. Надо доказать, что выполняются условия I, II и III.

Заметим, что на каждом шаге среди рассматриваемых узлов были два соседних узла старого базиса. Поэтому при $k \in [1 : n+1]$

$$\text{sign } \Delta_\sigma(t_{v_k}) = -\text{sign } \Delta_\sigma(t_{w_k})$$

и

$$|\Delta_{\sigma}(t_{vk})| \geq \rho(\sigma), \quad |\Delta_{\sigma}(t_{wk})| \geq \rho(\sigma).$$

Отсюда и из построения следует выполнение условий I и II. Нетрудно также показать, что в новый базис войдет один из узлов, в котором достигается $\max_{k \in [0: N]} \Delta_{\sigma}(t_k)$.

Таким образом, R -преобразование удовлетворяет всем поставленным условиям.

3. Опишем теперь R -алгоритм построения алгебраического полинома наилучшего приближения, использующий R -преобразование базисов.

Берем произвольным образом начальный базис σ_1 , $\rho(\sigma_1) > 0$, строим полином $P_n(A(\sigma_1), t)$ и базис $\sigma_2 = R\sigma_1$. Если оказалось, что

$$\rho(\sigma_2) = \rho(\sigma_1), \quad (4.2)$$

то процесс закончен и $P_n(A(\sigma_1), t)$ — полином наилучшего приближения. Действительно, в силу (4.1) равенство (4.2) возможно только тогда, когда $\varphi(A(\sigma_1)) = \rho(\sigma_1)$, а этого достаточно, чтобы $P_n(A(\sigma_1), t)$ был полиномом наилучшего приближения.

Если $\rho(\sigma_2) > \rho(\sigma_1)$, то строим полином $P_n(A(\sigma_2), t)$ и базис $\sigma_3 = R\sigma_2$.

Пусть уже построен полином $P_n(A(\sigma_v), t)$ и базис $\sigma_{v+1} = R\sigma_v$. Если оказалось, что $\rho(\sigma_{v+1}) = \rho(\sigma_v)$, то процесс закончен и $P_n(A(\sigma_v), t)$ — полином наилучшего приближения. В противном случае строим $P_n(A(\sigma_{v+1}), t)$ и $\sigma_{v+2} = R\sigma_{v+1}$.

Этот процесс за конечное число шагов приводит к экстремальному базису и, значит, к полиному наилучшего приближения. Конечность процесса следует из того, что конечно число базисов и при переборе в силу (4.1) один базис не может встретиться дважды.

Достигнув экстремального базиса σ^* , на следующем шаге получим: $\rho(R\sigma^*) = \rho(\sigma^*)$, а это и есть условие конца процесса.

З а м е ч а н и е. В R -алгоритме начальный базис σ_1 подчинен условию: $\rho(\sigma_1) > 0$, в то время как в алгоритме Валле-Пуссена в качестве начального базиса может быть взят любой базис.

4. Пример. На отрезке $[-1, 1]$ введем сетку T_{1000} , состоящую из точек

$$t_k = -1 + \frac{2k}{1000}, \quad k \in [0 : 1000].$$

Найдем полином наилучшего приближения седьмой степени $P_7(t)$ функции $|t|$ на сетке T_{1000} .

Решение задачи проводилось с помощью R -алгоритма. Результаты вычислений сведены в табл. 1.

Таблица 1

	I	II	III	IV
t_{i_0}	-1,000	-1,000	-1,000	-1,000
t_{i_1}	-0,750	-0,876	-0,884	-0,884
t_{i_2}	-0,500	-0,538	-0,572	-0,572
t_{i_3}	-0,250	-0,184	-0,194	-0,196
t_{i_4}	0,000	0,000	0,000	0,000
t_{i_5}	0,250	0,184	0,194	0,196
t_{i_6}	0,500	0,538	0,572	0,572
t_{i_7}	0,750	0,876	0,884	0,884
t_{i_8}	1,000	1,000	1,000	1,000
$\rho(\sigma)$	0,0390625	0,0455908	0,0459271	0,0459284

В столбце I стоят компоненты начального базиса и величина наилучшего приближения на этом базисе, в столбце IV — компоненты экстремального базиса и величина наилучшего приближения на экстремальном базисе, полученные после трех шагов. Полином наилучшего приближения $P_7(t)$ функции $|t|$ на сетке T_{1000} равен

$$P_7(t) = 2,31038t^6 - 4,17842t^4 + 2,86804t^2 + 0,0459284.$$

§ 5. Сведение к задаче линейного программирования

1. Дискретная задача наилучшего приближения функции алгебраическими полиномами заключается в нахождении величины ρ и вектора A^* коэффициентов полинома наилучшего приближения.

В этом параграфе мы сформулируем две задачи, эквивалентные задаче наилучшего приближения.

2. Будем использовать прежние обозначения:

$$A = (a_0, a_1, \dots, a_n);$$

$$\rho = \inf_{\{A\}} \max_{k \in [0 : N]} \left| y_k - \sum_{i=0}^n a_i t_k^i \right|.$$

Положим

$$\gamma_{ki} = t_k^i, \quad i \in [0 : n], \quad k \in [0 : N];$$

$$\gamma_{N+1+s, i} = -\gamma_{si}, \quad y_{N+1+s} = -y_s \quad \text{для} \quad s \in [0 : N].$$

В этом случае, ввиду равенства $|\xi| = \max\{\xi, -\xi\}$, получим

$$\begin{aligned} \max_{k \in [0 : N]} |y_k - P_n(A, t_k)| &= \max_{k \in [0 : N]} \left| \sum_{i=0}^n \gamma_{ki} a_i - y_k \right| = \\ &= \max_{k \in [0 : 2N+1]} \left(\sum_{i=0}^n \gamma_{ki} a_i - y_k \right) = \max_{k \in [0 : 2N+1]} \left(y_k - \sum_{i=0}^n \gamma_{ki} a_i \right). \end{aligned} \quad (5.1)$$

В $(n+2)$ -мерном евклидовом пространстве векторов V ,

$$V = (a_0, a_1, \dots, a_n, z),$$

рассмотрим полупространства

$$H_k = \left\{ V \mid z \geq \sum_{i=0}^n \gamma_{ki} a_i - y_k \right\}, \quad k \in [0 : 2N+1].$$

Обозначим через G пересечение этих полупространств (рис. 2)

$$G = \bigcap_{k=0}^{2N+1} H_k.$$

Справедлива

Теорема 5.1. Для того чтобы $A^* = (a_0^*, a_1^*, \dots, a_n^*)$ был вектором коэффициентов полинома наилучшего приближения и ρ — величиной наилучшего приближения, необходимо и достаточно, чтобы точка V^* ,

$$V^* = (a_0^*, a_1^*, \dots, a_n^*, \rho),$$

принадлежала G и имела наименьшую последнюю координату среди всех точек множества G .

Доказательство. Необходимость. Поскольку

$$\rho = \max_{k \in [0:N]} |y_k - P_n(A^*, t_k)|,$$

то в силу (5.1) точка V^* принадлежит всем H_k , $k \in [0:2N+1]$, и, значит, G .

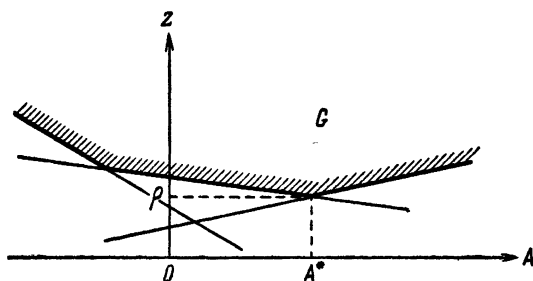


Рис. 2.

Предположим теперь, что существует точка $V_0 \in G$,

$$V_0 = (a_0^{(0)}, a_1^{(0)}, \dots, a_n^{(0)}, z^{(0)}),$$

у которой $z^{(0)} < \rho$. Тогда, полагая $A_0 = (a_0^{(0)}, a_1^{(0)}, \dots, a_n^{(0)})$, получим в силу определения G и (5.1)

$$\max_{k \in [0:N]} |y_k - P_n(A_0, t_k)| \leq z^{(0)} < \rho,$$

что противоречит определению ρ . Необходимость доказана.

Достаточность. Пусть $V_0 = (a_0^{(0)}, a_1^{(0)}, \dots, a_n^{(0)}, z^{(0)})$ — точка из G , имеющая наименьшую последнюю координату. В этом случае, очевидно,

$$\max_{k \in [0:N]} |y_k - P_n(A_0, t_k)| = z^{(0)}, \quad (5.2)$$

где $A_0 = (a_0^{(0)}, a_1^{(0)}, \dots, a_n^{(0)})$. Докажем, что A_0 — вектор коэффициентов полинома наилучшего приближения и $z^{(0)}$ — величина наилучшего приближения.

Пусть A — произвольный $(n+1)$ -мерный вектор. Тогда в силу определения V_0

$$\max_{k \in [0:N]} |y_k - P_n(A, t_k)| \geq z^{(0)}.$$

Учитывая (5.2), получим требуемое.

Теорема доказана.

З а м е ч а н и е. Точка $(a_0^*, a_1^*, \dots, a_n^*, z^*)$ из G с наименьшей последней координатой может быть определена как решение следующей задачи линейного программирования: найти $\min z$ при ограничениях

$$\sum_{i=0}^n \gamma_{ki} a_i - z \leq y_k, \quad k \in [0:2N+1].$$

Таким образом, решение этой задачи линейного программирования дает решение дискретной задачи наилучшего приближения.

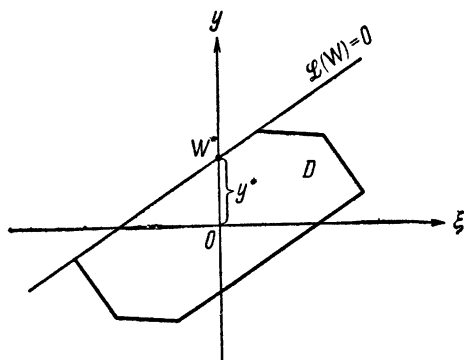


Рис. 3.

3. В $(n+2)$ -мерном евклидовом пространстве векторов $W = (\xi_0, \xi_1, \dots, \xi_n, y)$ рассмотрим точки

$$W_k = (\gamma_{k0}, \dots, \gamma_{kn}, y_k), \quad k \in [0:2N+1].$$

Заметим, что множество $\{W_k\}$ состоит из пар симметричных относительно начала координат точек: $W_{N+1+k} = -W_k$, $k \in [0:N]$. Обозначим через D

выпуклую оболочку, натянутую на $\{W_k\}$, т. е. множество векторов вида

$$W = \sum_{k=0}^{2N+1} \alpha_k W_k; \quad \alpha_k \geq 0, \quad \sum_{k=0}^{2N+1} \alpha_k = 1. \quad (5.3)$$

D содержит начало координат. Рассмотрим точку выхода луча $W = (0, \dots, 0, y)$, $y \geq 0$, из множества D (рис. 3). Обозначим ее через W^* :

$$W^* = (0, \dots, 0, y^*).$$

Лемма 5.1. *Справедливо равенство*

$$y^* = \rho.$$

Доказательство. Точка W^* является точкой из D с первыми $n+1$ нулевыми координатами и максимальной последней координатой. Таким образом, в силу (5.3)

$$y^* = \max \sum_{k=0}^{2N+1} \alpha_k y_k$$

при условии, что α_k удовлетворяют соотношениям

$$\left. \begin{aligned} \sum_{k=0}^{2N+1} \alpha_k \gamma_{ki} &= 0, \quad i \in [0 : n], \\ \alpha_k &\geq 0, \quad \sum_{k=0}^{2N+1} \alpha_k = 1. \end{aligned} \right\} \quad (5.4)$$

Докажем, что $y^* \leq \rho$.

Действительно, для любого вектора $A = (a_0, a_1, \dots, a_n)$ и любого набора чисел α_k , $k \in [0 : 2N+1]$, удовлетворяющего ограничениям (5.4), имеем

$$\begin{aligned} \sum_{k=0}^{2N+1} \alpha_k y_k &= \sum_{k=0}^{2N+1} \alpha_k \left(y_k - \sum_{i=0}^n \gamma_{ki} a_i \right) \leq \\ &\leq \max_{k \in [0 : 2N+1]} \left(y_k - \sum_{i=0}^n \gamma_{ki} a_i \right). \end{aligned}$$

Отсюда получаем

$$y^* \leq \min_{\{A\}} \max_{k \in [0 : 2N+1]} \left(y_k - \sum_{i=0}^n \gamma_{ki} a_i \right) \stackrel{\text{def}}{=} \rho.$$

Для доказательства обратного неравенства достаточно установить существование такого набора чисел $\bar{a}_k$, $k \in [0:2N+1]$, для которого выполняются условия (5.4) и

$$\sum_{k=0}^{2N+1} \bar{a}_k y_k = \rho.$$

Для этого обратимся к основной теореме 3.1 дискретных приближений. Возьмем экстремальный базис σ

$$\sigma = \{t_{k_0} < t_{k_1} < \dots < t_{k_{n+1}}\}.$$

Тогда на основании (3.7) и (3.2)

$$\rho = |h(\sigma)| = \left| \sum_{s=0}^{n+1} (-1)^s \alpha_{k_s}(\sigma) y_{k_s} \right|,$$

причем $\alpha_{k_s}(\sigma) > 0$, $\sum_{s=0}^{n+1} \alpha_{k_s}(\sigma) = 1$,

$$\sum_{s=0}^{n+1} (-1)^s \alpha_{k_s}(\sigma) y_{k_s}^i \stackrel{df}{=} \sum_{s=0}^{n+1} (-1)^s \alpha_{k_s}(\sigma) t_{k_s}^i = 0, \quad i \in [0:n].$$

Если $h(\sigma) \geq 0$, то полагаем:

$\bar{a}_{k_s} = \alpha_{k_s}(\sigma)$ для четных s ,

$\bar{a}_{N+1+k_s} = \alpha_{k_s}(\sigma)$ для нечетных s ,

$\bar{a}_k = 0$ для всех остальных $k \in [0:2N+1]$.

Если $h(\sigma) < 0$, то полагаем:

$\bar{a}_{k_s} = \alpha_{k_s}(\sigma)$ для нечетных s ,

$\bar{a}_{N+1+k_s} = \alpha_{k_s}(\sigma)$ для четных s ,

$\bar{a}_k = 0$ для всех остальных $k \in [0:2N+1]$. В этом случае, как нетрудно проверить,

$$\bar{a}_k \geq 0, \quad \sum_{k=0}^{2N+1} \bar{a}_k = 1; \quad \sum_{k=0}^{2N+1} \bar{a}_k y_{ki} = 0, \quad i \in [0:n];$$

$$\sum_{k=0}^{2N+1} \bar{a}_k y_k = \rho.$$

Лемма доказана.

Теорема 5.2. Для того чтобы $A^* = (a_0^*, a_1^*, \dots, a_n^*)$ был вектором коэффициентов полинома наилучшего

приближения и ρ — величиной наилучшего приближения, необходимо и достаточно чтобы гиперплоскость

$$\mathcal{L}(W) \stackrel{\text{df}}{=} - \sum_{i=0}^n a_i^* \xi_i + y - \rho = 0$$

была опорной к множеству D в точке W^* (см. рис. 3), т. е. чтобы $\mathcal{L}(W^*) = 0$ и $\mathcal{L}(W_k) \leq 0$, $k \in [0 : 2N + 1]$.

Доказательство. Необходимость. В силу (5.1) и определения полинома наилучшего приближения получаем

$$- \sum_{i=0}^n a_i^* \gamma_{ki} + y_k \leq \rho,$$

т. е. $\mathcal{L}(W_k) \leq 0$, $k \in [0 : 2N + 1]$. Остается заметить, что в силу леммы 5.1 $\mathcal{L}(W^*) = 0$.

Достаточность. Пусть гиперплоскость

$$\mathcal{L}_0(W) \stackrel{\text{df}}{=} - \sum_{i=0}^n a_i^{(0)} \xi_i + y - \rho^{(0)} = 0$$

является опорной к множеству D в точке W^* . Тогда из равенства $\mathcal{L}_0(W^*) = 0$ и леммы 5.1 следует, что $\rho^{(0)} = \rho$, т. е. $\rho^{(0)}$ является величиной наилучшего приближения.

Далее, поскольку $\mathcal{L}_0(W_k) \leq 0$, $k \in [0 : 2N + 1]$, то в силу (5.1) получаем

$$\max_{k \in [0 : N]} |y_k - P_n(A_0, t_k)| \leq \rho,$$

где $A_0 = (a_0^{(0)}, a_1^{(0)}, \dots, a_n^{(0)})$. Обратное неравенство очевидно в силу определения ρ . Значит, A_0 является вектором коэффициентов полинома наилучшего приближения.

Теорема доказана.

НЕПРЕРЫВНАЯ ЗАДАЧА НАИЛУЧШЕГО ПРИБЛИЖЕНИЯ ФУНКЦИЙ АЛГЕБРАИЧЕСКИМИ ПОЛИНОМАМИ

§ 1. Постановка задачи

Пусть $f(t)$ — непрерывная функция, заданная на отрезке $[c, d]$. Положим

$$\rho_n(f; c, d) = \rho_n(f) = \inf_{\{A\}} \max_{c \leq t \leq d} |f(t) - P_n(A, t)|.$$

Полином $P_n(A^*, t)$, для которого

$$\max_{c \leq t \leq d} |f(t) - P_n(A^*, t)| = \rho_n(f),$$

называется полиномом наилучшего приближения функции $f(t)$ на отрезке $[c, d]$ или просто полиномом наилучшего приближения.

Задача заключается в фактическом построении полинома наилучшего приближения.

По-прежнему базисом σ будем называть любую упорядоченную по возрастанию систему из $(n+2)$ -х точек, принадлежащих отрезку $[c, d]$,

$$\sigma = \{t_0 < t_1 < \dots < t_{n+1}\}.$$

Множество базисов $\{\sigma\}$, которое будем обозначать через Ξ , бесконечно (даже несчетно). Тем не менее, как и в дискретном случае, оказывается справедливым основное равенство

$$\inf_{\{A\}} \max_{c \leq t \leq d} |f(t) - P_n(A, t)| = \sup_{\sigma \in \Xi} \rho(\sigma),$$

где $\rho(\sigma)$ — наилучшее приближение функции $f(t)$ на базисе σ .

В дальнейшем на основе этого равенства развивается метод последовательных чебышевских интерполяций Е. Я. Ремеза для построения полинома наилучшего приближения.

Рассматривается и другой — сеточный — подход к задаче построения полинома наилучшего приближения.

Для этого на отрезке $[c, d]$ вводится равномерная сетка

$$z_k = c + k \frac{d-c}{N}; \quad w_k = f(z_k), \quad k \in [0 : N].$$

На этой сетке известными методами может быть решена дискретная задача наилучшего приближения.

Оказывается, что при $N \rightarrow \infty$ полиномы наилучшего приближения дискретной задачи равномерно сходятся к полиному наилучшего приближения непрерывной задачи.

В конце главы изучается поведение коэффициентов полиномов наилучшего приближения при $n \rightarrow \infty$.

§ 2. Теорема Чебышева. Полиномы Чебышева

1. Положим

$$\bar{\rho} = \sup_{\sigma \in \Xi} \rho(\sigma).$$

Величина $\bar{\rho}$ конечна. Действительно, если $M = \max_{c \leq t \leq d} |f(t)|$, то для любого базиса $\sigma = \{t_0 < t_1 < \dots < t_{n+1}\}$ будет

$$\begin{aligned} \rho(\sigma) &= \inf_{\{A\}} \max_{k \in [0 : n+1]} |f(t_k) - P_n(A, t_k)| \leq \\ &\leq \max_{k \in [0 : n+1]} |f(t_k) - P_n(0, t_k)| = \max_{k \in [0 : n+1]} |f(t_k)| \leq M. \end{aligned}$$

Значит, $\bar{\rho} \leq M$.

Базис $\sigma^* \in \Xi$, для которого $\rho(\sigma^*) = \bar{\rho}$, называется экстремальным базисом.

Нашей ближайшей целью является доказательство того, что экстремальный базис существует.

2. Определение. Будем говорить, что последовательность базисов $\{\sigma_s\}$, $s = 1, 2, \dots$,

$$\sigma_s = \{t_0^{(s)} < t_1^{(s)} < \dots < t_{n+1}^{(s)}\},$$

сходится к базису σ ,

$$\sigma = \{t_0 < t_1 < \dots < t_{n+1}\},$$

и писать $\sigma_s \xrightarrow{s \rightarrow \infty} \sigma$, если при каждом $k \in [0 : n+1]$

$$t_k^{(s)} \xrightarrow{s \rightarrow \infty} t_k.$$

Лемма 2.1. Если $\sigma_s \xrightarrow{s \rightarrow \infty} \sigma$, то

$$h(\sigma_s) \xrightarrow{s \rightarrow \infty} h(\sigma).$$

Доказательство. Напомним, что

$$h(\sigma_s) = \sum_{k=0}^{n+1} (-1)^k \alpha_k(\sigma_s) f(t_k^{(s)}).$$

Доказательство леммы сводится к установлению следующего факта: при $s \rightarrow \infty$

$$\alpha_k(\sigma_s) \rightarrow \alpha_k(\sigma), \quad k \in [0 : n+1]. \quad (2.1)$$

Вспоминая выражение для $\alpha_k(\sigma_s)$ и $\alpha_k(\sigma)$, заключаем, что для доказательства (2.1) достаточно проверить, что при $s \rightarrow \infty$

$$\frac{1}{t_k^{(s)} - t_i^{(s)}} \rightarrow \frac{1}{t_k - t_i}, \quad i \neq k; \quad i, k \in [0 : n+1].$$

Последнее же предельное соотношение очевидным образом следует из сходимости базисов. Лемма доказана.

Следствие. Если $\sigma_s \xrightarrow{s \rightarrow \infty} \sigma$, то

$$\rho(\sigma_s) \xrightarrow{s \rightarrow \infty} \rho(\sigma).$$

Надо учесть, что $\rho(\sigma_s) = |h(\sigma_s)|$ и $\rho(\sigma) = |h(\sigma)|$.

Лемма 2.2. Пусть последовательность базисов $\{\sigma_s\}$

$$\sigma_s = \{t_0^{(s)} < t_1^{(s)} < \dots < t_{n+1}^{(s)}\},$$

такова, что $\rho(\sigma_s) \geq \beta > 0$, $s = 1, 2, \dots$, и существуют конечные пределы

$$\lim_{s \rightarrow \infty} t_k^{(s)} = t_k, \quad k \in [0 : n+1].$$

Тогда справедливы неравенства

$$t_0 < t_1 < \dots < t_{n+1},$$

т. е. последовательность базисов $\{\sigma_s\}$ сходится к базису σ ,

$$\sigma = \{t_0 < t_1 < \dots < t_{n+1}\}.$$

Доказательство. Очевидно, что

$$t_0 \leq t_1 \leq \dots \leq t_{n+1}.$$

Допустим, что хотя бы одно из последних неравенств обращается в равенство. Тогда среди точек $\{t_k\}$ не более $(n+1)$ различных. Значит, можно построить интерполяционный полином $L_n(t)$, удовлетворяющий соотношениям

$$L_n(t_k) = f(t_k), \quad k \in [0 : n+1].$$

Функция $\Delta(t) = f(t) - L_n(t)$ непрерывна на отрезке $[c, d]$. Следовательно, по $\varepsilon = \beta > 0$ найдется такое $\delta > 0$, что как только $|t'_k - t_k| < \delta$, тотчас $|\Delta(t'_k)| < \beta$.

Выберем индекс s_0 столь большим, чтобы выполнялись неравенства

$$|t_k^{(s_0)} - t_k| < \delta, \quad k \in [0 : n+1].$$

Тогда

$$\begin{aligned} \rho(\sigma_{s_0}) &= \inf_{\{A\}} \max_{k \in [0 : n+1]} |f(t_k^{(s_0)}) - P_n(A, t_k^{(s_0)})| \leq \\ &\leq \max_{k \in [0 : n+1]} |f(t_k^{(s_0)}) - L_n(t_k^{(s_0)})| = \max_{k \in [0 : n+1]} |\Delta(t_k^{(s_0)})| < \beta, \end{aligned}$$

что противоречит условию.

Лемма доказана.

Лемма 2.3. *Экстремальный базис существует.*

Доказательство. Если $\bar{\rho} = 0$, то любой базис является экстремальным. Поэтому будем считать, что $\bar{\rho} > 0$. По определению $\bar{\rho}$ найдется последовательность базисов $\{\sigma_s\}$, $s = 1, 2, \dots$, такая, что

$$\rho(\sigma_s) \xrightarrow{s \rightarrow \infty} \bar{\rho}. \quad (2.2)$$

Выберем индекс $\bar{s}$ столь большим, чтобы при $s > \bar{s}$

$$\rho(\sigma_s) \geq \frac{\bar{\rho}}{2} > 0.$$

Поскольку при всех $k \in [0 : n+1]$ и $s > \bar{s}$

$$c \leq t_k^{(s)} \leq d,$$

то в силу леммы 2.2 найдется подпоследовательность базисов $\{\sigma_{s_i}\}$, сходящаяся к некоторому базису σ^* .

В силу следствия из леммы 2.1

$$\rho(\sigma_{s_i}) \xrightarrow{t \rightarrow \infty} \rho(\sigma^*).$$

Учитывая (2.2), получим

$$\rho(\sigma^*) = \bar{\rho}.$$

Таким образом, σ^* — экстремальный базис.

Лемма доказана.

3. Введем обозначения:

$$\varphi(A) = \max_{c \leq t \leq d} |f(t) - P_n(A, t)|;$$

$$\Delta(A, t) = f(t) - P_n(A, t).$$

Без каких-либо изменений по сравнению со случаем общей дискретной задачи (см. § 3 гл. I) доказываются следующие утверждения.

I. Если $\varphi(A(\sigma)) = \rho(\sigma)$, то $P_n(A(\sigma), t)$ — полином наилучшего приближения непрерывной задачи.

II. Если $\varphi(A(\sigma)) > \rho(\sigma)$, то можно построить новый базис $\sigma_1 = S\sigma$ такой, что $\rho(\sigma_1) > \rho(\sigma)$.

Преобразование S описывается дословно так же, как в дискретном случае; в новый базис σ_1 вместо одного из узлов старого базиса σ привлекается точка, в которой $|\Delta(A(\sigma), t)|$ достигает наибольшего значения $\varphi(A(\sigma))$.

Теперь можно сформулировать основную теорему.

Теорема 2.1 (П. Л. Чебышева). Полином наилучшего приближения функции $f(t)$ на отрезке $[c, d]$ существует и единствен. Для того чтобы полином $P_n(A^, t)$ был полиномом наилучшего приближения, необходимо и достаточно, чтобы он осуществлял чебышевскую интерполяцию на некотором экстремальном базисе.*

Наконец, $\bar{\rho} = \rho_n(f)$, т. е.

$$\inf_{\{A\}} \max_{c \leq t \leq d} |\Delta(A, t)| = \sup_{\sigma \in \Xi} \rho(\sigma). \quad (2.3)$$

Доказательство этой теоремы дословно повторяет доказательство теоремы 3.1 в дискретном случае. Различие по существу лишь в том, что в непрерывной задаче множество базисов бесконечно. Однако эта трудность уже преодолена в лемме 2.3, где доказано существование экстремального базиса,

Следствие. Для того чтобы базис $\sigma \in \Xi$,

$$\sigma = \{t_0 < t_1 < \dots < t_{n+1}\},$$

был экстремальным, и, значит, полином $P_n(A(\sigma), t)$ — полиномом наилучшего приближения, необходимо и достаточно, чтобы выполнялось равенство

$$\varphi(A(\sigma)) = \rho(\sigma), \quad (2.4)$$

или, что то же самое,

$$\max_{c \leq t \leq d} |\Delta(A(\sigma), t)| = \max_{k \in [c; n+1]} |\Delta(A(\sigma), t_k)|.$$

На рис. 4 изображен характерный вид функции $\Delta(A(\sigma^*), t)$ в случае, когда σ^* — экстремальный базис.

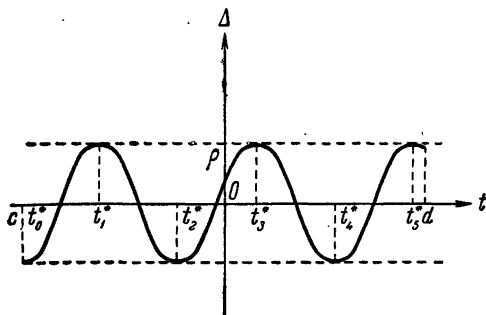


Рис. 4.

4. Рассмотрим на отрезке $[c, d]$ функцию

$$f(t) = t^{n+1}.$$

Найдем для нее полином наилучшего приближения степени $\leq n$ и величину наилучшего приближения $\rho_n(f)$. Для этого положим

$$P_n(A, t) = t^{n+1} - \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} T_{n+1} \left(-1 + \frac{2}{d-c} (t-c) \right),$$

где $T_{n+1}(u)$ — полином Чебышева:

$$T_{n+1}(u) = \cos(n+1) \arccos u, \quad -1 \leq u \leq 1.$$

Нетрудно проверить (см. лемму 2.1 гл. I), что $P_n(A, t)$ действительно является алгебраическим полиномом степени $\leq n$.

Имеем

$$\begin{aligned}\Delta(A, t) &\stackrel{\text{df}}{=} t^{n+1} - P_n(A, t) = \\ &= \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} T_{n+1} \left(-1 + \frac{2}{d-c} (t-c) \right),\end{aligned}$$

откуда следует, что

$$\varphi(A) \stackrel{\text{df}}{=} \max_{c \leq t \leq d} |\Delta(A, t)| = \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1}.$$

Введем базис $\sigma^* = \{t_0 < t_1 < \dots < t_{n+1}\}$, где

$$t_k = \frac{d+c}{2} + \frac{d-c}{2} \cos \frac{(n+1-k)\pi}{n+1}, \quad k \in [0 : n+1].$$

Найдем значения $\Delta(A, t)$ в узлах базиса σ^* :

$$\begin{aligned}\Delta(A, t_k) &= \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} T_{n+1} \left(\cos \frac{(n+1-k)\pi}{n+1} \right) = \\ &= \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} \cos(n+1) \arccos \cos \frac{(n+1-k)\pi}{n+1} = \\ &= (-1)^k \left[(-1)^{n+1} \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} \right].\end{aligned}$$

Получили, что $P_n(A, t)$ осуществляет чебышевскую интерполяцию на базисе σ^* :

$$P_n(A, t) \equiv P_n(A(\sigma^*), t),$$

причем

$$\rho(\sigma^*) = \left| (-1)^{n+1} \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} \right| = \varphi(A(\sigma^*)).$$

Значит, σ^* — экстремальный базис и $P_n(A, t)$ — иско-
мый полином наилучшего приближения. В частности,

$$\rho_n(t^{n+1}; c, d) = \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1}.$$

Полученный результат можно сформулировать иначе: для всякого алгебраического полинома $\tilde{P}_{n+1}(t)$ степени $n+1$ со старшим коэффициентом, равным единице, выполняется неравенство

$$\max_{c \leq t \leq d} |\tilde{P}_{n+1}(t)| \geq \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1};$$

при этом знак равенства достигается только для полинома

$$\tilde{P}_{n+1}^*(t) = \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1} T_{n+1} \left(-1 + \frac{2}{d-c} (t-c) \right).$$

Полином $\tilde{P}_{n+1}^*(t)$ называется полиномом, наименее уклоняющимся от нуля на отрезке $[c, d]$ среди всех полиномов степени $n+1$ со старшим коэффициентом, равным единице.

§ 3. Предельные теоремы

1. Пусть $A(\sigma)$ обозначает вектор коэффициентов полинома наилучшего приближения функции f на базисе σ .

Лемма 3.1. *Если последовательность базисов σ_s , $s = 1, 2, \dots$, сходится к базису σ_0 , то*

$$A(\sigma_s) \xrightarrow{s \rightarrow \infty} A(\sigma_0)$$

(сходимость векторов понимается как покоординатная — см. Приложение I, § 2).

Доказательство. Во-первых, в силу леммы 2.1

$$h(\sigma_s) \xrightarrow{s \rightarrow \infty} h(\sigma_0). \quad (3.1)$$

Далее заметим, что полиномы $P_n(A(\sigma_0), t)$ и $P_n(A(\sigma_s), t)$ удовлетворяют соотношениям

$$\left. \begin{aligned} P_n(A(\sigma_0), t_k^{(0)}) &= y_k^{(0)}, \\ P_n(A(\sigma_s), t_k^{(s)}) &= y_k^{(s)}, \end{aligned} \right\} \quad (3.2)$$

где

$$\begin{aligned} y_k^{(0)} &= f(t_k^{(0)}) - (-1)^k h(\sigma_0), \\ y_k^{(s)} &= f(t_k^{(s)}) - (-1)^k h(\sigma_s), \\ k &\in [0 : n], \quad s = 1, 2, \dots \end{aligned}$$

(равенства (3.2) справедливы и при $k = n+1$, однако сейчас это не важно).

Таким образом, полином $P_n(A(\sigma_0), t)$ и последовательность полиномов $P_n(A(\sigma_s), t)$, $s = 1, 2, \dots$, являются интерполяционными. Поскольку по условию $t_k^{(s)} \xrightarrow{s \rightarrow \infty} t_k^{(0)}$, $k \in [0 : n]$, и в силу непрерывности $f(t)$ и предельного

соотношения (3.1) $y_k^{(s)} \xrightarrow{s \rightarrow \infty} y_k^{(0)}$, то по лемме о непрерывной зависимости интерполяционного полинома (см. Приложение I, § 2)

$$A(\sigma_s) \xrightarrow{s \rightarrow \infty} A(\sigma_0).$$

Лемма доказана.

Положим

$$\varphi(A) = \max_{c \leq t \leq d} |f(t) - P_n(A, t)|.$$

Лемма 3.2. Если последовательность векторов A_s , $s = 1, 2, \dots$, сходится к вектору A^* , то

$$\varphi(A_s) \xrightarrow{s \rightarrow \infty} \varphi(A^*).$$

Доказательство. Воспользуемся следующим неравенством, справедливым для любых непрерывных на отрезке $[c, d]$ функций $F_1(t)$ и $F_2(t)$ (см. Приложение III, § 2):

$$|\max_{c \leq t \leq d} |F_1(t)| - \max_{c \leq t \leq d} |F_2(t)|| \leq \max_{c \leq t \leq d} |F_1(t) - F_2(t)|.$$

Получим

$$\begin{aligned} |\varphi(A_s) - \varphi(A^*)| &\leq \max_{c \leq t \leq d} |P_n(A^*, t) - P_n(A_s, t)| = \\ &= \max_{c \leq t \leq d} \left| \sum_{i=0}^n (a_i^* - a_i^{(s)}) t^i \right| \leq M \sum_{i=0}^n |a_i^* - a_i^{(s)}|. \end{aligned}$$

Остается перейти к пределу при $s \rightarrow \infty$.

Лемма доказана.

2. Теорема 3.1. Если последовательность алгебраических полиномов $P_n(A_s, t)$, $s = 1, 2, \dots$, такова, что при $s \rightarrow \infty$

$$\varphi(A_s) \rightarrow \rho_n(f), \quad (3.3)$$

то $A_s \rightarrow A^*$, где A^* — вектор коэффициентов полинома наилучшего приближения.

Доказательство. Нужно доказать, что при всех $k \in [0: n]$

$$a_k^{(s)} \xrightarrow{s \rightarrow \infty} a_k^*. \quad (3.4)$$

Предположим, что при $k = k_0$ предельное соотношение (3.4) не выполняется. Тогда найдутся $\varepsilon_0 > 0$ и последовательность индексов $\{s_i\}$, $i = 1, 2, \dots$, такие, что

$$|a_{k_0}^{(s_i)} - a_{k_0}^*| \geq \varepsilon_0. \quad (3.5)$$

Подпоследовательность полиномов $\{P_n(A_{s_i}, t)\}$, $i = 1, 2, \dots$, ограничена на $[c, d]$ равномерно по i . Действительно, в силу (3.3)

$$\varphi(A_{s_i}) \leq M_1.$$

Отсюда

$$\begin{aligned} \max_{c \leq t \leq d} |P_n(A_{s_i}, t)| &\leq \\ &\leq \max_{c \leq t \leq d} |P_n(A_{s_i}, t) - f(t)| + \max_{c \leq t \leq d} |f(t)| = \\ &= \varphi(A_{s_i}) + \max_{c \leq t \leq d} |f(t)| \leq M_2. \end{aligned}$$

В силу леммы о компактности (см. Приложение I, § 2) найдется подпоследовательность векторов $\{A_{s_{i_r}}\}$ и вектор A_0 такие, что

$$A_{s_{i_r}} \xrightarrow{r \rightarrow \infty} A_0.$$

На основании леммы 3.2 при $r \rightarrow \infty$

$$\varphi(A_{s_{i_r}}) \rightarrow \varphi(A_0).$$

Учитывая (3.3), получим, что $\varphi(A_0) = \rho_n(f)$, т. е. полином $P_n(A_0, t)$ есть полином наилучшего приближения. В силу единственности полинома наилучшего приближения $A_0 = A^*$, так что

$$A_{s_{i_r}} \xrightarrow{r \rightarrow \infty} A^*.$$

Но это противоречит (3.5). Теорема доказана.

Теорема 3.2. Если последовательность базисов $\{\sigma_s\}$, $s = 1, 2, \dots$, такова, что при $s \rightarrow \infty$

$$\rho(\sigma_s) \rightarrow \rho_n(f),$$

то $A(\sigma_s) \rightarrow A^*$, где A^* — вектор коэффициентов полинома наилучшего приближения.

Доказательство. Если $\rho_n(f) = 0$, то утверждение теоремы очевидно. Поэтому пусть $\rho_n(f) > 0$. В силу теоремы 3.1 достаточно доказать, что

$$\varphi(A(\sigma_s)) \xrightarrow{s \rightarrow \infty} \rho_n(f).$$

Допустим противное, т. е. что для некоторого $\varepsilon_0 > 0$ найдется подпоследовательность базисов $\{\sigma_{s_i}\}$, $i = 1, 2, \dots$, такая, что

$$\varphi(A(\sigma_{s_i})) - \rho_n(f) \geq \varepsilon_0. \quad (3.6)$$

Поскольку для достаточно больших номеров s_i

$$\rho(\sigma_{s_i}) \geq \frac{1}{2} \rho_n(f) > 0,$$

то в силу леммы 2.2 найдется подпоследовательность базисов $\{\sigma_{s_{i_r}}\}$, $r = 1, 2, \dots$, сходящаяся к некоторому базису σ_0 :

$$\sigma_{s_{i_r}} \xrightarrow{r \rightarrow \infty} \sigma_0. \quad (3.7)$$

Но тогда $\rho(\sigma_{s_{i_r}}) \rightarrow \rho(\sigma_0)$ (следствие из леммы 2.1) и, значит, $\rho(\sigma_0) = \rho_n(f)$.

Получили, что σ_0 есть экстремальный базис, так что

$$\varphi(A(\sigma_0)) = \rho(\sigma_0) = \rho_n(f).$$

Кроме того, из (3.7) на основании леммы 3.1 следует, что

$$A(\sigma_{s_{i_r}}) \xrightarrow{r \rightarrow \infty} A(\sigma_0).$$

Теперь воспользуемся леммой 3.2. Это дает

$$\varphi(A(\sigma_{s_{i_r}})) \rightarrow \varphi(A(\sigma_0)) = \rho_n(f),$$

что противоречит (3.6).

Теорема доказана.

§ 4. Метод последовательных чебышевских интерполяций Ремеза

1. Пусть на отрезке $[c, d]$ задана непрерывная функция $f(t)$. Опишем в самом общем виде метод последовательных чебышевских интерполяций для нахождения полинома наилучшего приближения функции $f(t)$.

Произвольным образом выбираем начальный базис σ_1 и осуществляем на нем чебышевскую интерполяцию, т. е. строим полином $P_n(A(\sigma_1), t)$. Предполагается, что $\rho(\sigma_1) > 0$.

Пусть уже построены базис σ_s и полином $P_n(A(\sigma_s), t)$. Если $\varphi(A(\sigma_s)) = \rho(\sigma_s)$, то $P_n(A(\sigma_s), t)$ — искомый полином наилучшего приближения, и процесс на этом заканчивается. Пусть $\varphi(A(\sigma_s)) > \rho(\sigma_s)$. Тогда строим новый базис σ_{s+1} так, чтобы выполнялись следующие три условия:

$$I. \quad m_s \stackrel{df}{=} \min_{k \in [0 : n+1]} |\Delta_s(t_k^{(s+1)})| \geq \rho(\sigma_s) \stackrel{df}{=} \rho_s,$$

где $\Delta_s(t) = f(t) - P_n(A(\sigma_s), t)$.

$$II. \quad \text{sign } \Delta_s(t_k^{(s+1)}) = -\text{sign } \Delta_s(t_{k+1}^{(s+1)}), \quad k \in [0 : n].$$

$$III. \quad l_s \stackrel{df}{=} \max_{k \in [0 : n+1]} |\Delta_s(t_k^{(s+1)})| > \rho_s.$$

На базисе σ_{s+1} осуществляем чебышевскую интерполяцию, т. е. строим полином $P_n(A(\sigma_{s+1}), t)$.

Продолжая аналогично, получим последовательность

$$P_n(A(\sigma_1), t), P_n(A(\sigma_2), t), \dots, P_n(A(\sigma_s), t), \dots$$

Если эта последовательность конечна, то последний ее элемент по построению является полиномом наилучшего приближения функции $f(t)$ на отрезке $[c, d]$.

Рассмотрим случай, когда последовательность $\{P_n(A(\sigma_s), t)\}$ бесконечна. Введем обозначение

$$\varphi_s = \max_{c \leq t \leq d} |\Delta_s(t)|.$$

Справедлива

Теорема 4.1. Если

$$\lim_{s \rightarrow \infty} \frac{l_s - \rho_s}{\varphi_s - \rho_s} > 0, \quad (4.1)$$

то $A(\sigma_s) \rightarrow A^*$, где A^* — вектор коэффициентов полинома наилучшего приближения.

При этом при некотором θ , $0 < \theta < 1$,

$$\rho_n(f) - \rho_s \leq M\theta^s. \quad (4.2)$$

Константа M не зависит от s .

2. Предварительно докажем одну лемму.

Лемма 4.1. Если для последовательности базисов $\{\sigma_s\}$, $s = 1, 2, \dots$, отрезка $[c, d]$ выполняется неравенство

$$\rho(\sigma_s) \geq \beta > 0,$$

то существует такое $\eta > 0$, что при любых $k_1 \neq k_2$ ($k_1, k_2 \in [0 : n + 1]$) и $s = 1, 2, \dots$ будет

$$|t_{k_1}^{(s)} - t_{k_2}^{(s)}| \geq \eta. \quad (4.3)$$

Кроме того,

$$\theta_1 \stackrel{df}{=} \inf_{\{s\}} \min_{k \in [0 : n + 1]} \alpha_k(\sigma_s) \geq \frac{1}{n + 2} \left(\frac{\eta}{d - c} \right)^{n + 1}. \quad (4.4)$$

Доказательство. Предположим противное, т. е. что требуемого η не существует. Тогда найдутся индексы k_1^0 и k_2^0 , $k_1^0 \neq k_2^0$, и подпоследовательность индексов $\{s_r\}$, $r = 1, 2, \dots$, такие, что

$$t_{k_1^0}^{(s_r)} - t_{k_2^0}^{(s_r)} \xrightarrow{r \rightarrow \infty} 0 \quad (4.5)$$

(пользуемся конечностью значений индексов k_1 и k_2 , $k_1, k_2 \in [0 : n + 1]$). Из подпоследовательности $\{s_r\}$ можно выделить такую подпоследовательность $\{s_{r_m}\}$, $m = 1, 2, \dots$, что будут существовать конечные пределы

$$\lim_{m \rightarrow \infty} t_k^{(s_{r_m})} = t_k^{(0)}, \quad k \in [0 : n + 1].$$

При этом в силу (4.5) $t_{k_1^0}^{(0)} = t_{k_2^0}^{(0)}$. Учитывая, что по-прежнему

$$\rho(\sigma_{s_{r_m}}) \geq \beta > 0,$$

приходим к противоречию с леммой 2.2.

Таким образом, доказано существование требуемого η . Неравенство (4.4) легко следует из (4.3).

Действительно,

$$\begin{aligned} \alpha_k(\sigma_s) &= \frac{(-1)^k}{(t_k^{(s)} - t_0^{(s)}) \dots (t_k^{(s)} - t_{n+1}^{(s)})} = \\ &= \frac{\sum_{i=0}^n \frac{(-1)^i}{(t_i^{(s)} - t_0^{(s)}) \dots (t_i^{(s)} - t_{n+1}^{(s)})}}{1} = \\ &= \frac{1}{\sum_{i=0}^n \frac{1}{|t_i^{(s)} - t_0^{(s)}| \dots |t_i^{(s)} - t_{n+1}^{(s)}|}} \geq \frac{1}{n+2} \left(\frac{\eta}{d-c} \right)^{n+1}. \end{aligned}$$

Лемма доказана.

3. Доказательство теоремы. Условие (4.1) означает, что найдется такое ε_1 , $0 < \varepsilon_1 < 1$, что начиная с некоторого номера $\bar{s}$, будет

$$\frac{l_s - \rho_s}{\varphi_s - \rho_s} \geq \varepsilon_1$$

или

$$l_s - \rho_s \geq \varepsilon_1 (\varphi_s - \rho_s). \quad (4.6)$$

В силу III можно считать, что неравенство (4.6) выполняется для всех натуральных s .

Положим

$$\theta_1 = \inf_{\{s\}} \min_{k \in [0 : n+1]} \alpha_k(\sigma_s).$$

По определению $0 \leq \theta_1 < 1$. Справедливо следующее неравенство:

$$\rho_{s+1} \geq \rho_s + \theta_1 (l_s - \rho_s). \quad (4.7)$$

Действительно, пусть $l_s = |\Delta_s(t_k^{(s+1)})|$. Учитывая I — III, будем иметь

$$\begin{aligned} \rho_{s+1} &= \left| \sum_{k=0}^{n+1} (-1)^k \alpha_k(\sigma_{s+1}) f(t_k^{(s+1)}) \right| = \\ &= \sum_{k=0}^{n+1} \alpha_k(\sigma_{s+1}) |\Delta_s(t_k^{(s+1)})| \geq \\ &\geq \sum_{k=0}^{n+1} \alpha_k(\sigma_{s+1}) \rho_s + \alpha_{k_0}(\sigma_{s+1}) (l_s - \rho_s) \geq \rho_s + \theta_1 (l_s - \rho_s). \end{aligned}$$

Из (4.7), в частности, следует, что $\rho(\sigma_s) \geq \rho(\sigma_1)$ при всех s . Поскольку, по предположению, $\rho(\sigma_1) > 0$, то в силу леммы 4.1 $0 < \theta_1 < 1$.

Заменив теперь в (4.7) s на $s-1$ и учитывая (4.6), получим

$$\begin{aligned} \rho_s - \rho_{s-1} &\geq \theta_1 (l_{s-1} - \rho_{s-1}) \geq \\ &\geq \theta_1 e_1 (\varphi_{s-1} - \rho_{s-1}) \geq \theta_1 e_1 (\rho_n(f) - \rho_{s-1}). \end{aligned}$$

Перепишем неравенство

$$\rho_s - \rho_{s-1} \geq \theta_1 e_1 (\rho_n(f) - \rho_{s-1})$$

в другом виде:

$$\rho_n(f) - \rho_s \leq \theta (\rho_n(f) - \rho_{s-1}), \quad (4.8)$$

где $\theta = 1 - \theta_1 e_1$, $0 < \theta < 1$.

Из (4.8) следует, что

$$\rho_n(f) - \rho_s \leq \theta^{s-1} (\rho_n(f) - \rho_1) \stackrel{df}{=} M \theta^s.$$

Тем самым доказано неравенство (4.2). Соотношение $A(\sigma_s) \xrightarrow{s \rightarrow \infty} A^*$ следует теперь из (4.2) и теоремы 3.2. Теорема доказана.

З а м е ч а н и е. Условие (4.1) наверняка выполняется, если $l_s = \varphi_s$ при всех натуральных s . Это соответствует тому, что на каждом шаге в новый базис σ_{s+1} включается узел, в котором достигается $\max_{c \leq t \leq d} |\Delta_s(t)|$ (см. алгоритм Валле-Пуссена и R -алгоритм).

В общем случае для сходимости процесса достаточно заботиться лишь о том, чтобы отношение

$$\frac{l_s - \rho_s}{\varphi_s - \rho_s}$$

было ограничено снизу положительным числом равномерно по s .

4. Остановимся на выборе начального базиса σ_1 . Нас интересует такой базис σ_1 , для которого величина $\rho(\sigma_1)$ была бы по возможности максимальной. Напомним, что рассматривается непрерывная на отрезке $[c, d]$ функция $f(t)$ и что

$$\rho(\sigma_1) = \left| \frac{f[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]}{\chi[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]} \right|.$$

Числитель этой дроби зависит от функции f , и добиться его увеличения за счет специального выбора базиса σ_1 сразу для всех непрерывных функций невозможно. Поэтому будем максимизировать величину

$$\frac{1}{|\chi[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]|}.$$

Если обозначить $\Phi(t) = t^{n+1}$, то в силу одного из свойств разделенных разностей (см. Приложение I, § 1)

$$\frac{1}{|\chi[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]|} = \left| \frac{\Phi[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]}{\chi[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]} \right| \stackrel{\text{df}}{=} \rho_\Phi(\sigma_1).$$

Однако

$$\rho_\Phi(\sigma_1) \leq \rho_n(\Psi; c, d),$$

и равенство достигается для экстремального базиса σ^* задачи наилучшего приближения функции $\Phi(t) = t^{n+1}$ на отрезке $[c, d]$ полиномами степени $\leq n$.

Учитывая результаты п. 4 § 2, получим

$$\frac{1}{|\chi[t_0^{(1)}, t_1^{(1)}, \dots, t_{n+1}^{(1)}]|} \leq \frac{1}{2^n} \left(\frac{d-c}{2} \right)^{n+1},$$

и равенство будет достигнуто, если в качестве компонент базиса σ_1 взять

$$t_k^{(1)} = \frac{d+c}{2} + \frac{d-c}{2} \cos \frac{(n+1-k)\pi}{n+1}, \quad k \in [0 : n+1].$$

§ 5. Метод сеток

1. На отрезке $[c, d]$ по-прежнему рассматривается непрерывная функция $f(t)$ и ищется ее полином наилучшего приближения $P_n(A^*, t)$.

Положим

$$z_k = c + k \frac{d-c}{N}, \quad w_k = f(z_k), \quad k \in [0 : N].$$

Множество точек $\{z_k\}$ будем обозначать через T_N .

Методами главы I можно построить алгебраический полином $P_n(A_N, t)$, являющийся решением следующей дискретной задачи:

$$\max_{k \in [0 : N]} |w_k - P_n(A_N, z_k)| = \rho_N, \quad (5.1)$$

где

$$\rho_N = \inf_{\{A\}} \max_{k \in [0:N]} |\omega_k - P_n(A, z_k)|.$$

Теорема 5.1 (о сходимости сеточного процесса).
При $N \rightarrow \infty$ справедливо предельное соотношение

$$A_N \rightarrow A^*.$$

Доказательство. Обозначим через $\sigma^* = \{t_0^* < t_1^* < \dots < t_{n+1}^*\}$ экстремальный базис в задаче наилучшего приближения на всем отрезке $[c, d]$ (в силу леммы 2.3 такой экстремальный базис существует). При достаточно больших N рассмотрим базисы $\sigma_N = \{z_0^{(N)} < z_1^{(N)} < \dots < z_{n+1}^{(N)}\}$, где $z_k^{(N)}$ — ближайший к t_k^* узел сетки T_N , $k \in [0:n+1]$.

Очевидно, $\sigma_N \rightarrow \sigma^*$ при $N \rightarrow \infty$, так что

$$\rho(\sigma_N) \rightarrow \rho(\sigma^*). \quad (5.2)$$

Пусть σ_N^* есть экстремальный базис сетки T_N . Тогда (см. (5.1))

$$\rho(\sigma_N^*) = \rho_N, \quad A(\sigma_N^*) = A_N.$$

Поскольку σ_N — один из возможных базисов сетки T_N , то

$$\rho(\sigma_N) \leq \rho(\sigma_N^*). \quad (5.3)$$

С другой стороны,

$$\rho(\sigma_N^*) \leq \rho(\sigma^*), \quad (5.4)$$

ибо σ_N^* есть один из возможных базисов отрезка $[c, d]$.

Учитывая (5.3), (5.4) и (5.2), получим

$$\rho(\sigma_N^*) \rightarrow \rho(\sigma^*) = \rho_n(f),$$

откуда на основании теоремы 3.2 следует, что

$$A_N \rightarrow A^*.$$

Теорема доказана.

Заметим, что неравенство (5.4) равносильно следующему неравенству:

$$\rho_N \leq \rho_n(f). \quad (5.5)$$

2. В этом пункте будет получена оценка скорости сходимости сеточного процесса.

Теорема 5.2. Пусть функция $f(t)$ имеет непрерывную вторую производную на отрезке $[c, d]$. Тогда, если A_N есть вектор коэффициентов полинома наилучшего приближения на сетке T_N , то

$$\varphi(A_N) - \rho_n(f) \leq M \left(\frac{d-c}{N} \right)^2, \quad (5.6)$$

где

$$\varphi(A_N) = \max_{c \leq t \leq d} |f(t) - P_n(A_N, t)|$$

и M — константа, не зависящая от N .

Доказательство. Положим

$$\Delta_N(t) = f(t) - P_n(A_N, t).$$

Учитывая (5.5), будем иметь

$$\begin{aligned} \varphi(A_N) - \rho_n(f) &\leq \varphi(A_N) - \rho_N = \\ &= \max_{c \leq t \leq d} |\Delta_N(t)| - \max_{k \in [0:N]} |\Delta_N(z_k)|. \end{aligned}$$

Обозначим через $t_N \in [c, d]$ точку, для которой

$$|\Delta_N(t_N)| = \max_{c \leq t \leq d} |\Delta_N(t)|. \quad (5.7)$$

Если $\Delta_N(t_N) = 0$ или t_N совпадает с одним из узлов сетки T_N , то

$$\varphi(A_N) - \rho_n(f) = 0,$$

и неравенство (5.6) выполняется при любом $M \geq 0$. Поэтому будем считать, что $\Delta_N(t_N) \neq 0$ и что t_N не совпадает ни с одним из узлов сетки T_N . В частности, t_N лежит внутри отрезка $[c, d]$, и в силу (5.7)

$$\Delta'_N(t_N) = 0. \quad (5.8)$$

Обозначим через z_s ближайший к t_N узел сетки T_N . Учитывая формулу Тейлора, будем иметь

$$\begin{aligned} \Delta_N(z_s) &= \Delta_N(t_N) + \\ &+ \frac{\Delta'_N(t_N)}{1!} (z_s - t_N) + \frac{\Delta''_N(t_N + \theta(z_s - t_N))}{2!} (z_s - t_N)^2, \end{aligned}$$

где $0 < \theta < 1$. Отсюда и из (5.8) следует неравенство

$$|\Delta_N(t_N)| - |\Delta_N(z_s)| \leq \frac{M_{N2}}{8} \left(\frac{d-c}{N} \right)^2,$$

где $M_{N2} = \max_{c \leq t \leq d} |\Delta_N''(t)|$. Тем более

$$\max_{c \leq t \leq d} |\Delta_N(t)| - \max_{k \in [0:N]} |\Delta_N(z_k)| \leq \frac{M_{N2}}{8} \left(\frac{d-c}{N} \right)^2.$$

Осталось доказать, что последовательность $\{M_{N2}\}$ ограничена равномерно по N . Для этого воспользуемся теоремой 5.1 и леммой 2.3 Приложения I. В силу теоремы 5.1 $A_N \xrightarrow{N \rightarrow \infty} A^*$, поэтому

$$\max_{c \leq t \leq d} |P_n(A_N, t)| \leq M_1.$$

Отсюда на основании леммы 2.3 Приложения I следует, что

$$\max_{c \leq t \leq d} |P_n''(A_N, t)| \leq M_1^{(2)}.$$

Теперь имеем

$$\begin{aligned} M_{N2} &\leq \max_{c \leq t \leq d} |f''(t)| + \max_{c \leq t \leq d} |P_n''(A_N, t)| \leq \\ &\leq \max_{c \leq t \leq d} |f''(t)| + M_1^{(2)} \stackrel{\text{df}}{=} M_2. \end{aligned}$$

Теорема доказана.

§ 6*. О поведении коэффициентов полиномов наилучшего приближения

1. В силу известной теоремы Вейерштрасса всякая непрерывная на отрезке $[c, d]$ функция $f(t)$ допускает равномерную аппроксимацию с любой степенью точности алгебраическими полиномами. Это означает, что по любому $\varepsilon > 0$ найдется такой алгебраический полином $P(t)$, что

$$\max_{c \leq t \leq d} |f(t) - P(t)| < \varepsilon.$$

Дополнением к этому классическому результату является следующая

Теорема 6.1. Для того чтобы непрерывная на отрезке $[-1, 1]$ функция $f(t)$ допускала равномерную аппроксимацию с любой степенью точности алгебраическими полиномами $Q_n(t)$,

$$Q_n(t) = \sum_{k=0}^n a_k^{(n)} t^k,$$

с равномерно ограниченными коэффициентами

$$|a_k^{(n)}| \leq M, \quad k \in [0 : n],$$

необходимо и достаточно, чтобы для $t \in (-1, 1)$ имело место представление

$$f(t) = \sum_{k=0}^{\infty} a_k t^k; \quad |a_k| \leq M, \quad k = 0, 1, 2, \dots \quad (6.1)$$

Доказательство. Достаточность. Пусть функция $f(t)$ непрерывна на отрезке $[-1, 1]$ и для $t \in (-1, 1)$ имеет место представление (6.1).

Положим

$$F_n(t) = \sum_{k=0}^n a_k t^k.$$

По $\varepsilon > 0$ выберем такое r , $0 < r < 1$, чтобы для всех $t \in [-1, 1]$ выполнялось неравенство

$$|f(t) - f(rt)| < \frac{\varepsilon}{2}.$$

Поскольку ряд в (6.1) равномерно сходится на отрезке $[-r, r]$, то найдется натуральное n , при котором

$$|f(rt) - F_n(rt)| < \frac{\varepsilon}{2}$$

равномерно по $t \in [-1, 1]$. Имеем для $t \in [-1, 1]$

$$|f(t) - F_n(rt)| < \varepsilon.$$

Остается заметить, что коэффициенты алгебраического полинома $F_n(rt)$ не превосходят по модулю M .

Необходимость. Выделим последовательность полиномов

$$Q_n(t) = \sum_{k=0}^n a_k^{(n)} t^k, \quad |a_k^{(n)}| \leq M,$$

равномерно сходящуюся при $n \rightarrow \infty$ к $f(t)$ на отрезке $[-1, 1]$. Докажем, что при каждом k числовые последовательности $\{a_k^{(n)}\}$ сходятся. Действительно,

$$a_0^{(n)} = Q_n(0) \xrightarrow{n \rightarrow \infty} f(0).$$

Предположим, что сходятся последовательности $\{a_k^{(n)}\}$ при $k \in [0 : \nu - 1]$. Докажем, что сходимость имеет место и при $k = \nu$.

Положим для $n \geq \nu$

$$q_n^{(\nu)}(t) \stackrel{\text{df}}{=} \frac{Q_n(t) - \sum_{k=0}^{\nu-1} a_k^{(n)} t^k}{t^\nu} = a_\nu^{(n)} + a_{\nu+1}^{(n)} t + \dots + a_n^{(n)} t^{n-\nu}.$$

По определению $Q_n(t)$ и индукционному предположению последовательность $\{q_n^{(\nu)}(t)\}$ сходится для любого $t \in [-1, 1]$, отличного от нуля. Докажем, что сходимость имеет место и при $t = 0$. Для этого заметим, что при $t \in (-1, 1)$

$$|q_n^{(\nu)}(t) - q_n^{(\nu)}(0)| = \left| t \sum_{k=\nu+1}^n a_k^{(n)} t^{k-\nu-1} \right| \leq \frac{M|t|}{1-|t|}.$$

Далее,

$$|q_n^{(\nu)}(0) - q_{n+m}^{(\nu)}(0)| \leq \frac{2M|t|}{1-|t|} + |q_n^{(\nu)}(t) - q_{n+m}^{(\nu)}(t)|. \quad (6.2)$$

Возьмем $\varepsilon > 0$. Выберем такое $t \neq 0$, чтобы первое слагаемое в правой части неравенства (6.2) было меньше $\varepsilon/2$. После этого, учитывая, что последовательность $\{q_n^{(\nu)}(t)\}$ сходится, выберем такое n , чтобы для всех натуральных m второе слагаемое было меньше $\varepsilon/2$. Тогда для всех натуральных m получим

$$|q_n^{(\nu)}(0) - q_{n+m}^{(\nu)}(0)| < \varepsilon.$$

Отсюда следует, что последовательность $\{q_n^{(\nu)}(0)\}$ сходится. Учитывая равенство $a_\nu^{(n)} = q_n^{(\nu)}(0)$, заключаем, что сходится и последовательность $\{a_\nu^{(n)}\}$.

Итак, $a_k^{(n)} \xrightarrow{n \rightarrow \infty} a_k$ для всех $k = 0, 1, 2, \dots$, причем

$$|a_k| \leq M. \text{ Рассмотрим ряд } \sum_{k=0}^{\infty} a_k t^k.$$

Докажем, что при $-1 < t < 1$ имеет место равенство

$$f(t) = \sum_{k=0}^{\infty} a_k t^k. \quad (6.3)$$

Зафиксируем t , $|t| < 1$, и возьмем произвольное $\varepsilon > 0$. Выберем N так, чтобы

$$\frac{M|t|^{N+1}}{1-|t|} < \frac{\varepsilon}{4}$$

и чтобы для $n > N$

$$|f(t) - Q_n(t)| < \frac{\varepsilon}{4}.$$

Теперь найдем $\bar{n} > N$, при котором

$$\sum_{k=0}^N |a_k^{(\bar{n})} - a_k| < \frac{\varepsilon}{4}.$$

В этом случае получим

$$\begin{aligned} \left| \sum_{k=0}^{\infty} a_k t^k - f(t) \right| &\leq \left| \sum_{k=0}^{\infty} a_k t^k - Q_{\bar{n}}(t) \right| + |Q_{\bar{n}}(t) - f(t)| \leq \\ &\leq \sum_{k=0}^N |a_k^{(\bar{n})} - a_k| + \frac{2M|t|^{N+1}}{1-|t|} + \frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

Отсюда ввиду произвольности ε следует (6.3).

Теорема доказана.

2. Обозначим через $P_n(A_n, t)$, $A_n = (a_0^{(n)}, \dots, a_n^{(n)})$, полином наилучшего приближения степени $\leq n$ непрерывной на отрезке $[-1, 1]$ функции $f(t)$. В силу теоремы Вейерштрасса последовательность $\{P_n(A_n, t)\}$ при $n \rightarrow \infty$ равномерно на $[-1, 1]$ сходится к $f(t)$.

Положим

$$|A_n| = \max_{k \in [0: n]} |a_k^{(n)}|.$$

Для того чтобы последовательность $\{A_n\}$ была ограничена, необходимо в силу теоремы 6.1, чтобы для функции $f(t)$ при $-1 < t < 1$ имело место представление (6.1). В противном случае $|A_n| \xrightarrow{n \rightarrow \infty} \infty$.

Положим

$$\delta_n = \frac{|A_n|}{\max_{-1 \leq t \leq 1} |P_n(A_n, t)|}.$$

При $n \rightarrow \infty$

$$\max_{-1 \leq t \leq 1} |P_n(A_n, t)| \rightarrow \max_{-1 \leq t \leq 1} |f(t)|.$$

Поэтому, если $|A_n| \rightarrow \infty$, то $\delta_n \rightarrow \infty$.

При больших δ_n говорят о плохой обусловленности полинома $P_n(A_n, t)$.

В качестве примера плохо обусловленного полинома приведем полином наилучшего приближения 14-й степени $P_{14}(t)$ функции $|t|$ на $[-1, 1]$:

$$P_{14}(t) = 198,46t^{14} - 714,90t^{12} + 1022,98t^{10} - 739,91t^8 + \\ + 285,33t^6 - 57,607t^4 + 6,6443t^2 + 0,019949.$$

В этом случае $\delta_{14} \approx 10^3$.

Заметим, что на практике редко используются полиномы наилучшего приближения степени, большей чем 10, а плохая обусловленность появляется лишь при больших степенях n .

ДИСКРЕТНАЯ МИНИМАКСНАЯ ЗАДАЧА

§ 1. Постановка задачи

Пусть $f_i(X)$, $i \in [0 : N]$, $X = (x_1, x_2, \dots, x_n)$ — функции, заданные на всем n -мерном евклидовом пространстве E_n . Положим

$$\mu = \inf_{\{X\}} \max_{i \in [0 : N]} f_i(X).$$

Задача заключается в нахождении точки X^* , для которой

$$\max_{i \in [0 : N]} f_i(X^*) = \mu.$$

Введем обозначение

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X).$$

Поскольку

$$\mu = \inf_{X \in E_n} \varphi(X),$$

то исходную минимаксную задачу можно рассматривать как задачу минимизации функции $\varphi(X)$ на E_n .

В связи с этим в дальнейшем подробно изучаются свойства функции максимума $\varphi(X)$. Устанавливается, в частности, что при естественных предположениях функция $\varphi(X)$ дифференцируема по всем направлениям. Получены необходимые условия минимума функции $\varphi(X)$ и дана их геометрическая интерпретация.

Много внимания уделяется методам последовательных приближений для нахождения стационарных точек функции $\varphi(X)$, т. е. точек, удовлетворяющих необходимому условию минимума. Приводится, в частности, пример, показывающий, что метод наискорейшего спуска может сходиться к точке, не являющейся стационарной.

Один параграф посвящен выяснению достаточных условий локального минимума функции $\varphi(X)$.

Все необходимые понятия, обозначения и свойства, относящиеся к n -мерному евклидову пространству и функциям n переменных, можно найти в Приложениях II и III.

§ 2. Свойства функции максимума

1. Рассмотрим функции $f_i(X)$, $i \in [0 : N]$, $X \in E_n$, и пусть

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X).$$

Отметим простейшие свойства функции $\varphi(X)$.

I. Если все функции $f_i(X)$, $i \in [0 : N]$, непрерывны в точке X_0 , то и функция $\varphi(X)$ непрерывна в той же точке.

Справедливость этого утверждения следует из неравенства (см. Приложение III, § 2)

$$\begin{aligned} |\varphi(X) - \varphi(X_0)| &\stackrel{df}{=} \max_{i \in [0 : N]} f_i(X) - \max_{i \in [0 : N]} f_i(X_0) \leq \\ &\leq \max_{i \in [0 : N]} |f_i(X) - f_i(X_0)|. \end{aligned}$$

II. Если все функции $f_i(X)$, $i \in [0 : N]$, непрерывны на E_n и при некотором $X_0 \in E_n$ множество

$$M(X_0) = \{X \in E_n \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено, то существует точка X^* , в которой функция $\varphi(X)$ достигает своего минимального значения.

Доказательство. В силу свойства I функция $\varphi(X)$ непрерывна на E_n . Отсюда, в частности, следует, что множество $M(X_0)$ замкнуто. Поскольку по предположению $M(X_0)$ — ограниченное множество, то найдется точка X^* , для которой

$$\varphi(X^*) = \min_{X \in M(X_0)} \varphi(X).$$

Осталось заметить, что

$$\min_{X \in M(X_0)} \varphi(X) = \inf_{X \in E_n} \varphi(X),$$

поэтому

$$\varphi(X^*) = \inf_{X \in E_n} \varphi(X).$$

Утверждение доказано.

III. Пусть $\Omega \subset E_n$ — некоторое выпуклое множество. Если все функции $f_i(X)$, $i \in [0 : N]$, выпуклы на Ω , то функция $\varphi(X)$ также выпукла на Ω .

Доказательство. Пусть $X_1, X_2 \in \Omega$ и $\alpha \in [0, 1]$. Тогда

$$f_i(\alpha X_1 + (1 - \alpha) X_2) \leq \alpha f_i(X_1) + (1 - \alpha) f_i(X_2) \leq \leq \alpha \varphi(X_1) + (1 - \alpha) \varphi(X_2). \quad (2.1)$$

Поскольку эти неравенства выполняются для всех $i \in [0 : N]$, то

$$\varphi(\alpha X_1 + (1 - \alpha) X_2) \leq \alpha \varphi(X_1) + (1 - \alpha) \varphi(X_2), \quad (2.2)$$

что и требовалось доказать.

2. При фиксированном X рассмотрим множество индексов $R(X)$:

$$R(X) = \{i \in [0 : N] \mid f_i(X) = \varphi(X)\}.$$

Таким образом, в $R(X)$ входят те индексы i , при которых $f_i(X)$ достигает максимального значения, равного $\varphi(X)$.

Лемма 2.1. Если функции $f_i(X)$, $i \in [0 : N]$, непрерывны в точке X_0 , то найдется число $\alpha_0 > 0$ такое, что для $\alpha \in [0, \alpha_0]$ и $g \in E_n$, $\|g\| = 1$, будет выполняться равенство

$$\varphi(X_0 + \alpha g) \stackrel{df}{=} \max_{i \in [0 : N]} f_i(X_0 + \alpha g) = \max_{i \in R(X_0)} f_i(X_0 + \alpha g). \quad (2.3)$$

Доказательство. Достаточно рассмотреть случай, когда $R(X_0) \neq [0 : N]$. Положим

$$\beta = \varphi(X_0) - \max_{i \notin R(X_0)} f_i(X_0).$$

Ясно, что $\beta > 0$. По непрерывности функций $f_i(X)$ в точке X_0 найдется такое $\alpha_0 > 0$, что для всех $g \in E_n$, $\|g\| = 1$, при $\alpha \in [0, \alpha_0]$ будет

$$f_i(X_0 + \alpha g) \geq \varphi(X_0) - \frac{1}{3}\beta \quad \text{для } i \in R(X_0),$$

$$f_i(X_0 + \alpha g) \leq \varphi(X_0) - \frac{2}{3}\beta \quad \text{для } i \notin R(X_0).$$

Отсюда очевидным образом следует (2.3).

Лемма доказана.

3. Определение. Функция $\varphi(X)$, $X \in E_n$, называется дифференцируемой в точке X_0 по направлению $g \in E_n$, $\|g\|=1$, если существует конечный предел

$$\lim_{\alpha \rightarrow +0} \frac{\varphi(X_0 + \alpha g) - \varphi(X_0)}{\alpha}.$$

Этот предел называется производной функции $\varphi(X)$ в точке X_0 по направлению g и обозначается $\frac{\partial \varphi(X_0)}{\partial g}$.

Теорема 2.1. Пусть функции $f_i(X)$, $i \in [0 : N]$, являются непрерывно дифференцируемыми в некоторой окрестности $S_\delta(X_0)$ точки X_0 ,

$$S_\delta(X_0) = \{X \mid \|X - X_0\| < \delta\}, \quad \delta > 0.$$

Тогда функция максимума $\varphi(X)$ дифференцируема в точке X_0 по любому направлению g , $\|g\|=1$, причем

$$\frac{\partial \varphi(X_0)}{\partial g} = \max_{i \in R(X_0)} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) \quad (2.4)$$

(внешние круглые скобки в правой части равенства обозначают скалярное произведение в E_n).

Доказательство. Для любого $i \in [0 : N]$, $\alpha \in (0, \delta)$ и g , $\|g\|=1$, имеем (см. Приложение III, § 3)

$$f_i(X_0 + \alpha g) = f_i(X_0) + \alpha \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) + o_i(g; \alpha), \quad (2.5)$$

где $o_i(g; \alpha)$ — величина, удовлетворяющая соотношению

$$\frac{o_i(g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$$

равномерно по g , $\|g\|=1$.

В силу леммы 2.1 найдется такое $\alpha_0 > 0$, $\alpha_0 < \delta$, что для $\alpha \in (0, \alpha_0)$ будет

$$\begin{aligned} \varphi(X_0 + \alpha g) &= \max_{i \in R(X_0)} \left\{ f_i(X_0) + \alpha \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) + o_i(g; \alpha) \right\} \leq \\ &\leq \varphi(X_0) + \alpha \max_{i \in R(X_0)} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) + \max_{i \in [0 : N]} o_i(g; \alpha). \end{aligned}$$

Воспользуемся теперь следующим числовым неравенством (см. Приложение III, § 2):

$$\max_{i \in [0 : N]} (\beta_i + \gamma_i) \geq \max_{i \in [0 : N]} \beta_i + \max_{i \in R} \gamma_i, \quad (2.6)$$

где $R = \{i \mid \beta_i = \max_{k \in [0 : N]} \beta_k\}$.

Получим, учитывая (2.5),

$$\begin{aligned} \varphi(X_0 + \alpha g) &\geq \varphi(X_0) + \max_{i \in R(X_0)} \left\{ \alpha \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) + o_i(g; \alpha) \right\} \geq \\ &\geq \varphi(X_0) + \alpha \max_{i \in R(X_0)} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) + \min_{i \in [0 : N]} o_i(g; \alpha). \end{aligned}$$

Итак, для $\alpha \in (0, \alpha_0)$ и любого $g, \|g\| = 1$, справедливы неравенства

$$\begin{aligned} \min_{i \in [0 : N]} o_i(g; \alpha) &\leq \\ &\leq \varphi(X_0 + \alpha g) - \varphi(X_0) - \alpha \max_{i \in R(X_0)} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) \leq \\ &\leq \max_{i \in [0 : N]} o_i(g; \alpha). \end{aligned}$$

Деля эти неравенства на $\alpha > 0$ и переходя к пределу при $\alpha \rightarrow +0$, получим требуемое. Более того, можно считать доказанным следующее равенство:

$$\varphi(X_0 + \alpha g) = \varphi(X_0) + \alpha \frac{\partial \varphi(X_0)}{\partial g} + o(g; \alpha), \quad (2.7)$$

где $o(g; \alpha)$ — величина, удовлетворяющая соотношению

$$\frac{o(g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$$

равномерно по $g, \|g\| = 1$.

4. Пример. Пусть (см. рис. 5)

$$f_0(x) = \sin x, \quad f_1(x) = \cos x,$$

$$\bar{\varphi}(x) = \max \{ \sin x, \cos x \}.$$

Поскольку функция $\varphi(x)$ 2π -периодична, будем рассматривать $x \in [0, 2\pi]$. Заметим, во-первых, что

$$R(x) = \begin{cases} \{1\}, & \text{если } x \in \left[0, \frac{\pi}{4}\right) \text{ или } x \in \left(\frac{5\pi}{4}, 2\pi\right]; \\ \{0, 1\}, & \text{если } x = \frac{\pi}{4} \text{ или } x = \frac{5\pi}{4}; \\ \{0\}, & \text{если } x \in \left(\frac{\pi}{4}, \frac{5\pi}{4}\right). \end{cases}$$

Функция $\varphi(x)$ непрерывна, однако она не является непрерывно дифференцируемой. Первая производная функции $\varphi(x)$ терпит разрыв в точках $x = \frac{\pi}{4}$ и $x = \frac{5\pi}{4}$.

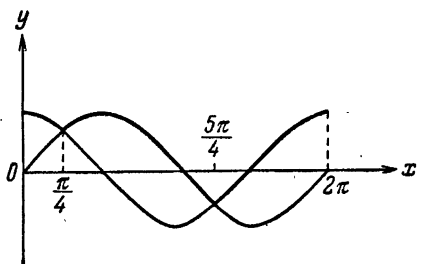


Рис. 5.

Тем не менее даже в этих точках $\varphi(x)$ имеет производные по двум возможным направлениям $g = (+1)$ и $g = (-1)$.

Возьмем, например, $x = \frac{\pi}{4}$. При $g = (-1)$ получим на основании (2.4)

$$\begin{aligned} \frac{\partial \varphi\left(\frac{\pi}{4}\right)}{\partial g} &= \max_{i \in R\left(\frac{\pi}{4}\right)} \left(\frac{df_i\left(\frac{\pi}{4}\right)}{dx}, g \right) = \\ &= \max \left\{ -\cos \frac{\pi}{4}, \sin \frac{\pi}{4} \right\} = \frac{\sqrt{2}}{2}. \end{aligned}$$

Аналогично при $g = (+1)$

$$\frac{\partial \varphi\left(\frac{\pi}{4}\right)}{\partial g} = \frac{\sqrt{2}}{2}.$$

Заметим, что в точке $x = \frac{\pi}{4}$ функция $\varphi(x)$ возрастает в обоих возможных направлениях.

Легко проверить, что в общем случае для $g = (+1)$ и $g = (-1)$

$$\frac{\partial \varphi(x)}{\partial g} = \begin{cases} -g \sin x, & \text{если } x \in \left[0, \frac{\pi}{4}\right) \text{ или } x \in \left(\frac{5\pi}{4}, 2\pi\right]; \\ \max \{g \cos x, -g \sin x\}, & \text{если } x = \frac{\pi}{4} \text{ или } x = \frac{5\pi}{4}; \\ g \cos x, & \text{если } x \in \left(\frac{\pi}{4}, \frac{5\pi}{4}\right). \end{cases}$$

5*. В заключение этого параграфа получим обобщение формулы (2.7).

Пусть функции $f_i(X)$, $i \in [0 : N]$, $X = (x_1, x_2, \dots, x_n)$, непрерывны и имеют непрерывные производные до l -го порядка включительно в некоторой окрестности $S_\delta(X_0)$, $\delta > 0$, точки X_0 . Заметим, что в этом случае для любого $\alpha \in (0, \delta)$ и $g = (g_1, g_2, \dots, g_n)$, $\|g\| = 1$, справедливо равенство (см. Приложение III, § 3)

$$f_i(X_0 + \alpha g) = f_i(X_0) + \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} + o_i(g; \alpha^l), \quad (2.8)$$

где

$$\frac{\partial^k f_i(X_0)}{\partial g^k} = \sum_{j_1, \dots, j_k=1}^n \frac{\partial^k f_i(X_0)}{\partial x_{j_1} \dots \partial x_{j_k}} g_{j_1} \dots g_{j_k}$$

и

$$\frac{o_i(g; \alpha^l)}{\alpha^l} \xrightarrow{\alpha \rightarrow 0} 0$$

равномерно по g , $\|g\| = 1$.

Введем обозначения:

$$\frac{\partial^0 f_i(X)}{\partial g^0} = f_i(X), \quad i \in [0 : N];$$

$$R_0(X, g) = [0 : N];$$

$$R_k(X, g) = \left\{ i \mid i \in R_{k-1}(X, g), \right.$$

$$\left. \frac{\partial^{k-1} f_i(X)}{\partial g^{k-1}} = \max_{j \in R_{k-1}(X, g)} \frac{\partial^{k-1} f_j(X)}{\partial g^{k-1}} \right\}, \quad k \in [1 : l].$$

Так как $R_0(X, g)$ не зависит от X и g , а $R_1(X, g)$ не зависит от g , то будем писать

$$R_0(X, g) = R_0, \quad R_1(X, g) = R(X).$$

Очевидно,

$$R_0 \supset R(X) \supset R_2(X, g) \supset \dots$$

Теорема 2.2. Пусть функции $f_i(X)$, $i \in [0 : N]$, непрерывны и имеют непрерывные производные до l -го порядка включительно в некоторой окрестности $S_\delta(X_0)$ точки X_0 . Тогда справедливо следующее представление функции максимума $\varphi(X)$ в окрестности точки X_0 в направлении g , $\|g\| = 1$,

$$\varphi(X_0 + \alpha g) = \varphi(X_0) + \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k \varphi(X_0)}{\partial g^k} + o(g; \alpha^l),$$

где

$$\frac{\partial^k \varphi(X_0)}{\partial g^k} = \max_{i \in R_k(X_0, g)} \frac{\partial^k f_i(X_0)}{\partial g^k}$$

и

$$\frac{o(g; \alpha^l)}{\alpha^l} \xrightarrow{\alpha \rightarrow +0} 0$$

равномерно по g , $\|g\| = 1$.

Доказательство. Так же, как при доказательстве теоремы 2.1, найдем число $\alpha_0 > 0$, $\alpha_0 < \delta$, такое, что для $\alpha \in (0, \alpha_0)$ в силу (2.8)

$$\varphi(X_0 + \alpha g) \leq$$

$$\leq \varphi(X_0) + \max_{i \in R(X_0)} \left\{ \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\} + \max_{i \in R_0} o_i(g; \alpha^l). \quad (2.9)$$

Нетрудно показать, что существует столь малое $\alpha_1 > 0$, $\alpha_1 \leq \alpha_0$, что для $\alpha \in (0, \alpha_1)$ будет

$$\begin{aligned} \max_{i \in R(X_0)} \left\{ \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\} &= \\ &= \max_{i \in R(X_0)} \left\{ \frac{\alpha}{1!} \frac{\partial f_i(X_0)}{\partial g} + \sum_{k=2}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\} = \end{aligned}$$

$$\begin{aligned}
&= \max_{i \in R_2(X_0, g)} \left\{ \frac{\alpha}{1!} \frac{\partial f_i(X_0)}{\partial g} + \sum_{k=2}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\} = \\
&= \frac{\alpha}{1!} \max_{i \in R(X_0)} \frac{\partial f_i(X_0)}{\partial g} + \max_{i \in R_2(X_0, g)} \left\{ \sum_{k=2}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\}.
\end{aligned}$$

Продолжая аналогично, найдем такое $\alpha_{l-1} > 0$, $\alpha_{l-1} \leq \alpha_{l-2}$, что для $\alpha \in (0, \alpha_{l-1})$

$$\max_{i \in R(X_0)} \left\{ \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\} = \sum_{k=1}^l \frac{\alpha^k}{k!} \max_{i \in R_k(X_0, g)} \frac{\partial^k f_i(X_0)}{\partial g^k}.$$

Отсюда и из (2.9) следует

$$\begin{aligned}
\varphi(X_0 + \alpha g) &\leq \varphi(X_0) + \sum_{k=1}^l \frac{\alpha^k}{k!} \max_{i \in R_k(X_0, g)} \frac{\partial^k f_i(X_0)}{\partial g^k} + \\
&\quad + \max_{i \in R_0} o_i(g; \alpha^l).
\end{aligned}$$

С другой стороны, применяя последовательно неравенство (2.6), получим

$$\begin{aligned}
\varphi(X_0 + \alpha g) &\geq \varphi(X_0) + \max_{i \in R(X_0)} \left\{ \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k f_i(X_0)}{\partial g^k} \right\} + \\
&\quad + \min_{i \in R_0} o_i(g; \alpha^l) \geq \varphi(X_0) + \\
&\quad + \sum_{k=1}^l \frac{\alpha^k}{k!} \max_{i \in R_k(X_0, g)} \frac{\partial^k f_i(X_0)}{\partial g^k} + \min_{i \in R_0} o_i(g; \alpha^l).
\end{aligned}$$

Итак, для $\alpha \in (0, \alpha_{l-1})$ и любого g , $\|g\| = 1$, справедливы неравенства

$$\min_{i \in R_0} o_i(g; \alpha^l) \leq$$

$$\leq \varphi(X_0 + \alpha g) - \varphi(X_0) - \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k \varphi(X_0)}{\partial g^k} \leq \max_{i \in R_0} o_i(g; \alpha^l).$$

Отсюда очевидным образом следует утверждение теоремы. Теорема доказана.

§ 3. Необходимые условия минимакса

1. В дальнейшем будем предполагать, что функции $f_i(X)$, $i \in [0:N]$, непрерывно дифференцируемы на E_n и что при некотором X_0 множество

$$M(X_0) = \{X \in E_n \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено. Пусть, как и раньше,

$$\varphi(X) = \max_{i \in [0:N]} f_i(X).$$

Теорема 3.1. *Для того чтобы точка X^* была точкой минимума функции $\varphi(X)$ на E_n , необходимо, а в случае выпуклости $\varphi(X)$ и достаточно, чтобы выполнялось неравенство*

$$\inf_{\|g\|=1} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) \geq 0, \quad (3.1)$$

или, что то же самое (см. (2.4)),

$$\inf_{\|g\|=1} \frac{\partial \varphi(X^*)}{\partial g} \geq 0. \quad (3.2)$$

Доказательство. Необходимость. Предположим, что точка X^* есть точка минимума функции $\varphi(X)$, однако неравенство (3.2) не выполняется. Тогда найдется такой вектор $g_1 \in E_n$, $\|g_1\| = 1$, что

$$\frac{\partial \varphi(X^*)}{\partial g_1} = -a < 0. \quad (3.3)$$

В силу (2.7)

$$\varphi(X^* + \alpha g_1) = \varphi(X^*) + \alpha \frac{\partial \varphi(X^*)}{\partial g_1} + o(g_1; \alpha), \quad (3.4)$$

где

$$\frac{o(g_1; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0.$$

Зафиксируем достаточно малое $\alpha_1 > 0$, при котором

$$|o(g_1; \alpha_1)| \leq \frac{a}{2} \alpha_1.$$

Тогда, учитывая (3.4) и (3.3), получим

$$\varphi(X^* + \alpha_1 g_1) \leq \varphi(X^*) - \frac{a}{2} \alpha_1,$$

что противоречит предположению о том, что X^* — точка минимума функции $\varphi(X)$.

Достаточность. Пусть $\varphi(X)$ — выпуклая функция и в точке X^* выполняется неравенство (3.2). Покажем, что в этом случае X^* — точка минимума функции $\varphi(X)$ на E_n . Предположим противное. Тогда найдется точка X_0 , в которой $\varphi(X_0) < \varphi(X^*)$. В силу выпуклости $\varphi(X)$ имеем для $\alpha \in [0, 1]$

$$\begin{aligned}\varphi(X^* + \alpha(X_0 - X^*)) &= \varphi((1 - \alpha)X^* + \alpha X_0) \leq \\ &\leq (1 - \alpha)\varphi(X^*) + \alpha\varphi(X_0).\end{aligned}\quad (3.5)$$

Положим $g_1 = \frac{X_0 - X^*}{\|X_0 - X^*\|}$, $\|g_1\| = 1$, и найдем $\frac{\partial \varphi(X^*)}{\partial g_1}$,

$$\frac{\partial \varphi(X^*)}{\partial g_1} = \lim_{\alpha \rightarrow +0} \frac{1}{\alpha} \left[\varphi \left(X^* + \frac{\alpha}{\|X_0 - X^*\|} (X_0 - X^*) \right) - \varphi(X^*) \right].$$

Учитывая (3.5), получим для $0 < \alpha < \|X_0 - X^*\|$

$$\begin{aligned}\varphi \left(X^* + \frac{\alpha}{\|X_0 - X^*\|} (X_0 - X^*) \right) - \varphi(X^*) &\leq \\ &\leq \left(1 - \frac{\alpha}{\|X_0 - X^*\|} \right) \varphi(X^*) + \frac{\alpha}{\|X_0 - X^*\|} \varphi(X_0) - \varphi(X^*) = \\ &= \frac{\alpha}{\|X_0 - X^*\|} (\varphi(X_0) - \varphi(X^*)).\end{aligned}$$

Отсюда

$$\frac{\partial \varphi(X^*)}{\partial g_1} \leq \frac{\varphi(X_0) - \varphi(X^*)}{\|X_0 - X^*\|} < 0,$$

что противоречит (3.2).

Теорема доказана.

Точка X^* , для которой выполняются неравенства (3.1) и (3.2), называется *стационарной точкой* функции максимума $\varphi(X)$ на E_n .

2. Дадим геометрическую интерпретацию необходимого условия минимакса (3.1). Для этого при фиксированном $X \in E_n$ образуем множество

$$H(X) = \left\{ Z \in E_n \mid Z = \frac{\partial f_i(X)}{\partial X}, i \in R(X) \right\}.$$

Обозначим через $L(X)$ выпуклую оболочку множества $H(X)$ (см. Приложение II, § 1)

$$L(X) = \left\{ Z = \sum_{i \in R(X)} \alpha_i \frac{\partial f_i(X)}{\partial X} \mid \alpha_i \geq 0, \sum_{i \in R(X)} \alpha_i = 1 \right\}.$$

Заметим, что $L(X)$ является ограниченным замкнутым выпуклым множеством.

Теорема 3.2. *Нервенство (3.1) равносильно включению*

$$0 \in L(X^*). \quad (3.6)$$

Доказательство. Покажем вначале, что из (3.1) следует (3.6). Допустим противное: $0 \notin L(X^*)$. Тогда по теореме отделимости (см. Приложение II, § 1) найдутся вектор g_0 , $\|g_0\| = 1$, и число $a > 0$ такие, что для всех $Z \in L(X^*)$ будет

$$(Z, g_0) \leq -a < 0.$$

Так как $H(X^*) \subset L(X^*)$, то из последнего неравенства получаем

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) \leq -a < 0,$$

что противоречит (3.1).

Покажем теперь, что из (3.6) следует (3.1). Допустим противное. Тогда найдется вектор g_0 , $\|g_0\| = 1$, и число $a > 0$ такие, что для всех $i \in R(X^*)$

$$\left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) \leq -a < 0. \quad (3.7)$$

Возьмем произвольный вектор $Z \in L(X^*)$. В силу (3.7) и определения множества $L(X^*)$ получим

$$\begin{aligned} (Z, g_0) &= \left(\sum_{i \in R(X^*)} \alpha_i \frac{\partial f_i(X^*)}{\partial X}, g_0 \right) = \\ &= \sum_{i \in R(X^*)} \alpha_i \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) \leq -a \sum_{i \in R(X^*)} \alpha_i = -a < 0. \end{aligned}$$

Но тогда точка $Z = 0$ не может принадлежать $L(X^*)$, что противоречит (3.6).

Теорема доказана.

Следствие. Для того чтобы функция $\varphi(X)$ достигала минимума в точке X^ , необходимо, а в случае выпуклости $\varphi(X)$ и достаточно, чтобы*

$$0 \in L(X^*).$$

З а м е ч а н и е. Условие (3.6) является обобщением классического необходимого условия минимума непрерывно дифференцируемой функции. Последнее получается из (3.6), если положить $N=0$. Действительно, в этом случае

$$\varphi(X) = f_0(X).$$

Множества $H(X)$ и $L(X)$ состоят из единственной точки $\left\{ \frac{\partial f_0(X)}{\partial X} \right\}$. Включение (3.6) заменяется равенством

$$\frac{\partial f_0(X^*)}{\partial X} = 0. \quad (3.8)$$

Таким образом, для того чтобы непрерывно дифференцируемая функция $f_0(X)$ достигала минимума на E_n в точке X^* , необходимо, а в случае выпуклости $f_0(X)$ и достаточно, чтобы выполнялось равенство (3.8).

3. При фиксированном $X \in E_n$ введем функцию

$$\chi(g) = \max_{i \in R(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right).$$

В силу общих свойств функции максимума (см. § 2, п. 1) $\chi(g)$ является непрерывной и выпуклой функцией на всем пространстве E_n . Кроме того, очевидно, что для любого $\lambda \geq 0$

$$\chi(\lambda g) = \lambda \chi(g). \quad (3.9)$$

Нас интересует $\chi(g)$ при $g \in S^0$, где

$$S^0 = \{g \in E_n \mid \|g\| = 1\}.$$

Поскольку S^0 — ограниченное замкнутое множество, то $\chi(g)$ достигает минимума на S^0 . Отсюда, в частности, следует, что инфимум, стоящий в левой части неравенства (3.1), достигается, и поэтому неравенство (3.1) можно переписать в виде

$$\min_{\|g\|=1} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) \geq 0. \quad (3.1')$$

Л е м м а 3.1. *Справедливо равенство*

$$\chi(g) = \max_{Z \in L(X)} (Z, g),$$

где $L(X)$ — множество, определенное в предыдущем пункте.

Доказательство. Очевидно, что

$$\chi(g) = \max_{Z \in H(X)} (Z, g). \quad (3.10)$$

Поскольку $H(X) \subset L(X)$, то

$$\chi(g) \leq \max_{Z \in L(X)} (Z, g). \quad (3.11)$$

С другой стороны, любая точка $Z' \in L(X)$ может быть представлена в виде

$$Z' = \sum_{i \in R(X)} \alpha'_i Z_i; \quad Z_i \in H(X); \quad \alpha'_i \geq 0; \quad \sum_{i \in R(X)} \alpha'_i = 1.$$

Поэтому

$$\begin{aligned} (Z', g) &= \sum_{i \in R(X)} \alpha'_i (Z_i, g) \leq \max_{Z \in H(X)} (Z, g) \sum_{i \in R(X)} \alpha'_i = \\ &= \max_{Z \in H(X)} (Z, g). \end{aligned}$$

Так как это неравенство справедливо для произвольного $Z' \in L(X)$, то

$$\max_{Z \in L(X)} (Z, g) \leq \max_{Z \in H(X)} (Z, g).$$

Учитывая (3.10), получим

$$\max_{Z \in L(X)} (Z, g) \leq \chi(g). \quad (3.12)$$

Утверждение леммы следует теперь из (3.11) и (3.12).
Лемма доказана.

4. Введем функцию

$$\psi(X) = \min_{\|g\|=1} \max_{i \in R(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right).$$

С помощью функции $\psi(X)$ необходимое условие минимакса (3.1) можно переписать в виде

$$\psi(X^*) \geq 0. \quad (3.1'')$$

Пусть $\chi(X, g) = \chi(g)$. В силу леммы 3.1 получим

$$\psi(X) = \min_{\|g\|=1} \chi(X, g) = \min_{\|g\|=1} \max_{Z \in L(X)} (Z, g). \quad (3.13)$$

Лемма 3.2. Если $\psi(X^*) \geq 0$, т. е. X^* является стационарной точкой функции максимума $\phi(X)$, то

$$\psi(X^*) = r(X^*),$$

где $r(X^*)$ — радиус максимального шара с центром в начале координат, который можно вписать в $L(X^*)$ (рис. 6).

Доказательство. Обозначим через $\bar{S}_\delta$ шар радиуса $\delta \geq 0$ с центром в начале координат:

$$\bar{S}_\delta = \{Z \in E_n \mid \|Z\| \leq \delta\}.$$

Если $\bar{S}_\delta \subset L(X^*)$, то для любого g , $\|g\|=1$, получим

$$\delta = \max_{Z \in \bar{S}_\delta} (Z, g) \leq \max_{Z \in L(X^*)} (Z, g).$$

Отсюда и из (3.13) следует, что $\delta \leq \psi(X^*)$. В частности,

$$r(X^*) \leq \psi(X^*). \quad (3.14)$$

Докажем теперь, что

$$\bar{S}_{\psi(X^*)} \subset L(X^*). \quad (3.15)$$

Допустим противное. Тогда найдется точка $Z_0 \in \bar{S}_{\psi(X^*)}$,

$$\|Z_0\| \leq \psi(X^*), \quad (3.16)$$

не принадлежащая $L(X^*)$. По теореме отделимости (см. Приложение II, § 1) существуют вектор g_0 , $\|g_0\|=1$, и число $a > 0$ такие, что для любого $Z \in L(X^*)$

$$(Z - Z_0, g_0) \leq -a.$$

Отсюда, учитывая (3.16), получим

$$\max_{Z \in L(X^*)} (Z, g_0) \leq (Z_0, g_0) - a \leq \psi(X^*) - a,$$

что противоречит (3.13).

Итак, соотношение (3.15), а с ним и неравенство

$$\psi(X^*) \leq r(X^*), \quad (3.17)$$

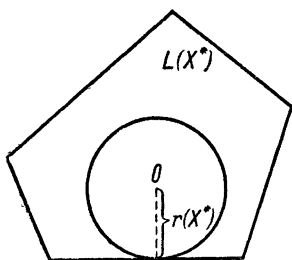


Рис. 6.

установлены. Утверждение леммы следует теперь из (3.14) и (3.17).

Лемма доказана.

5. Обратимся к тем $X \in E_n$, для которых $\psi(X) < 0$. В этом случае X не является стационарной точкой, и по теореме 3.2 $0 \notin L(X)$.

Лемма 3.3. Если $\psi(X) < 0$, то

$$\psi(X) = \max_{Z \in L(X)} \left(Z, -\frac{Z^*}{\|Z^*\|} \right) = -\|Z^*\|,$$

где Z^* есть ближайшая к началу координат точка множества $L(X)$.

Доказательство. Для любого $Z \in L(X)$ справедливо неравенство (см. Приложение II, лемма 1.3)

$$(Z, Z^*) \geq (Z^*, Z^*).$$

$$\text{Положим } \bar{g} = -\frac{Z^*}{\|Z^*\|}, \quad \|\bar{g}\| = 1.$$

Имеем для любого $Z \in L(X)$

$$(Z, \bar{g}) = -\frac{1}{\|Z^*\|} (Z, Z^*) \leq -\frac{1}{\|Z^*\|} (Z^*, Z^*) = -\|Z^*\|.$$

Отсюда следует, что

$$\max_{Z \in L(X)} (Z, \bar{g}) = -\|Z^*\|. \quad (3.18)$$

Далее, для любого g , $\|g\| = 1$,

$$-\|Z^*\| \leq (Z^*, g) \leq \max_{Z \in L(X)} (Z, g).$$

Значит, для любого g , $\|g\| = 1$, в силу (3.18)

$$\max_{Z \in L(X)} (Z, \bar{g}) \leq \max_{Z \in L(X)} (Z, g),$$

т. е.

$$\max_{Z \in L(X)} (Z, \bar{g}) = \min_{\|g\|=1} \max_{Z \in L(X)} (Z, g) = \psi(X). \quad (3.19)$$

Объединяя (3.18) и (3.19) и учитывая определение вектора $\bar{g}$, получим требуемое.

Лемма доказана.

Равенство (3.19) можно понимать следующим образом (см. (3.13)): функция $\chi(X, g)$ при фиксированном $X \in E_n$, $\psi(X) < 0$, достигает минимума на единичной сфере S^0 при $g = \bar{g}$. Ниже это утверждение уточняется.

Лемма 3.4. Если $\psi(X) < 0$, то функция $\chi(X, g)$ достигает минимума по $g \in S^0$ в единственной точке.

Доказательство. Поскольку X фиксировано, будем писать $\chi(X, g) = \chi(g)$. Предположим, что существуют два вектора $g_1 \neq g_2$, $\|g_1\| = \|g_2\| = 1$, такие, что

$$\chi(g_1) = \chi(g_2) = \min_{\|g\|=1} \chi(g) = \psi(X).$$

Заметим, что в этом случае

$$(g_1, g_2) < 1. \quad (3.20)$$

В силу выпуклости функции $\chi(g)$ на всем E_n (см. п. 3) имеем для любого $\alpha \in (0, 1)$

$$\chi(\alpha g_1 + (1 - \alpha) g_2) \leq \alpha \chi(g_1) + (1 - \alpha) \chi(g_2) = \psi(X). \quad (3.21)$$

Зафиксируем $\alpha_0 \in (0, 1)$ так, чтобы

$$\beta \stackrel{\text{df}}{=} \|\alpha_0 g_1 + (1 - \alpha_0) g_2\| > 0.$$

Имеем в силу (3.20)

$$\begin{aligned} \beta^2 &= \|\alpha_0 g_1 + (1 - \alpha_0) g_2\|^2 = \alpha_0^2 + 2\alpha_0(1 - \alpha_0)(g_1, g_2) + \\ &\quad + (1 - \alpha_0)^2 < \alpha_0^2 + 2\alpha_0(1 - \alpha_0) + (1 - \alpha_0)^2 = 1. \end{aligned}$$

Итак, $0 < \beta < 1$. Положим

$$g_0 = \frac{1}{\beta} (\alpha_0 g_1 + (1 - \alpha_0) g_2), \quad \|g_0\| = 1.$$

Поскольку $\frac{1}{\beta} > 1$ и $\psi(X) < 0$, то получим в силу 3.9) и (3.21)

$$\chi(g_0) = \frac{1}{\beta} \chi(\alpha_0 g_1 + (1 - \alpha_0) g_2) \leq \frac{1}{\beta} \psi(X) < \psi(X),$$

что невозможно. Лемма доказана.

6. Определение. Вектор $g(\bar{X})$, $\|g(\bar{X})\| = 1$, будем называть направлением наискорейшего спуска функции максимума $\varphi(X)$ в точке $\bar{X}$, если

$$\frac{\partial \varphi(\bar{X})}{\partial g(\bar{X})} = \min_{\|g\|=1} \frac{\partial \varphi(\bar{X})}{\partial g}.$$

Вектор $g_1(\bar{X})$, $\|g_1(\bar{X})\|=1$, будем называть *направлением наискорейшего подъема функции максимума $\varphi(X)$ в точке $\bar{X}$* , если

$$\frac{\partial \varphi(\bar{X})}{\partial g_1(\bar{X})} = \max_{\|g\|=1} \frac{\partial \varphi(\bar{X})}{\partial g}.$$

Теорема 3.3. Если $\psi(\bar{X}) < 0$, то функция максимума $\varphi(X)$ имеет в точке $\bar{X}$ единственное направление наискорейшего спуска $g(\bar{X})$, $\|g(\bar{X})\|=1$. При этом

$$g(\bar{X}) = - \frac{Z^*}{\|Z^*\|}, \quad \frac{\partial \varphi(\bar{X})}{\partial g(\bar{X})} = -\|Z^*\|,$$

где Z^* — ближайшая к началу координат точка множества $L(\bar{X})$.

Эта теорема по существу уже доказана. Нужно заметить, что

$$\frac{\partial \varphi(\bar{X})}{\partial g} = \chi(\bar{X}, g),$$

и воспользоваться результатами лемм 3.3 и 3.4.

Замечание. Пусть $N=0$, т. е.

$$\varphi(X) = f_0(X).$$

В этом случае при всех $X \in E_n$

$$L(X) = \left\{ \frac{\partial f_0(X)}{\partial X} \right\}.$$

Если $\psi(\bar{X}) < 0$, т. е. $\frac{\partial f(\bar{X})}{\partial X} \neq 0$, то направлением наискорейшего спуска функции $f_0(X)$ в точке $\bar{X}$ по теореме 3.3 будет вектор

$$g(\bar{X}) = - \frac{\partial f_0(\bar{X})}{\partial X} / \left\| \frac{\partial f_0(\bar{X})}{\partial X} \right\|.$$

Очевидно, что направлением наискорейшего подъема функции $f_0(X)$ в точке $\bar{X}$ является вектор

$$g_1(\bar{X}) = \frac{\partial f_0(\bar{X})}{\partial X} / \left\| \frac{\partial f_0(\bar{X})}{\partial X} \right\|.$$

Таким образом, для непрерывно дифференцируемой функции $f_0(X)$ во всякой точке $\bar{X}$, в которой $\frac{\partial f_0(\bar{X})}{\partial X} \neq 0$, выполняется соотношение

$$g_1(\bar{X}) = -g(\bar{X}). \quad (3.22)$$

Заметим, что в общем случае при $N > 0$ равенство (3.22) может нарушаться. Более того, в общем случае направление наискорейшего подъема может быть не единственным.

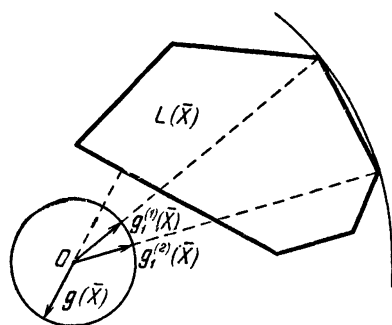


Рис. 7.

В ситуации, изображенной на рис. 7, $g(\bar{X})$ — направление наискорейшего спуска, $g_1^{(1)}(\bar{X})$ и $g_1^{(2)}(\bar{X})$ — направления наискорейшего подъема.

§ 4. Достаточные условия локального минимакса. Некоторые оценки

1. В предыдущем параграфе было показано, что если функция максимума $\varphi(X)$ выпукла, то стационарная точка X^* является в то же время точкой минимума функции $\varphi(X)$ на E_n . Теперь нас интересуют условия, при выполнении которых стационарная точка X^* является точкой локального минимума функции $\varphi(X)$.

Как и раньше, считаем, что функции $f_i(X)$, $i \in [0 : N]$, непрерывно дифференцируемы на E_n ; пусть

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X),$$

$$\psi(X) = \min_{\|g\|=1} \max_{i \in R(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right).$$

2. **Определение.** Точка $X^* \in E_n$ называется *точкой строгого локального минимума функции* $\varphi(X)$, если существует окрестность $S_\delta(X^*)$, $\delta > 0$, точки X^*

$$S_\delta(X^*) = \{X \in E_n \mid \|X - X^*\| < \delta\}$$

такая, что для всех $X \in S_\delta(X^*)$, $X \neq X^*$, выполняется неравенство

$$\varphi(X) > \varphi(X^*).$$

Теорема 4.1. Если $\psi(X^*) = r > 0$, то X^* является точкой строгого локального минимума функции $\varphi(X)$.

Доказательство. Как известно (см. (2.7)),

$$\varphi(X^* + \alpha g) = \varphi(X^*) + \alpha \frac{\partial \varphi(X^*)}{\partial g} + o(g; \alpha), \quad (4.1)$$

где $\frac{o(g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$ равномерно по g , $\|g\| = 1$.

По определению $\psi(X)$ имеем для всех g , $\|g\| = 1$,

$$\frac{\partial \varphi(X^*)}{\partial g} \geq r.$$

Выберем столь малое $\alpha_1 > 0$, чтобы для всех $\alpha \in (0, \alpha_1)$ и всех g , $\|g\| = 1$, выполнялось неравенство

$$|o(g; \alpha)| \leq \frac{r}{2} \alpha.$$

Получим в силу (4.1) для любого $X \in S_{\alpha_1}(X^*)$, $X \neq X^*$,

$$\varphi(X) = \varphi(X^* + \alpha g) \geq \varphi(X^*) + \frac{r}{2} \alpha > \varphi(X^*).$$

Теорема доказана.

З а м е ч а н и е. Если $\psi(X^*) = r > 0$, то найдется $\alpha_0 > 0$ такое, что для всех $X \in S_{\alpha_0}(X^*)$, $X \neq X^*$, индексное множество $R(X)$ строго включается в $R(X^*)$:

$$R(X) \subset R(X^*), \quad R(X) \neq R(X^*).$$

Действительно, в силу леммы 2.1 найдется такое $\alpha_1 > 0$, что для всех $X \in S_{\alpha_1}(X^*)$ будет $R(X) \subset R(X^*)$.

Заметим далее, что для любого g , $\|g\| = 1$, найдется индекс $i_1(g) \in R(X^*)$ такой, что

$$\left(\frac{\partial f_{i_1(g)}(X^*)}{\partial X}, g \right) \geq r. \quad (4.2)$$

Положим $i_2(g) = i_1(-g) \in R(X^*)$. В этом случае

$$\left(\frac{\partial f_{i_2(g)}(X^*)}{\partial X}, g \right) \leq -r. \quad (4.3)$$

Выпишем теперь соотношение

$$f_i(X^* + \alpha g) = f_i(X^*) + \alpha \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) + o_i(g; \alpha). \quad (4.4)$$

Требуемое α_0 , $0 < \alpha_0 \leq \alpha_1$, выберем так, чтобы для $\alpha \in (0, \alpha_0)$ выполнялось неравенство

$$|o_i(g; \alpha)| \leq \frac{r}{2} \alpha \quad (4.5)$$

равномерно по $i \in [0 : N]$ и g , $\|g\| = 1$.

В этом случае для любого $X \in S_{\alpha_0}(X^*)$, $X = X^* + \alpha g$, $X \neq X^*$, имеем в силу (4.4), (4.2), (4.3) и (4.5)

$$f_{i_1(g)}(X) \geq f_{i_1(g)}(X^*) + \frac{1}{2} r \alpha = \varphi(X^*) + \frac{1}{2} r \alpha;$$

$$f_{i_2(g)}(X) \leq f_{i_2(g)}(X^*) - \frac{1}{2} r \alpha = \varphi(X^*) - \frac{1}{2} r \alpha.$$

Таким образом, $f_{i_1(g)}(X) > f_{i_2(g)}(X)$. Отсюда следует, что $i_2(g) \notin R(X)$, хотя $i_2(g) \in R(X^*)$.

Утверждение доказано.

3*. Теорема 4.2. Пусть X^* — стационарная точка функции $\varphi(X)$ на E_n , причем $\psi(X^*) = 0$. Пусть далее функции $f_i(X)$, $i \in [0 : N]$, дважды непрерывно дифференцируемы в некоторой окрестности $S_\delta(X^*)$, $\delta > 0$, точки X^* .

Положим

$$\frac{\partial^2 f_i(X)}{\partial X^2} = \begin{bmatrix} \frac{\partial^2 f_i(X)}{\partial x_1 \partial x_1} & \frac{\partial^2 f_i(X)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f_i(X)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f_i(X)}{\partial x_2 \partial x_1} & \frac{\partial^2 f_i(X)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f_i(X)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f_i(X)}{\partial x_n \partial x_1} & \frac{\partial^2 f_i(X)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f_i(X)}{\partial x_n \partial x_n} \end{bmatrix}.$$

Если при некотором $\gamma > 0$

$$\min_{g \in G_\gamma} \max_{i \in R_2(X^*, g)} \left(\frac{\partial^2 f_i(X^*)}{\partial X^2} g, g \right) \stackrel{\text{def}}{=} a(X^*) > 0, \quad (4.6)$$

где

$$G_\gamma = \left\{ g \mid \|g\| = 1, 0 \leq \frac{\partial \varphi(X^*)}{\partial g} \leq \gamma \right\},$$

$$R_2(X^*, g) = \left\{ i \in R(X^*) \mid \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) = \frac{\partial \varphi(X^*)}{\partial g} \right\},$$

$$\left(\frac{\partial^2 f_i(X^*)}{\partial X^2} g, g \right) = \sum_{l_1, l_2=1}^n \frac{\partial^2 f_i(X^*)}{\partial x_{l_1} \partial x_{l_2}} g_{l_1} g_{l_2},$$

то X^* является точкой строгого локального минимума функции $\varphi(X)$.

Доказательство. Заметим, во-первых, что для всех $i \in [0:N]$, любых $g, \|g\|=1$, и $\alpha \in (0, \delta)$ справедливо равенство (см. Приложение III, § 3)

$$f_i(X^* + \alpha g) = f_i(X^*) + \alpha \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) + \frac{1}{2} \alpha^2 \left(\frac{\partial^2 f_i(X^* + \theta_i g)}{\partial X^2} g, g \right), \quad (4.7)$$

где $\theta_i \in (0, \alpha)$.

Далее из (4.6) следует, что для любого $g \in G_\gamma$ (множество G_γ не пусто, поскольку по условию $\psi(X^*)=0$) найдется индекс $i(g) \in R_2(X^*, g)$ такой, что

$$\left(\frac{\partial^2 f_{i(g)}(X^*)}{\partial X^2} g, g \right) \geq a(X^*).$$

Так как функция $\left(\frac{\partial^2 f_i(X)}{\partial X^2} g, g \right)$ непрерывна по X в точке X^* равномерно по $g, \|g\|=1$, и $i \in [0:N]$, то найдется такое $\alpha_0, 0 < \alpha_0 < \delta$, что для всех $\alpha \in (0, \alpha_0)$ и всех $g \in G_\gamma$ будет

$$\left(\frac{\partial^2 f_{i(g)}(X^* + \alpha g)}{\partial X^2} g, g \right) \geq \frac{1}{2} a(X^*).$$

Кроме того, из определения $R_2(X^*, g)$ и из того, что $\psi(X^*)=0$, следует

$$\left(\frac{\partial f_{i(g)}(X^*)}{\partial X}, g \right) \geq \gamma.$$

Теперь имеем в силу (4.7) для любого $X = X^* + \alpha g, \alpha \in (0, \alpha_0), g \in G_\gamma$,

$$f_{i(g)}(X) \geq f_{i(g)}(X^*) + \frac{1}{4} \alpha^2 a(X^*) = \varphi(X^*) + \frac{1}{4} \alpha^2 a(X^*).$$

Тем более $\varphi(X) > \varphi(X^*)$.

Введем обозначение

$$\tilde{G}_\gamma = \left\{ g \mid \|g\|=1, \frac{\partial \varphi(X^*)}{\partial g} \geq \gamma \right\}.$$

Поскольку $\psi(X^*) = 0$, т. е.

$$\min_{\|g\|=1} \frac{\partial \varphi(X^*)}{\partial g} = 0,$$

то $G_\gamma \cup \tilde{G}_\gamma = S^0 \stackrel{\text{def}}{=} \{g \mid \|g\| = 1\}$.

Легко показать на основании (4.7), что найдется столь малое α_1 , $0 < \alpha_1 < \alpha_0$, что для всех $X = X^* + \alpha g$, $\alpha \in (0, \alpha_1)$, $g \in \tilde{G}_\gamma$, и любого $i \in R_2(X^*, g)$ будет

$$f_i(X) \geq \varphi(X^*) + \frac{1}{2} \alpha \gamma.$$

Тем более $\varphi(X) > \varphi(X^*)$.

Итак, установлено, что для любого $X \in S_{\alpha_1}(X^*)$, $X \neq X^*$, выполняется неравенство $\varphi(X) > \varphi(X^*)$.

Теорема доказана.

З а м е ч а н и е 1. Неравенство (4.6) выполняется, например, если для всех $i \in R(X^*)$ матрицы $\frac{\partial^2 f_i(X^*)}{\partial X^2}$ являются строго положительно определенными, т. е. если для всех $i \in R(X^*)$ и $V \in E_n$

$$\left(\frac{\partial^2 f_i(X^*)}{\partial X^2} V, V \right) \geq m_i \|V\|^2,$$

где $m_i > 0$ не зависят от V .

З а м е ч а н и е 2. Теорема 4.2, разумеется, остается справедливой, если вместо (4.6) потребовать выполнения более сильного неравенства

$$\min_{\|g\|=1} \max_{i \in R_2(X^*, g)} \left(\frac{\partial^2 f_i(X^*)}{\partial X^2} g, g \right) > 0.$$

З а м е ч а н и е 3. Приведем пример, показывающий, что условие $\gamma > 0$ в теореме 4.2 существенно, т. е. что выполнение неравенства

$$\min_{g \in G_0} \max_{i \in R_2(X^*, g)} \left(\frac{\partial^2 f_i(X^*)}{\partial X^2} g, g \right) > 0, \quad (4.8)$$

где $G_0 = \{g \mid \|g\| = 1, \frac{\partial \varphi(X^*)}{\partial g} = 0\}$, может оказаться недостаточным для того, чтобы точка X^* являлась точкой строгого локального минимума функции $\varphi(X)$.

Пример. Пусть

$$X = (x_1, x_2), \\ f_0(X) = -x_1 - x_2^2, \quad f_1(X) = 3x_1 + x_1^2 + x_2^2.$$

Рассмотрим точку $X^* = (0, 0)$. Ясно, что

$$\varphi(X^*) = \max_{t \in [0; 1]} f_t(X^*) = 0, \\ R(X^*) = \{0, 1\}, \quad H(X^*) = \{(-1, 0), (3, 0)\}, \\ L(X^*) = \{Z_\alpha \mid Z_\alpha = (3 - 4\alpha, 0), 0 \leq \alpha \leq 1\}.$$

Поскольку $0 = Z_{1/4} \in L(X^*)$, то X^* является стационарной точкой функции $\varphi(X)$; при этом

$$\psi(X^*) = \min_{\|g\|=1} \max_{Z \in L(X^*)} (Z, g) = 0.$$

Легко усмотреть, что множество

$$G_0 \triangleq \left\{ g \mid \|g\| = 1, \frac{\partial \varphi(X^*)}{\partial g} = 0 \right\}$$

состоит из двух векторов: $g_1 = (0, 1)$ и $g_2 = (0, -1)$, и что

$$R_2(X^*, g_1) = R_2(X^*, g_2) = \{0, 1\}.$$

Далее

$$\frac{\partial^2 f_0(X^*)}{\partial X^2} = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}, \quad \frac{\partial^2 f_1(X^*)}{\partial X^2} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Поскольку

$$\left(\frac{\partial^2 f_0(X^*)}{\partial X^2} g_1, g_1 \right) = \left(\frac{\partial^2 f_0(X^*)}{\partial X^2} g_2, g_2 \right) = -2, \\ \left(\frac{\partial^2 f_1(X^*)}{\partial X^2} g_1, g_1 \right) = \left(\frac{\partial^2 f_1(X^*)}{\partial X^2} g_2, g_2 \right) = 2,$$

то получаем

$$\min_{g \in G_0} \max_{i \in R_2(X^*, g)} \left(\frac{\partial^2 f_i(X^*)}{\partial X^2} g, g \right) = 2 > 0.$$

Итак, неравенство (4.8) выполнено. Покажем, что, несмотря на это, точка X^* не является точкой локального минимума функции $\varphi(X)$.

Пусть $V_\varepsilon = (-\varepsilon, 1)$, $\varepsilon > 0$. Имеем при $\alpha \geq 0$

$$f_0(X^* + \alpha V_\varepsilon) = \alpha \varepsilon - \alpha^2, \\ f_1(X^* + \alpha V_\varepsilon) = -3\alpha \varepsilon + \alpha^2 \varepsilon^2 + \alpha^2.$$

Рассмотрим, как ведет себя функция максимума

$$\varphi(X^* + \alpha V_\varepsilon) = \max \{f_0(X^* + \alpha V_\varepsilon), f_1(X^* + \alpha V_\varepsilon)\}.$$

Нетрудно проверить, что (рис. 8)

$$\varphi(X^* + \alpha V_\varepsilon) = \begin{cases} f_0(X^* + \alpha V_\varepsilon) & \text{при } \alpha \in \left[0, \frac{4\varepsilon}{2+\varepsilon^2}\right], \\ f_1(X^* + \alpha V_\varepsilon) & \text{при } \alpha \geq \frac{4\varepsilon}{2+\varepsilon^2}. \end{cases}$$

Если обозначить $\alpha_\varepsilon = \frac{4\varepsilon}{2+\varepsilon^2}$, то при $0 < \varepsilon < 1$ получим

$$\varphi(X^* + \alpha_\varepsilon V_\varepsilon) = \frac{4\varepsilon^2(-2+\varepsilon^2)}{(2+\varepsilon^2)^2} < 0 = \varphi(X^*).$$

Остается заметить, что

$$\alpha_\varepsilon V_\varepsilon \xrightarrow{\varepsilon \rightarrow +0} 0.$$

4. Переходим к получению некоторых оценок для величины μ ,

$$\mu = \min_{X \in E_n} \max_{i \in [0:N]} f_i(X) = \min_{X \in E_n} \varphi(X).$$

Функции $f_i(X)$, $i \in [0:N]$, по-прежнему считаются непрерывно дифференцируемыми на всем E_n .

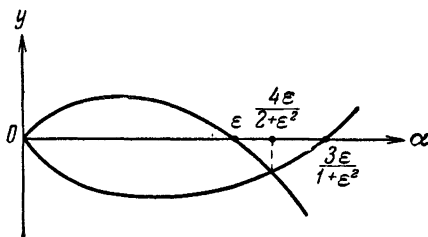


Рис. 8.

Теорема 4.3. *Предположим, что для некоторой точки $X_0 \in E_n$ нашлось такое множество индексов $Q \subset [0:N]$, что*

$$\min_{\|g\|=1} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) \geq 0. \quad (4.9)$$

Если функции $f_i(X)$ для всех $i \in Q$ выпуклы, то справедливы неравенства

$$\min_{i \in Q} f_i(X_0) \leq \mu \leq \varphi(X_0).$$

Доказательство. Неравенство $\mu \leq \varphi(X_0)$ очевидно в силу определения μ . Докажем второе неравенство. Для этого введем функции

$$\begin{aligned}\tilde{f}_i(X) &= f_i(X) - f_i(X_0), \quad i \in Q, \\ \tilde{\varphi}(X) &= \max_{i \in Q} \tilde{f}_i(X).\end{aligned}$$

Очевидно,

$$\frac{\partial \tilde{f}_i(X)}{\partial X} = \frac{\partial f_i(X)}{\partial X}, \quad \tilde{\varphi}(X_0) = 0$$

и

$$\tilde{R}(X_0) \stackrel{\text{df}}{=} \{i \in Q \mid \tilde{f}_i(X_0) = \tilde{\varphi}(X_0)\} = Q. \quad (4.10)$$

Поскольку функции $\tilde{f}_i(X)$ при всех $i \in Q$ выпуклые, то заключаем в силу (4.9), (4.10) и теоремы 3.1, что X_0 есть точка минимума функции $\tilde{\varphi}(X)$ на E_n .

Теперь имеем для всех $X \in E_n$:

$$\begin{aligned}0 = \tilde{\varphi}(X_0) \leq \tilde{\varphi}(X) &= \max_{i \in Q} (f_i(X) - f_i(X_0)) \leq \\ &\leq \max_{i \in [0:N]} f_i(X) - \min_{i \in Q} f_i(X_0), \quad (4.11)\end{aligned}$$

откуда и следует требуемое неравенство

$$\min_{i \in Q} f_i(X_0) \leq \mu.$$

Теорема доказана.

5. Теорема 4.4. Пусть для некоторой точки $X_0 \in E_n$ нашлось множество $Q \subset [0:N]$ такое, что

$$\min_{\|g\|=1} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) \stackrel{\text{df}}{=} \varepsilon < 0. \quad (4.12)$$

Пусть далее функции $f_i(X)$ для всех $i \in Q$ являются дважды непрерывно дифференцируемыми на E_n , причем для всех $i \in Q$, $X \in E_n$ и $V \in E_n$

$$\left(\frac{\partial^2 f_i(X)}{\partial X^2} V, V \right) \geq m \|V\|^2, \quad (4.13)$$

где $m > 0$ не зависит от i , X и V . Тогда

$$\min_{i \in Q} f_i(X_0) - \frac{\varepsilon^2}{2m} \leq \mu \leq \varphi(X_0).$$

Доказательство. Для любых $i \in [0 : N]$, $\alpha \in (-\infty, \infty)$ и g , $\|g\| = 1$, имеет место равенство

$$f_i(X_0 + \alpha g) = f_i(X_0) + \alpha \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) + \frac{1}{2} \alpha^2 \left(\frac{\partial^2 f_i(X_0 + \theta_i g)}{\partial X^2} g, g \right), \quad (4.14)$$

где $\theta_i \in (0, \alpha)$.

Для любого g , $\|g\| = 1$, найдется в силу (4.12) индекс $i(g) \in Q$ такой, что

$$\left(\frac{\partial f_{i(g)}(X_0)}{\partial X}, g \right) \geq -\varepsilon. \quad (4.15)$$

На основании (4.14), (4.15) и (4.13) получим

$$\begin{aligned} f_{i(g)}(X_0 + \alpha g) &\geq f_{i(g)}(X_0) - \alpha\varepsilon + \frac{1}{2} \alpha^2 m \geq \\ &\geq \min_{i \in Q} f_i(X_0) - \alpha\varepsilon + \frac{1}{2} \alpha^2 m. \end{aligned}$$

Так как для любого $\alpha \in (-\infty, \infty)$

$$\frac{1}{2} \alpha^2 m - \alpha\varepsilon \geq \min_{-\infty < \beta < \infty} \left[\frac{1}{2} \beta^2 m - \beta\varepsilon \right] = -\frac{\varepsilon^2}{2m},$$

то

$$\varphi(X_0 + \alpha g) \geq f_{i(g)}(X_0 + \alpha g) \geq \min_{i \in Q} f_i(X_0) - \frac{\varepsilon^2}{2m}.$$

Окончательно имеем

$$\mu \geq \min_{i \in Q} f_i(X_0) - \frac{\varepsilon^2}{2m}.$$

Неравенство $\varphi(X_0) \geq \mu$ очевидно в силу определения μ . Теорема доказана.

Следствие. Если в условиях теоремы $Q \subset R(X_0)$, то

$$0 \leq \varphi(X_0) - \mu \leq \frac{\varepsilon^2}{2m}.$$

§ 5. Метод покоординатного спуска. Метод наискорейшего спуска. Отрицательные примеры

1. В этом параграфе мы обратимся к методам минимизации функции максимума $\varphi(X)$. Начнем с простейшего из них — с метода покоординатного спуска.

Положим

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, 0, \dots, 0), \dots \\ \dots, e_n = (0, \dots, 0, 1).$$

Произвольным образом выберем начальное приближение $X_0 \in E_n$. Допустим, что множество

$$M(X_0) = \{X \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено. В силу непрерывности $\varphi(X)$ множество $M(X_0)$ является к тому же замкнутым.

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Опишем построение X_{k+1} . Рассмотрим прямую

$$X = X_{k1}(\alpha) \stackrel{\text{df}}{=} X_k + \alpha e_1, \quad -\infty < \alpha < \infty,$$

и найдем $\alpha_{k1} \in (-\infty, \infty)$ такое, что

$$\varphi(X_{k1}(\alpha_{k1})) = \min_{\alpha \in (-\infty, \infty)} \varphi(X_{k1}(\alpha)).$$

Поскольку $M(X_0)$ — ограниченное замкнутое множество, минимум здесь достигается, при этом $X_{k1}(\alpha_{k1}) \in M(X_0)$.

Рассмотрим далее прямую

$$X = X_{k2}(\alpha) \stackrel{\text{df}}{=} X_{k1}(\alpha_{k1}) + \alpha e_2, \quad -\infty < \alpha < \infty,$$

и найдем $\alpha_{k2} \in (-\infty, \infty)$ такое, что

$$\varphi(X_{k2}(\alpha_{k2})) = \min_{\alpha \in (-\infty, \infty)} \varphi(X_{k2}(\alpha)).$$

Очевидно, $X_{k2}(\alpha_{k2}) \in M(X_0)$. Продолжая аналогично, придем к прямой

$$X = X_{kn}(\alpha) \stackrel{\text{df}}{=} X_{k, n-1}(\alpha_{k, n-1}) + \alpha e_n, \quad -\infty < \alpha < \infty.$$

Найдем $\alpha_{kn} \in (-\infty, \infty)$, при котором

$$\varphi(X_{kn}(\alpha_{kn})) = \min_{\alpha \in (-\infty, \infty)} \varphi(X_{kn}(\alpha)).$$

Положим $X_{k+1} = X_{kn}(\alpha_{kn})$. Ясно, что $X_{k+1} \in M(X_0)$ и $\varphi(X_{k+1}) \leq \varphi(X_k)$.

Продолжая описанный процесс, получим последовательность точек $X_k \in M(X_0)$, $k = 0, 1, 2, \dots$, причем

$$\varphi(X_0) \geq \varphi(X_1) \geq \dots \geq \varphi(X_k) \geq \dots$$

Можно было ожидать, что предельная точка последовательности $\{X_k\}$ является точкой локального минимума

функции $\varphi(X)$. Однако в общем случае это не так. Более того, предельная точка последовательности $\{X_k\}$ может не быть даже стационарной точкой функции $\varphi(X)$.

Пример 1. Пусть

$$X = (x_1, x_2); \quad e_1 = (1, 0), \quad e_2 = (0, 1).$$

$$f_1(X) = 2x_1 + x_2; \quad f_2(X) = -x_1 - 2x_2;$$

$$f_3(X) = -x_1 + x_2 - 3.$$

$$\varphi(X) = \max_{i \in [1:3]} f_i(X).$$

В качестве начального приближения для метода покоординатного спуска возьмем точку $X_0 = (0, 0)$. Заметим, что

$$\begin{aligned} \varphi(X_0 + \alpha e_1) &= \max \{2\alpha, -\alpha, -\alpha - 3\} = \\ &= \begin{cases} 2\alpha, & \text{если } \alpha \geq 0, \\ -\alpha, & \text{если } \alpha \leq 0. \end{cases} \end{aligned}$$

Отсюда $\alpha_{01} = 0$, $X_{01}(\alpha_{01}) = X_0$.

Далее

$$\begin{aligned} \varphi(X_{01}(\alpha_{01}) + \alpha e_2) &= \max \{\alpha, -2\alpha, \alpha - 3\} = \\ &= \begin{cases} \alpha, & \text{если } \alpha \geq 0, \\ -2\alpha, & \text{если } \alpha \leq 0. \end{cases} \end{aligned}$$

Отсюда $\alpha_{02} = 0$, $X_1 \stackrel{df}{=} X_{02}(\alpha_{02}) = X_0$.

Продолжая аналогично, получим $X_k = X_0$, $k = 1, 2, \dots$, так что в данном случае единственной предельной точкой последовательности $\{X_k\}$ является X_0 .

Проверим точку X_0 на стационарность. Имеем

$$R(X_0) = \{1, 2\}, \quad H(X_0) = \{(2, 1), (-1, -2)\},$$

$$L(X_0) = \{Z_\alpha \mid Z_\alpha = (3\alpha - 1, 3\alpha - 2), 0 \leq \alpha \leq 1\}.$$

Очевидно, $0 \notin L(X_0)$. Значит, X_0 не является стационарной точкой функции $\varphi(X)$.

Замечание. Легко показать, что ближайшей к началу координат точкой множества $L(X_0)$ будет $Z_{1/2} = (\frac{1}{2}, -\frac{1}{2})$ (рис. 9). Пользуясь теоремой 3.3, заключаем, что функция $\varphi(X)$, не убывающая в точке X_0 .

по каждой координате в отдельности, имеет следующее направление наискорейшего спуска:

$$g(X_0) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right).$$

2. Переходим к описанию метода наискорейшего спуска минимизации функции максимума $\varphi(X)$. В основе этого метода лежит то обстоятельство, что если точка $\bar{X}$ не является стационарной точкой функции $\varphi(X)$, т. е. если $\psi(\bar{X}) < 0$, то существует единственное направление наискорейшего спуска $g(\bar{X})$ функции $\varphi(X)$ в точке $\bar{X}$ (см. теорему 3.3).

Возьмем начальное приближение $X_0 \in E_n$.

Предположим, что множество $M(X_0) = \{X | \varphi(X) \leq \varphi(X_0)\}$ ограничено. В силу непрерывности $\varphi(X)$ множество $M(X_0)$ является, кроме того, замкнутым.

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $\psi(X_k) \geq 0$, то X_k — стационарная точка, и процесс на этом заканчивается. В противном случае строим вектор $g_k = g(X_k)$. Рассмотрим луч

$$X = X_k(\alpha) \stackrel{\text{дф}}{=} X_k + \alpha g_k, \quad \alpha \geq 0,$$

и найдем $\alpha_k \in [0, \infty)$, при котором

$$\varphi(X_k(\alpha_k)) = \min_{\alpha \in [0, \infty)} \varphi(X_k(\alpha))$$

(минимум здесь достигается в силу ограниченности и замкнутости $M(X_0)$).

Положим $X_{k+1} = X_k(\alpha_k)$. Ясно, что $X_{k+1} \in M(X_0)$ и $\varphi(X_{k+1}) < \varphi(X_k)$. Продолжая описанный процесс, получим последовательность $\{X_k\}$, причем $X_k \in M(X_0)$, $k = 0, 1, 2, \dots$,

$$\varphi(X_0) > \varphi(X_1) > \dots > \varphi(X_k) > \dots,$$

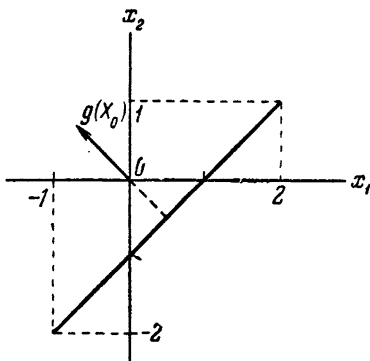


Рис. 9.

и если последовательность $\{X_k\}$ конечна, то последний ее элемент по построению является стационарной точкой функции $\varphi(X)$.

Применим метод наискорейшего спуска к минимизации функции $\varphi(X)$, рассмотренной в примере 1. В качестве начального приближения возьмем точку $X_0 = (0, 0)$. Как уже отмечалось, направлением наискорейшего спуска в этом случае будет

$$g_0 = g(X_0) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right).$$

Имеем

$$\begin{aligned} \varphi(X_0 + \alpha g_0) &= \max \left\{ -\frac{\sqrt{2}}{2} \alpha, -\frac{\sqrt{2}}{2} \alpha, \sqrt{2} \alpha - 3 \right\} = \\ &= \frac{\sqrt{2}}{2} \max \{ -\alpha, 2\alpha - 3 \sqrt{2} \}. \end{aligned}$$

Очевидно, $\alpha_0 = \sqrt{2}$ и $X_1 \stackrel{df}{=} X_0 + \alpha_0 g_0 = (-1, 1)$. Далее $f_1(X_1) = f_2(X_1) = f_3(X_1) = -1$,

$$R(X_1) = \{1, 2, 3\}, \quad H(X_1) = \{(2, 1), (-1, -2), (-1, 1)\}.$$

Начало координат принадлежит треугольнику $L(X_1)$ (рис. 10). Значит, X_1 — стационарная точка. Более того,

X_1 является точкой минимума функции $\varphi(X)$, ибо в данном случае $\varphi(X)$ — выпуклая функция. Наконец, отметим, что $\varphi(X_1) = -1$.

Метод наискорейшего спуска является более гибким методом, чем метод по координатного спуска. Однако и он в общем случае не обладает основным свойством — сходимостью. Это означает, что при минимизации функции максимума $\varphi(X)$ предельная точка последовательности $\{X_k\}$, построенной по методу наискорейшего спуска, может не

быть стационарной точкой функции $\varphi(X)$. Причина здесь кроется в том, что функция максимума $\varphi(X)$ не является гладкой (непрерывно дифференцируемой) функцией.

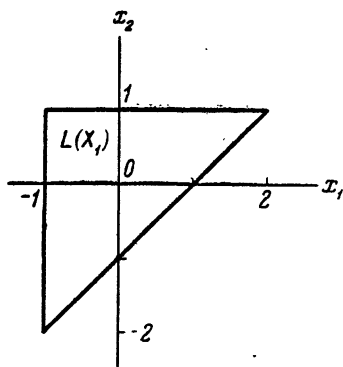


Рис. 10.

Ниже приводится пример возможного в рассматриваемом методе эффекта «заедания».

3*. Пример 2. Пусть

$$\begin{aligned} X &= (x_1, x_2); \\ f_1(X) &= -5x_1 + x_2; \quad f_2(X) = 4x_2 + x_1^2 + x_2^2; \\ f_3(X) &= 5x_1 + x_2; \\ \varphi(X) &= \max_{i \in [1:3]} f_i(X). \end{aligned}$$

Поскольку функции $f_i(X)$, $i \in [1:3]$, выпуклы, то выпуклой будет и функция $\varphi(X)$.

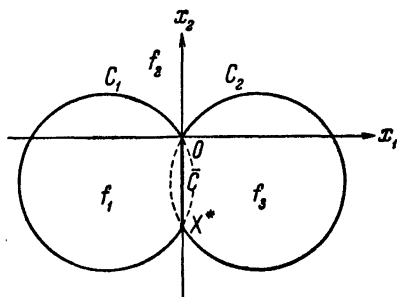


Рис. 11.

Введем обозначения:

$$\begin{aligned} C_1 &\stackrel{df}{=} \{X \mid f_1(X) = f_2(X)\} = \\ &= \left\{ X \mid \left(x_1 + \frac{5}{2}\right)^2 + \left(x_2 + \frac{3}{2}\right)^2 = \frac{34}{4} \right\}, \end{aligned}$$

$$\begin{aligned} C_2 &\stackrel{df}{=} \{X \mid f_2(X) = f_3(X)\} = \\ &= \left\{ X \mid \left(x_1 - \frac{5}{2}\right)^2 + \left(x_2 + \frac{3}{2}\right)^2 = \frac{34}{4} \right\}, \end{aligned}$$

$$\bar{C} \stackrel{df}{=} \{X \mid f_1(X) = f_3(X)\} = \{X \mid x_1 = 0\}.$$

Нетрудно проверить, что (рис. 11)

$$\varphi(X) = \begin{cases} f_1(X), & \text{если } X \text{ лежит внутри } C_1 \text{ и левее } \bar{C}, \\ f_3(X), & \text{если } X \text{ лежит внутри } C_2 \text{ и правее } \bar{C}, \\ f_2(X), & \text{если } X \text{ лежит вне } C_1 \text{ и } C_2. \end{cases}$$

В точках $\bar{X} = (0, 0)$ и $X^* = (0, -3)$ имеем

$$\begin{aligned} f_1(\bar{X}) &= f_2(\bar{X}) = f_3(\bar{X}) = 0, \\ f_1(X^*) &= f_2(X^*) = f_3(X^*) = -3. \end{aligned}$$

Покажем, что X^* является стационарной точкой, а $\bar{X}$ не является стационарной точкой функции $\varphi(X)$. Для этого заметим, что

$$\begin{aligned} \frac{\partial f_1(X)}{\partial X} &= (-5, 1), & \frac{\partial f_2(X)}{\partial X} &= (2x_1, 4 + 2x_2), \\ \frac{\partial f_3(X)}{\partial X} &= (5, 1). \end{aligned}$$

Далее

$$R(X^*) = \{1, 2, 3\}, \quad H(X^*) = \{(-5, 1), (0, -2), (5, 1)\}.$$

Начало координат принадлежит треугольнику $L(X^*)$ (рис. 12). Значит, $X^* = (0, -3)$ является стационарной

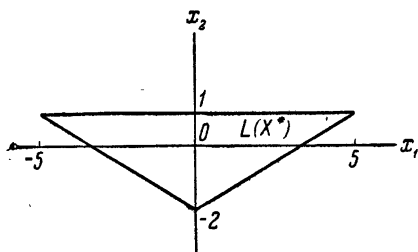


Рис. 12.

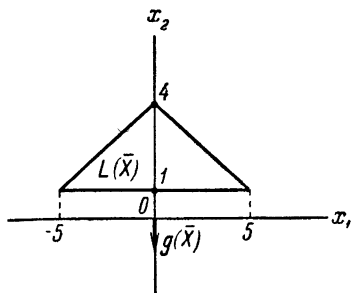


Рис. 13.

точкой. Поскольку функция $\varphi(X)$ выпукла, то X^* является точкой минимума функции $\varphi(X)$. Отметим, что $\varphi(X^*) = -3$.

Теперь обратимся к точке $\bar{X}$. Имеем

$$R(\bar{X}) = \{1, 2, 3\}, \quad H(\bar{X}) = \{(-5, 1), (0, 4), (5, 1)\}.$$

Начало координат не принадлежит треугольнику $L(\bar{X})$ (рис. 13), значит, $\bar{X} = (0, 0)$ не является стационарной точкой функции $\varphi(X)$.

Наша цель будет заключаться в построении методом наискорейшего спуска последовательности $\{X_k\}$, сходящейся к точке $\bar{X}$.

Основная трудность при этом сводится к выбору начального приближения X_0 .

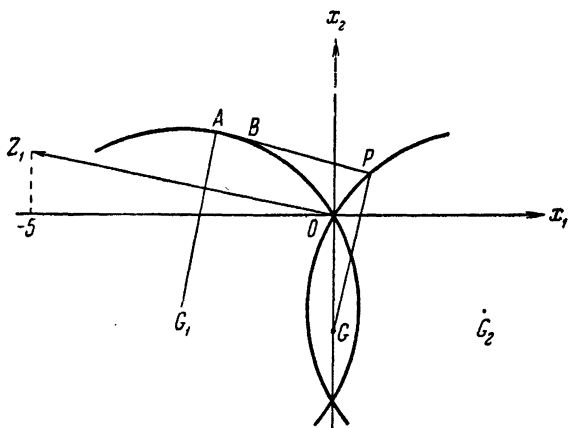


Рис. 14.

Введем обозначения (рис. 14):

$G_1 = \left(-\frac{5}{2}, -\frac{3}{2}\right)$ — центр окружности C_1 ;

$G = (0, -2)$ — точка минимума функции $f_2(X)$;

$Z_1 = (-5, 1)$ — градиент функции $f_1(X)$;

$Z(X) \stackrel{df}{=} (2x_1, 4 + 2x_2) = 2(X - G)$ — градиент функции $f_2(X)$ (точку мы воспринимаем и как вектор, компоненты которого совпадают с координатами этой точки);

$$C'_1 = \left\{ X = (x_1, x_2) \mid X \in C_1, -\frac{5}{2} < x_1 < 0, x_2 > 0 \right\};$$

A — точка, принадлежащая C'_1 и удовлетворяющая условию $(A - G_1, Z_1) = 0$ (отрезок $\overline{AG_1}$ перпендикулярен отрезку $\overline{OZ_1}$);

B — точка, принадлежащая C'_1 и удовлетворяющая следующему условию: если P — первая точка пересечения касательной к окружности C_1 в точке B с окружностью C_2 , то $(P - B, P - G) = 0$;

$\tilde{C}_1$ — множество точек окружности C_1 , расположенных правее точек A и B и левее точки O .

Утверждается, что если взять $X_0 \in \tilde{C}_1$ и построить методом наискорейшего спуска последовательность $\{X_k\}$, то эта последовательность будет сходиться к $\bar{X} = (0, 0)$.

Доказательство разбивается на три этапа. На первом этапе будет показано, что направлением наискорейшего спуска функции $\varphi(X)$ в точке X_0 является направление по касательной к окружности C_1 в точке X_0 в сторону окружности C_2 . На втором этапе устанавливается, что минимум функции $\varphi(X)$ в направлении наискорейшего спуска достигается в первой точке пересечения рассматриваемой касательной с окружностью C_2 . Таким образом, X_1 есть первая точка пересечения касательной к окружности C_1 в точке X_0 с окружностью C_2 . Теперь понятно, какой будет вся последовательность $\{X_k\}$. На третьем этапе показывается, что $X_k \xrightarrow{k \rightarrow \infty} \bar{X}$.

Переходим к подробному проведению всего доказательства.

Первый этап. Пусть $X_0 = (x_1^{(0)}, x_2^{(0)}) \in \tilde{C}_1$. В этом случае

$$R(X_0) = \{1, 2\}, \quad H(X_0) = \{Z_1, Z(X_0)\}.$$

Множество $L(X_0)$ есть отрезок, соединяющий точки Z_1 и $Z(X_0)$.

Легко проверить, что вектор

$$V_0 = \left(-1, \frac{2x_1^{(0)} + 5}{2x_2^{(0)} + 3} \right) \quad (5.1)$$

перпендикулярен к прямой, проходящей через точки Z_1 и $Z(X_0)$, т. е. что $(Z_1 - Z(X_0), V_0) = 0$. С другой стороны, тот же вектор V_0 перпендикулярен к радиусу — вектору точки X_0 , проведенному из точки $G_1: (X_0 - G_1, V_0) = 0$. Итак, вектор V_0 , перпендикулярный к прямой, проходящей через точки Z_1 и $Z(X_0)$, является касательным к окружности C_1 в точке X_0 . Учитывая, что X_0 лежит на C_1 правее A , заключаем, что вектор V_0 направлен на ближайшую к началу координат точку множества $L(X_0)$ (рис. 15, где векторы $Z_1, Z(X_0) \stackrel{df}{=} 2(X_0 - G)$ и V_0 проведены из точки X_0). Таким образом, направлением

наискорейшего спуска функции $\varphi(X)$ в точке X_0 является вектор

$$g(X_0) = -\frac{V_0}{\|V_0\|}. \quad (5.2)$$

Это есть направление по касательной к окружности C_1 в точке X_0 в сторону окружности C_2 .

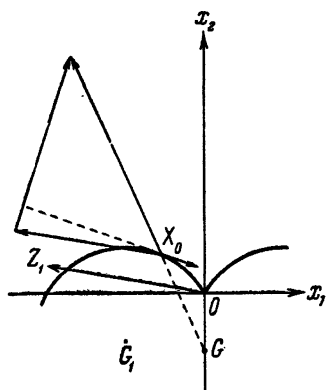


Рис. 15.

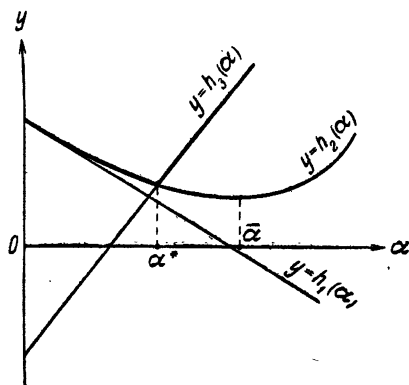


Рис. 16.

Второй этап. Положим $g_0 = g(X_0)$ и рассмотрим функцию $\varphi(X)$ на луче

$$X = X_0(\alpha) \stackrel{\text{df}}{=} X_0 + \alpha g_0, \quad \alpha \geq 0.$$

Введем обозначение

$$h_i(\alpha) = f_i(X_0 + \alpha g_0), \quad i \in [1:3].$$

Очевидно,

$$\varphi(X_0 + \alpha g_0) = \max_{i \in [1:3]} h_i(\alpha).$$

После несложных вычислений, учитывая (5.1) и (5.2), получим (рис. 16)

$$h_1(\alpha) = f_1(X_0) + 2\alpha(X_0 - G, g_0),$$

$$h_2(\alpha) = f_2(X_0) + 2\alpha(X_0 - G, g_0) + \alpha^2(g_0, g_0),$$

$$h_3(\alpha) = f_3(X_0) - 2\alpha(X_0 - G_3, g_0),$$

где $G_3 = (0, -1)$.

Заметим, что прямая $y = h_1(\alpha)$ касается параболы $y = h_2(\alpha)$ в точке $\alpha = 0$, ибо $f_1(X_0) = f_2(X_0)$ и $h'_1(0) = h'_2(0)$. Далее $h_2(\alpha)$ достигает минимума при

$$\bar{\alpha} = - \frac{(X_0 - G, g_0)}{(g_0, g_0)}.$$

Обозначим через α^* первую точку пересечения прямой $y = h_3(\alpha)$ с параболой $y = h_2(\alpha)$ (соответствующая точка $X_0(\alpha^*) \stackrel{df}{=} X_0 + \alpha^* g_0$ лежит на пересечении касательной к окружности C_1 в точке X_0 с окружностью C_2).

Если мы докажем, что

$$\alpha^* < \bar{\alpha}, \quad (5.3)$$

то это будет означать, что $\varphi(X)$ достигает минимума на луче $X = X_0(\alpha)$ в точке $X_0(\alpha^*)$.

Неравенство (5.3) следует из того, что X_0 лежит на C_1 правее B . Действительно, легко проверить, что

$$(X_0 + \bar{\alpha} g_0 - G, g_0) = 0.$$

Это означает, что радиус-вектор точки $X_0(\bar{\alpha})$, проведенный из точки G , перпендикулярен к касательной к окружности C_1 в точке X_0 . Таким образом, получаем простое правило для нахождения $X_0(\bar{\alpha})$: нужно из точки G опустить перпендикуляр на касательную к окружности C_1 в точке X_0 ; основание этого перпендикуляра и есть $X_0(\bar{\alpha})$. Точка B обладает тем свойством, что если взять $X_0 = B$, то соответствующее $\bar{\alpha}$ будет совпадать с α^* . Так как X_0 лежит на C_1 правее B , то $\alpha^* < \bar{\alpha}$.

Итак, доказано, что минимум функции $\varphi(X)$ на луче $X_0(\alpha)$ достигается в точке $X_0(\alpha^*)$. Это означает, что X_1 есть первая точка пересечения касательной к окружности C_1 в точке X_0 с окружностью C_2 .

Проводя аналогичные рассуждения для X_1 , получим, что метод наискорейшего спуска приводит нас к точке X_2 , являющейся первой точкой пересечения касательной к окружности C_2 в точке X_1 с окружностью C_1 и т. д. (рис. 17).

Таким образом, можно считать построенной всю последовательность $\{X_k\}$. Осталось доказать, что

$$X_k \xrightarrow[k \rightarrow \infty]{} \bar{X}.$$

Третий этап. Введем точки P_0, P_1, P_2, P_3 , как это показано на рис. 18. Обозначим $\angle X_1 X_0 P_0 = \alpha_0$, $\angle X_0 O P_3 = \beta$. Очевидно, $\|X_1\| \leq \|P_1\|$ и

$$\frac{\|X_1\|}{\|X_0\|} \leq \frac{\|P_1\|}{\|P_0\|} = \frac{\|P_2\|}{\|P_3\|} = 1 - \operatorname{tg} \beta \operatorname{tg} \alpha_0 \leq 1 - \operatorname{tg} \alpha_0 \operatorname{tg} \beta_0,$$

где β_0 — острый угол между касательной к окружности C_1 в точке O с осью $x_1 = 0$.

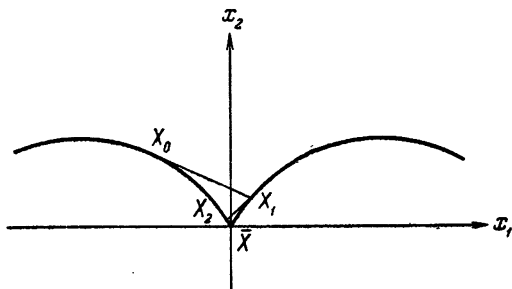


Рис. 17.

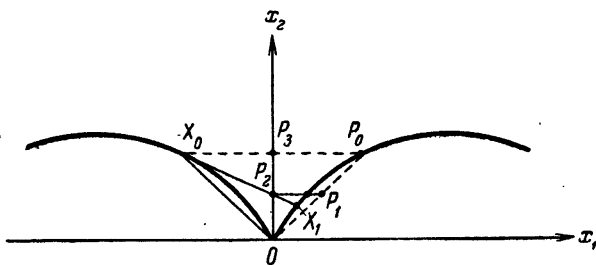


Рис. 18.

Положив $1 - \operatorname{tg} \alpha_0 \operatorname{tg} \beta_0 = q$, $0 < q < 1$, получим

$$\frac{\|X_1\|}{\|X_0\|} \leq q.$$

Аналогично показывается, что для любого натурального k

$$\frac{\|X_k\|}{\|X_{k-1}\|} \leq q.$$

Отсюда $\|X_k\| \leq q^k \|X_0\|$, $0 < q < 1$. Значит, $X_k \xrightarrow{k \rightarrow \infty} 0$, или, что то же самое,

$$X_k \xrightarrow{k \rightarrow \infty} \bar{X}.$$

Остается напомнить, что $\bar{X}$ не является стационарной точкой функции $\varphi(X)$.

Замечание. В рассмотренном примере функция $\varphi(X)$ терпит разрыв во всех точках $X = \bar{X} + \alpha(X^* - \bar{X})$, $0 \leq \alpha \leq 1$. Доказательство этого утверждения, опирающееся на лемму 3.3, предоставляется читателю.

§ 6. Первый метод последовательных приближений

1. Для дальнейшего нам понадобятся некоторые новые понятия.

Как и раньше, будем считать, что функции $f_i(X)$, $i \in [0 : N]$, непрерывно дифференцируемы на всем E_n ; пусть

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X).$$

Положим при фиксированном $X \in E_n$

$$R_\varepsilon(X) = \{i \mid \varphi(X) - f_i(X) \leq \varepsilon\}, \quad \varepsilon \geq 0.$$

Очевидно, $R_0(X) = R(X)$, и если $\varepsilon' \geq \varepsilon$, то

$$R_{\varepsilon'}(X) \supset R_\varepsilon(X). \quad (6.1)$$

Опишем более подробно зависимость множества $R_\varepsilon(X)$ от ε . Для этого перенумеруем числа $f_i \stackrel{\text{df}}{=} f_i(X)$ (X у нас фиксировано) в порядке убывания:

$$f_{i_{01}} = \dots = f_{i_{0j_0}} > f_{i_{11}} = \dots \\ \dots = f_{i_{1j_1}} > \dots > f_{i_{m1}} = \dots = f_{i_{mj_m}}.$$

Очевидно, $0 \leq m \leq N$, причем $m \stackrel{\text{df}}{=} m(X)$ зависит от X . Обозначим

$$J_k(X) = \{i_{k1}, \dots, i_{kj_k}\}, \quad k \in [0 : m]; \\ a_k(X) = \varphi(X) - f_{i_k}(X), \quad i \in J_k(X).$$

Ясно, что $a_0(X) = 0$ и (рис. 19)

$$0 = a_0(X) < a_1(X) < \dots < a_m(X).$$

По определению будем считать, что $a_{m+1}(X) = \infty$,

Введем индексные множества

$$R^s(X) = \bigcup_{k=0}^s J_k(X), \quad s \in [0 : m].$$

Имеем $R^0(X) = R(X)$, $R^m(X) = [0 : N]$. Теперь можно записать основное представление для $R_\varepsilon(X)$:

$$R_\varepsilon(X) = R^s(X), \text{ если } a_s(X) \leq \varepsilon < a_{s+1}(X); \quad s \in [0 : m]. \quad (6.2)$$

В частности,

$$R_\varepsilon(X) = R(X), \quad \text{если } 0 \leq \varepsilon < a_1(X);$$

$$R_\varepsilon(X) = [0 : N], \quad \text{если } a_m(X) \leq \varepsilon < \infty.$$

Лемма 6.1. Для произвольного $X_0 \in E_n$ существуют числа $\gamma_0 > 0$ и $\delta_0 > 0$ такие, что для любых $\varepsilon \in [0, \gamma_0]$ и $X' \in E_n$, $\|X' - X_0\| < \delta_0$ будет

$$R_\varepsilon(X') \subset R(X_0).$$

Доказательство. Если $R(X_0) = [0 : N]$, то в качестве γ_0 и δ_0 можно взять любые положительные числа.

Пусть $R(X_0) \neq [0 : N]$. Тогда положим $\gamma_0 = \frac{1}{2} a_1(X_0)$, а $\delta_0 > 0$ выберем таким, чтобы из неравенства $\|X' - X_0\| < \delta_0$ следовало

$$|f_i(X') - f_i(X_0)| < \frac{1}{4} a_1(X_0)$$

равномерно по $i \in [0 : N]$. Заметим, что в этом случае выполняется также неравенство

$$|\varphi(X') - \varphi(X_0)| < \frac{1}{4} a_1(X_0).$$

Пусть теперь $\varepsilon \in [0, \gamma_0]$ и $\|X' - X_0\| < \delta_0$. Для $i \in R_\varepsilon(X')$ имеем

$$\begin{aligned} \varphi(X_0) - f_i(X_0) &= [\varphi(X_0) - \varphi(X')] + [\varphi(X') - f_i(X')] + \\ &+ [f_i(X') - f_i(X_0)] < \frac{1}{4} a_1(X_0) + \varepsilon + \frac{1}{4} a_1(X_0) \leq a_1(X_0). \end{aligned}$$

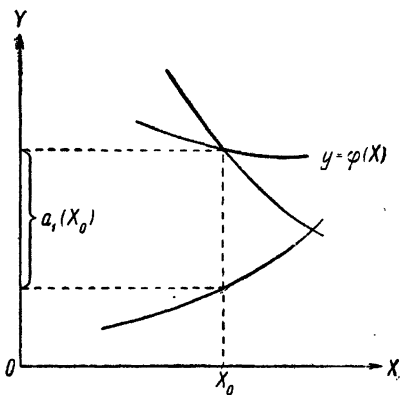


Рис. 19.

Итак, $\varphi(X_0) - f_i(X_0) < a_1(X_0)$, что возможно лишь тогда, когда $i \in R(X_0)$.

Лемма доказана.

2. Введем функцию

$$\psi_\varepsilon(X) = \min_{\|g\|=1} \max_{i \in R_\varepsilon(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right), \quad \varepsilon \geq 0.$$

Очевидно, $\psi_0(X) = \psi(X)$.

В силу (6.1) и (6.2) при фиксированном X функция $\psi_\varepsilon(X)$ как функция от ε , $0 \leq \varepsilon < \infty$, является

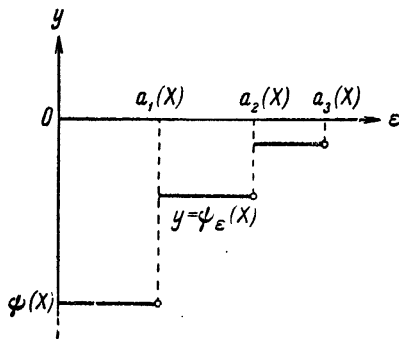


Рис. 20.

неубывающей и кусочно-постоянной, причем (рис. 20) $\psi_\varepsilon(X) = \psi_{a_s(X)}(X)$, если $a_s(X) \leq \varepsilon < a_{s+1}(X)$, $s \in [0:m]$. (6.3)

Заметим, в частности, что

$$\psi_\varepsilon(X) = \psi(X), \quad \text{если} \quad 0 \leq \varepsilon < a_1(X). \quad (6.4)$$

Определение 1. Точку $X^* \in E_n$, для которой выполняется неравенство

$$\psi_\varepsilon(X^*) \geq 0,$$

будем называть ε -стационарной точкой функции $\varphi(X)$.

Определение 2. Вектор $g_\varepsilon(\bar{X})$, $\|g_\varepsilon(\bar{X})\| = 1$, называется направлением ε -наискорейшего спуска функции $\varphi(X)$ в точке $\bar{X}$, если

$$\max_{i \in R_\varepsilon(\bar{X})} \left(\frac{\partial f_i(\bar{X})}{\partial X}, g_\varepsilon(\bar{X}) \right) = \psi_\varepsilon(\bar{X}).$$

Положим

$$H_\varepsilon(X) = \left\{ Z = \frac{\partial f_i(X)}{\partial X} \mid i \in R_\varepsilon(X) \right\}.$$

Через $L_\varepsilon(X)$ будем обозначать выпуклую оболочку, натянутую на точки множества $H_\varepsilon(X)$.

Справедливы следующие два утверждения.

I. Для того чтобы точка X^* была ε -стационарной точкой функции $\varphi(X)$, необходимо и достаточно, чтобы начало координат принадлежало $L_\varepsilon(X^*)$:

$$0 \in L_\varepsilon(X^*).$$

II. Если $\bar{X}$ не является ε -стационарной точкой функции $\varphi(X)$, т. е. если $\psi_\varepsilon(\bar{X}) < 0$, то функция $\varphi(X)$ имеет в точке $\bar{X}$ единственное направление ε -наискорейшего спуска $g_\varepsilon(\bar{X})$. При этом

$$g_\varepsilon(\bar{X}) = - \frac{Z_\varepsilon^*}{\|Z_\varepsilon^*\|},$$

где Z_ε^* — ближайшая к началу координат точка множества $L_\varepsilon(\bar{X})$ и $\psi_\varepsilon(\bar{X}) = -\|Z_\varepsilon^*\|$.

Доказательство этих утверждений вполне аналогично доказательству теорем 3.2 и 3.3. —

3. Теперь можно перейти к описанию первого метода последовательных приближений для нахождения стационарных точек функции $\varphi(X)$.

Зафиксируем прежде всего два параметра $\varepsilon_0 > 0$ и $\rho_0 > 0$. Возьмем начальное приближение $X_0 \in E_n$. Предположим, что множество $M(X_0) = \{X \mid \varphi(X) \leq \varphi(X_0)\}$ ограничено. В силу непрерывности $\varphi(X)$ множество $M(X_0)$ является к тому же замкнутым.

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $\psi(X_k) \geq 0$, то X_k — стационарная точка, и процесс на этом заканчивается.

Предположим, что $\psi(X_k) < 0$. Тогда перебираем числовую последовательность $\varepsilon_\nu = \varepsilon_0/2^\nu$, $\nu = 0, 1, 2, \dots$, до тех пор, пока не выполнится неравенство (рис. 21)

$$\psi_{\varepsilon_\nu}(X_k) \leq - \frac{\rho_0}{\varepsilon_0} \varepsilon_\nu \quad (6.5)$$

(в частности, это может случиться и при $v=0$). Первое v , при котором выполняется неравенство (6.5), обозначим через v_k . Заметим, что v_k — конечное число. Это

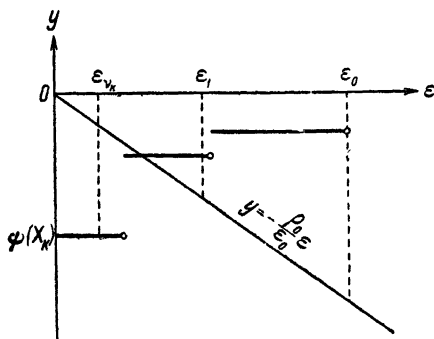


Рис. 21.

следует из того, что правая часть неравенства (6.5) стремится к нулю при $v \rightarrow \infty$, а левая в силу (6.4) — к $\psi(X_k) < 0$.

Итак,

$$\psi_{\varepsilon_{v_k}}(X_k) \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon_{v_k},$$

и если $v_k > 0$, то

$$\psi_{\varepsilon_{v_k-1}}(X_k) > -\frac{\rho_0}{\varepsilon_0} \varepsilon_{v_k-1}.$$

В дальнейшем для простоты будем использовать обозначение

$$\hat{\varepsilon}_k = \varepsilon_{v_k}.$$

Положим $g_k = g_{\hat{\varepsilon}_k}(X_k)$, где $g_{\hat{\varepsilon}_k}(X_k)$ — направление $\hat{\varepsilon}_k$ -наискорейшего спуска функции $\varphi(X)$ в точке X_k , и рассмотрим луч

$$X = X_k(\alpha) \stackrel{df}{=} X_k + \alpha g_k, \quad \alpha \geq 0.$$

Найдем $\alpha_k \in [0, \infty)$, при котором

$$\varphi(X_k(\alpha_k)) = \min_{\alpha \in [0, \infty)} \varphi(X_k(\alpha))$$

(инфимум здесь достигается в силу ограниченности и замкнутости $M(X_0)$).

В качестве $(k+1)$ -го приближения X_{k+1} возьмем $X_k(\alpha_k)$:

$$X_{k+1} = X_k(\alpha_k).$$

Ясно, что $X_{k+1} \in M(X_0)$. Нетрудно показать, что

$$\varphi(X_{k+1}) < \varphi(X_k). \quad (6.6)$$

Действительно, для достаточно малых α имеем

$$\begin{aligned} \varphi(X_k + \alpha g_k) &= \\ &= \max_{i \in R(X_k)} \left\{ f_i(X_k) + \alpha \left(\frac{\partial f_i(X_k)}{\partial X}, g_k \right) + o_i(g_k; \alpha) \right\} \leq \\ &\leq \varphi(X_k) + \alpha \max_{i \in R_{\hat{e}_k}(X_k)} \left(\frac{\partial f_i(X_k)}{\partial X}, g_{\hat{e}_k}(X_k) \right) + \max_{i \in [0:N]} o_i(g_k; \alpha) = \\ &= \varphi(X_k) + \alpha \psi_{\hat{e}_k}(X_k) + \max_{i \in [0:N]} o_i(g_k; \alpha) \leq \varphi(X_k) - \frac{1}{2} \alpha \frac{\rho_0}{\varepsilon_0} \hat{e}_k. \end{aligned}$$

Отсюда очевидным образом следует неравенство (6.6).

Итак, можно считать построенной последовательность $\{X_k\}$ и связанные с ней последовательности

$$\{\hat{e}_k\}, \{g_k\}, \{\alpha_k\},$$

причем $X_k \in M(X_0)$, $k = 0, 1, 2, \dots$,

$$\varphi(X_0) > \varphi(X_1) > \dots > \varphi(X_k) > \dots, \quad (6.7)$$

и если последовательность $\{X_k\}$ конечна, то последний ее элемент по построению является стационарной точкой функции $\varphi(X)$.

Рассмотрим случай, когда последовательность $\{X_k\}$ бесконечна.

Лемма 6.2. При $k \rightarrow \infty$ выполняется предельное соотношение

$$\hat{e}_k \rightarrow 0.$$

Доказательство. Предположим противное. Тогда найдутся число $\varepsilon^* > 0$ и подпоследовательность $\{\hat{e}_{k_j}\}$ такие, что

$$\hat{e}_{k_j} \geq \varepsilon^*.$$

Учитывая, что $R_{e^*}(X_{k_j}) \subset R_{e_{k_j}}(X_{k_j})$, получим для $i \in R_{e^*}(X_{k_j})$:

$$\begin{aligned} f_i(X_{k_j} + \alpha g_{k_j}) &\leq f_i(X_{k_j}) + \alpha \psi_{e_{k_j}}(X_{k_j}) + o_{ik_j}(\alpha) \leq \\ &\leq f_i(X_{k_j}) - \alpha \rho^* + o_{ik_j}(\alpha), \end{aligned}$$

где $\rho^* = \frac{\rho_0}{\varepsilon_0} \varepsilon^*$. Поскольку $X_{k_j} \in M(X_0)$ и $M(X_0)$ — ограниченное замкнутое множество, то в силу замечания 2 к лемме 3.1 Приложения III найдется $\alpha_0 > 0$, не зависящее от i и k_j , такое, что для $i \in R_{e^*}(X_{k_j})$ и $\alpha \in [0, \alpha_0]$ будет

$$f_i(X_{k_j} + \alpha g_{k_j}) \leq \varphi(X_{k_j}) - \frac{1}{2} \alpha \rho^*. \quad (6.8)$$

Теперь положим

$$C_1 = \max_{X \in M(X_0)} \max_{i \in [0: N]} \left\| \frac{\partial f_i(X)}{\partial X} \right\|.$$

Для $i \notin R_{e^*}(X_{k_j})$ будем иметь

$$\begin{aligned} f_i(X_{k_j} + \alpha g_{k_j}) &\leq f_i(X_{k_j}) + \alpha C_1 + o_{ik_j}(\alpha) \leq \\ &\leq \varphi(X_{k_j}) - \varepsilon^* + \alpha C_1 + o_{ik_j}(\alpha). \end{aligned}$$

Найдется столь малое α_1 , $0 < \alpha_1 \leq \alpha_0$, не зависящее от i и k_j , что для $i \notin R_{e^*}(X_{k_j})$ и $\alpha \in [0, \alpha_1]$ будет

$$f_i(X_{k_j} + \alpha g_{k_j}) \leq \varphi(X_{k_j}) - \frac{1}{2} \varepsilon^*. \quad (6.9)$$

Зафиксировав теперь некоторое α' , $0 < \alpha' \leq \alpha_1$, получим в силу (6.8) и (6.9) равномерно по k_j

$$\varphi(X_{k_j} + \alpha' g_{k_j}) \leq \varphi(X_{k_j}) - \beta, \quad (6.10)$$

где $\beta = \min \left\{ \frac{1}{2} \alpha' \rho^*, \frac{1}{2} \varepsilon^* \right\}$.

Это неравенство немедленно приводит нас к противоречию. Действительно, в силу (6.7) вся последовательность $\{\varphi(X_k)\}$, ограниченная снизу числом $\bar{\varphi} = \min_{X \in M(X_0)} \varphi(X)$,

имеет предел

$$\varphi(X_k) \xrightarrow{k \rightarrow \infty} \varphi^*,$$

причем при каждом $k = 0, 1, 2, \dots$

$$\varphi(X_k) > \varphi^*. \quad (6.11)$$

Выберем $\bar{k}_j$ столь большим, чтобы

$$\varphi(X_{\bar{k}_j}) < \varphi^* + \frac{1}{2}\beta.$$

Получим в силу (6.10)

$$\varphi(X_{\bar{k}_j+1}) \leq \varphi(X_{\bar{k}_j} + \alpha' g_{\bar{k}_j}) \leq \varphi^* - \frac{1}{2}\beta,$$

что противоречит (6.11).

Лемма доказана.

4. Если последовательность $\{X_k\}$, построенная в предыдущем пункте, бесконечна, то она имеет хотя бы одну предельную точку, поскольку все члены этой последовательности принадлежат ограниченному замкнутому множеству $M(X_0)$.

Теорема 6.1. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$.*

Доказательство. Пусть

$$X_{k_j} \xrightarrow{k_j \rightarrow \infty} X^*.$$

Ясно, что $X^* \in M(X_0)$. Требуется показать, что X^* является стационарной точкой функции $\varphi(X)$, т. е. что

$$\psi(X^*) \geq 0.$$

Предположим противное:

$$\psi(X^*) \stackrel{\text{def}}{=} \min_{\|g\|=1} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) = -b < 0.$$

В силу леммы 6.1 найдется число $\gamma^* > 0$ такое, что для достаточно больших номеров $k_j > K_1$ и $\varepsilon \in [0, \gamma^*]$ будет

$$R_\varepsilon(X_{k_j}) \subset R(X^*).$$

Отсюда следует, что для $\varepsilon \in [0, \gamma^*]$ и $k_j > K_1$

$$\psi_\varepsilon(X_{k_j}) \leq \min_{\|g\|=1} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, g \right).$$

Выберем столь большое натуральное $K \geq K_1$, чтобы для $k_j > K$ выполнялось неравенство

$$\left| \min_{\|g\|=1} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, g \right) - \min_{\|g\|=1} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) \right| < \frac{b}{2}.$$

Получим для $k_j > K$ и $\varepsilon \in [0, \gamma^*]$

$$\psi_\varepsilon(X_{k_j}) \leq -\frac{b}{2}.$$

Положим

$$\varepsilon' = \min \left\{ \gamma^*, \frac{\varepsilon_0}{\rho_0} \cdot \frac{b}{2} \right\}.$$

Через ε^* обозначим наибольший из членов последовательности $\varepsilon_v = \varepsilon_0/2^v$, $v = 0, 1, 2, \dots$, меньших ε' . В этом случае для всех $k_j > K$ будет

$$\psi_{\varepsilon^*}(X_{k_j}) \leq -\frac{b}{2} \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon' \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon^*.$$

Если вспомнить, что $\hat{\varepsilon}_{k_j}$ есть первый член последовательности $\{\varepsilon_v\}$, для которого выполняется аналогичное неравенство, то получим для всех $k_j > K$

$$\hat{\varepsilon}_{k_j} \geq \varepsilon^*,$$

что противоречит лемме 6.2.

Теорема доказана.

З а м е ч а н и е. На каждом шаге описанного метода последовательных приближений приходится находить $\hat{\varepsilon}_k$, g_k и α_k . Наиболее трудоемкой задачей является выбор $\hat{\varepsilon}_k = \varepsilon_{v_k}$ — первого члена последовательности $\varepsilon_v = \varepsilon_0/2^v$, $v = 0, 1, 2, \dots$, для которого выполняется неравенство (6.5). В силу леммы 6.2 $\hat{\varepsilon}_k \xrightarrow{k \rightarrow \infty} 0$. Отсюда следует, что с ростом k число испытаний (сравнений $\psi_{\varepsilon_v}(X_k)$ с $-\frac{\rho_0}{\varepsilon_0} \varepsilon_v$, $v = 0, 1, 2, \dots$) для выбора $\hat{\varepsilon}_k$ возрастает.

Для нахождения $\hat{\varepsilon}_k$, $k = 1, 2, \dots$, может быть использовано следующее соображение: испытания можно на-

чинать не с ε_0 , а с $\hat{\varepsilon}_{k-1}$, полученного на предыдущем шаге.

Если

$$\psi_{\varepsilon_{k-1}}(X_k) > -\frac{\rho_0}{\varepsilon_0} \hat{\varepsilon}_{k-1},$$

то перебираем последовательность $\hat{\varepsilon}_{k-1}/2^v$, $v = 1, 2, \dots$, пока не дойдем до $\hat{\varepsilon}_k$. Если же

$$\psi_{\varepsilon_{k-1}}(X_k) \leq -\frac{\rho_0}{\varepsilon_0} \hat{\varepsilon}_{k-1},$$

то найдем наибольшее $v \in [0 : \log_2 \frac{\varepsilon_0}{\hat{\varepsilon}_{k-1}}]$, при котором остается справедливым неравенство

$$\psi_{\varepsilon_{k-1}2^v}(X_k) \leq -\frac{\rho_0}{\varepsilon_0} \hat{\varepsilon}_{k-1}2^v.$$

Соответствующее $\hat{\varepsilon}_{k-1}2^v$ и есть $\hat{\varepsilon}_k$.

5*. В связи с предыдущим замечанием опишем одну модификацию первого метода последовательных приближений.

Пусть уже найдено k -е приближение X'_k , $k = 0, 1, 2, \dots$, и $\psi(X'_k) < 0$. Для построения X'_{k+1} предлагается перебирать последовательность $\bar{\varepsilon}_{k-1}/2^v$, $v = 0, 1, 2, \dots$; $\bar{\varepsilon}_{-1} = \varepsilon_0$. Через $\bar{\varepsilon}_k$ обозначается первый член этой последовательности, для которого

$$\psi_{\bar{\varepsilon}_k}(X'_k) \leq -\frac{\rho_0}{\varepsilon_0} \bar{\varepsilon}_k.$$

В остальном построения такие же, как в основном методе: полагаем $g'_k = g_{\bar{\varepsilon}_k}(X'_k)$ и в качестве X'_{k+1} берем точку минимума функции $\varphi(X)$ на луче $X = X'_k + \alpha g'_k$, $\alpha \geq 0$. Очевидно, $\bar{\varepsilon}_k \leq \bar{\varepsilon}_{k-1}$ и $\varphi(X'_{k+1}) < \varphi(X'_k)$.

Если последовательность $\{X'_k\}$ бесконечна, то

$$\bar{\varepsilon}_k \rightarrow 0 \quad \text{при} \quad k \rightarrow \infty. \quad (6.12)$$

Доказательство этого утверждения дословно совпадает с доказательством леммы 6.2.

Теорема 6.2. Пусть $\{X'_{k_p}\}$ — такая подпоследовательность последовательности $\{X'_k\}$, что

$$\bar{\varepsilon}_{k_p} < \bar{\varepsilon}_{k_{p-1}}. \quad (6.13)$$

Если $M(X'_0)$ — ограниченное множество, то любая предельная точка подпоследовательности $\{X'_{k_p}\}$ является стационарной точкой функции $\varphi(X)$.

Доказательство. Заметим, во-первых, что неравенство (6.13) равносильно следующему неравенству:

$$\psi_{\bar{\varepsilon}_{k_p-1}}(X'_{k_p}) > -\frac{\rho_0}{\varepsilon_0} \bar{\varepsilon}_{k_p-1}. \quad (6.14)$$

Далее в силу (6.12) подпоследовательность $\{X'_{k_p}\}$ бесконечна. Без ограничения общности будем считать, что вся подпоследовательность $\{X'_{k_p}\}$ сходится:

$$X'_{k_p} \xrightarrow{k_p \rightarrow \infty} X^*.$$

Требуется доказать, что X^* — стационарная точка функции $\varphi(X)$. Предположим противное. Пусть

$$\psi(X^*) = -b < 0.$$

Повторяя рассуждения, приведенные при доказательстве теоремы 6.1, получим, что найдутся такие $K > 0$ и $\gamma^* > 0$, что при $k_p > K$ и $\varepsilon \in [0, \gamma^*]$ будет

$$\psi_\varepsilon(X'_{k_p}) \leq -\frac{b}{2}. \quad (6.15)$$

Из (6.12) и (6.15) следует, что при достаточно больших k_p справедливо неравенство

$$\psi_{\bar{\varepsilon}_{k_p-1}}(X'_{k_p}) \leq -\frac{\rho_0}{\varepsilon_0} \bar{\varepsilon}_{k_p-1},$$

что, однако, противоречит (6.14).

Теорема доказана.

Следствие. Если функция максимума $\varphi(X)$ выпукла, а множество $M(X'_0)$ ограничено, то любая предельная точка последовательности $\{X'_k\}$, построенной модифицированным методом, является точкой минимума функции $\varphi(X)$.

Действительно, последовательность $\{\varphi(X'_k)\}$ сходится:

$$\varphi(X'_k) \xrightarrow{k \rightarrow \infty} \varphi^*.$$

В силу теорем 6.2 и 3.1 некоторая подпоследовательность последовательности $\{X'_k\}$ сходится к точке ми-

нимума функции $\varphi(X)$. Отсюда и из непрерывности $\varphi(X)$ следует, что

$$\varphi^* = \min_{X \in E_n} \varphi(X) \stackrel{df}{=} \mu.$$

Итак, $\varphi(X'_k) \xrightarrow{k \rightarrow \infty} \mu$.

Для любой предельной точки X^* последовательности $\{X'_k\}$ имеем

$$\varphi(X^*) = \mu.$$

Значит, X^* является точкой минимума функции $\varphi(X)$.
Утверждение доказано.

§ 7. ε -стационарные точки. Второй метод последовательных приближений

1. Напомним, что точка $X^* \in E_n$ называется ε -стационарной точкой функции $\varphi(X)$, если

$$\psi_\varepsilon(X^*) \geq 0,$$

или, что то же самое,

$$\min_{\|g\|=1} \max_{i \in R_\varepsilon(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) \geq 0.$$

Выясним смысл этого определения. Пусть X^* — ε -стационарная точка и функции $f_i(X)$ выпуклы на E_n при $i \in R_\varepsilon(X^*)$. Тогда выполняются неравенства

$$0 \leq \varphi(X^*) - \mu \leq \varepsilon,$$

где $\mu = \inf_{X \in E_n} \varphi(X)$. Действительно, в силу теоремы 4.3

$$\min_{i \in R_\varepsilon(X^*)} f_i(X^*) \leq \mu \leq \varphi(X^*).$$

Однако, для $i \in R_\varepsilon(X^*)$

$$\varphi(X^*) - f_i(X^*) \leq \varepsilon.$$

Теперь имеем

$$0 \leq \varphi(X^*) - \mu \leq \varphi(X^*) - \min_{i \in R_\varepsilon(X^*)} f_i(X^*) \leq \varepsilon,$$

что и требовалось доказать.

Таким образом, если X^* — ε -стационарная точка и $f_i(X)$, $i \in R_\varepsilon(X^*)$, — выпуклые функции, то $\varphi(X^*)$ есть приближенное значение для μ с абсолютной погрешностью, не превышающей ε .

Зафиксируем $\varepsilon > 0$. Нашей ближайшей целью является описание метода последовательных приближений для нахождения ε -стационарных точек.

В качестве начального приближения возьмем произвольную точку $X_0 \in E_n$. Предположим, что множество $M(X_0)$ ограничено. Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $\psi_\varepsilon(X_k) \geq 0$, то X_k — ε -стационарная точка, и процесс на этом заканчивается. Предположим, что $\psi_\varepsilon(X_k) < 0$. Тогда полагаем $g_k = g_\varepsilon(X_k)$, где $g_\varepsilon(X_k)$ — направление ε -наискорейшего спуска функции $\varphi(X)$ в точке X_k . На луче

$$X = X_k(\alpha) \stackrel{\text{df}}{=} X_k + \alpha g_k, \quad \alpha \geq 0,$$

найдем точку $X_k(\alpha_k)$, для которой

$$\varphi(X_k(\alpha_k)) = \min_{\alpha \geq 0} \varphi(X_k(\alpha)).$$

В качестве $(k+1)$ -го приближения X_{k+1} возьмем точку $X_k(\alpha_k)$:

$$X_{k+1} = X_k(\alpha_k).$$

Ясно, что $X_{k+1} \in M(X_0)$ и что $\varphi(X_{k+1}) < \varphi(X_k)$.

Далее продолжаем аналогично. Если последовательность $\{X_k\}$, построенная описанным выше методом, конечна, то последний ее элемент по построению является ε -стационарной точкой функции $\varphi(X)$. В противном случае справедлива

Теорема 7.1. *Любая предельная точка последовательности $\{X_k\}$ является ε -стационарной точкой функции $\varphi(X)$.*

Предварительно докажем две леммы.

Лемма 7.1. *Для произвольного $X \in E_n$ и произвольного $\varepsilon > 0$ найдется такое $\delta > 0$, что из неравенства $\|X' - X\| < \delta$ будет следовать*

$$R_\varepsilon(X') \subset R_\varepsilon(X).$$

Доказательство. В силу формулы (6.2) существует $\varepsilon' > \varepsilon$, при котором

$$R_{\varepsilon'}(X) = R_\varepsilon(X). \quad (7.1)$$

Пусть $\delta > 0$ таково, что из неравенства $\|X' - X\| < \delta$ следует

$$|f_i(X') - f_i(X)| < \frac{1}{2}(\varepsilon' - \varepsilon)$$

равномерно по $i \in [0 : N]$. Для тех же X' будет выполняться неравенство

$$|\varphi(X') - \varphi(X)| < \frac{1}{2}(\varepsilon' - \varepsilon).$$

Возьмем $i \in R_\varepsilon(X')$. Если $\|X' - X\| < \delta$, то получим

$$\begin{aligned} \varphi(X) - f_i(X) &= [\varphi(X) - \varphi(X')] + [\varphi(X') - f_i(X')] + \\ &+ [f_i(X') - f_i(X)] \leq \frac{1}{2}(\varepsilon' - \varepsilon) + \varepsilon + \frac{1}{2}(\varepsilon' - \varepsilon) = \varepsilon'. \end{aligned}$$

Следовательно, $i \in R_{\varepsilon'}(X)$, и в силу (7.1) $i \in R_\varepsilon(X)$.

Лемма доказана.

Лемма 7.2. Для построенной выше последовательности $\{X_k\}$ выполняется предельное соотношение

$$\lim_{k \rightarrow \infty} \psi_\varepsilon(X_k) \geq 0. \quad (7.2)$$

Доказательство. Допустим противное. Тогда найдется подпоследовательность $\{X_{k_j}\}$ и число $b > 0$ такие, что

$$\psi_\varepsilon(X_{k_j}) \leq -b.$$

Существует столь малое $\alpha_1 > 0$, что для $i \in R_\varepsilon(X_{k_j})$, $\alpha \in [0, \alpha_1]$ и всех k_j будет

$$f_i(X_{k_j} + \alpha g_{k_j}) \leq \varphi(X_{k_j}) - \frac{1}{2}\alpha b,$$

а для $i \notin R_\varepsilon(X_{k_j})$, $\alpha \in [0, \alpha_1]$ и всех k_j

$$f_i(X_{k_j} + \alpha g_{k_j}) \leq \varphi(X_{k_j}) - \frac{1}{2}\varepsilon$$

(см. доказательство леммы 6.2).

Отсюда следует, что равномерно по k_j

$$\varphi(X_{k_j} + \alpha_1 g_{k_j}) \leq \varphi(X_{k_j}) - \beta, \quad (7.3)$$

где $\beta = \min \left\{ \frac{1}{2}\alpha_1 b, \frac{1}{2}\varepsilon \right\}$.

Неравенство (7.3), так же, как в лемме 6.2, приводит нас к противоречию.

Лемма доказана.

Доказательство теоремы. Пусть

$$X_{k_j} \xrightarrow{k_j \rightarrow \infty} X^*.$$

Требуется доказать, что $\psi_\varepsilon(X^*) \geq 0$.

Предположим противное:

$$\psi_\varepsilon(X^*) = -b < 0.$$

В силу леммы 7.1 для достаточно больших номеров $k_j > K_1$ будет

$$R_\varepsilon(X_{k_j}) \subset R_\varepsilon(X^*).$$

Теперь, так же как при доказательстве теоремы 6.1, устанавливаем, что найдется такое натуральное число $K \geq K_1$, что при $k_j > K$

$$\psi_\varepsilon(X_{k_j}) \leq -\frac{b}{2}.$$

Последнее неравенство, однако, противоречит (7.2).

Теорема доказана.

2. Пример. Рассмотрим функцию

$$F(X, t) = (\cos t + \cos(2t + x_1) + \cos(3t + x_2))^2,$$

где $X = (x_1, x_2)$, $t \in [0, 2\pi]$.

Положим

$$t_i = \frac{2\pi i}{N}, \quad i \in [0 : N]; \quad f_i(X) = F(X, t_i).$$

Пусть $N = 100$ и $\varepsilon = 0,001$. Найдем ε -стационарную точку функции

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X)$$

методом, описанным в п. 1. В качестве начального приближения возьмем точку $X_0 = (0,785; 0,785)$. Результаты вычислений приведены в табл. 2. Заметим, что r_k есть расстояние от начала координат до многогранника $L_\varepsilon(X_k)$; s_k — количество индексов i , входящих в $R_\varepsilon(X_k)$.

Таблица 2

k	X_k		$\Phi(X_k)$	r_k	s_k	α_k
0	0,7850	0,7850	8,5762	1,6419	1	1,3530
1	2,1297	0,9347	4,6316	4,8784	1	0,0020
2	2,1281	0,9359	4,6236	4,5216	1	0,0002
3	2,1283	0,9358	4,6228	0,2381	2	0,0578
4	2,1606	0,9837	4,6045	0,0989	3	1,4549
5	2,8870	2,2443	3,9573	3,9563	1	0,0005
6	2,8874	2,2441	3,9558	0,3282	2	0,0766
7	2,8634	2,1713	3,9215	3,8887	1	0,0005
8	2,8639	2,1712	3,9204	0,3377	2	0,0221
9	2,8595	2,1495	3,9120	0,4220	2	0,0005
10	2,8595	2,1491	3,9118	0	3	—

На десятом шаге последовательных приближений мы получили ε -стационарную точку $X^* = (2,8595; 2,1491)$, причем $\Phi(X^*) = 3,9118$.

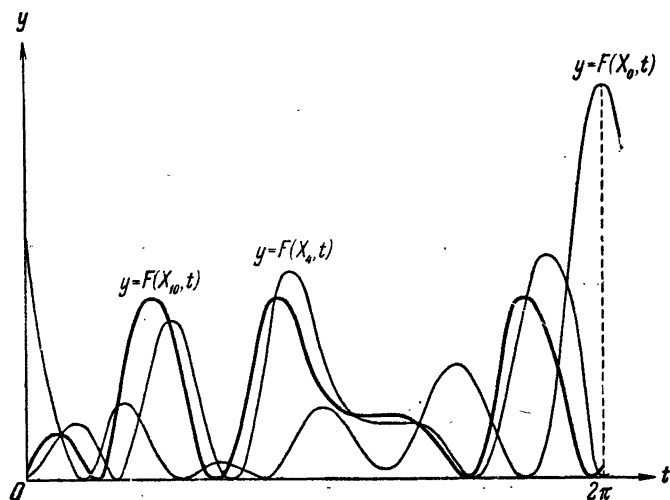


Рис. 22.

На рис. 22 приводится вид функций $F(X_k, t)$ при k , равном 0, 4 и 10.

3*. Опишем одну модификацию изложенного выше метода последовательных приближений для нахождения ε -стационарных точек. При этом будем предполагать, что функции $f_i(X)$, $i \in [0 : N]$, дважды непрерывно дифференцируемы на всем E_n , так что для любых $X \in E_n$ и $Y \in E_n$ справедливо равенство

$$f_i(X + Y) = f_i(X) + \left(\frac{\partial f_i(X)}{\partial X}, Y \right) + \frac{1}{2} \left(\frac{\partial^2 f_i(X + \theta_i Y)}{\partial X^2} Y, Y \right), \quad (7.4)$$

где $\theta_i \in (0, 1)$ и $i \in [0 : N]$.

Зафиксируем $\varepsilon > 0$. Возьмем начальное приближение $X'_0 \in E_n$. Будем предполагать, что множество $M(X'_0)$ ограничено.

Пусть $\alpha > 0$ — некоторая константа,

$$C_1 = \max_{X \in M(X'_0)} \max_{i \in [0 : N]} \left\| \frac{\partial f_i(X)}{\partial X} \right\|,$$

$$C_2 = \max_{X \in M(X'_0), \|\bar{Y}\| \leq \alpha C_1} \max_{i \in [0 : N]} \left\| \frac{\partial^2 f_i(X + \bar{Y})}{\partial X^2} \right\|,$$

где

$$\left\| \frac{\partial^2 f_i(X + \bar{Y})}{\partial X^2} \right\| = \sqrt{\sum_{j_1, j_2=1}^n \left(\frac{\partial^2 f_i(X + \bar{Y})}{\partial x_{j_1} \partial x_{j_2}} \right)^2}.$$

Нетрудно показать, пользуясь неравенством Коши — Буняковского, что для $X \in M(X'_0)$ и $\|\bar{Y}\| \leq \alpha C_1$

$$\left| \left(\frac{\partial^2 f_i(X + \bar{Y})}{\partial X^2} Y, Y \right) \right| \leq C_2 \|Y\|^2.$$

Отсюда и из (7.4) следует, что для всех $i \in [0 : N]$, $X \in M(X'_0)$, $\|Y\| \leq \alpha C_1$ выполняется неравенство

$$f_i(X + Y) \leq f_i(X) + \left(\frac{\partial f_i(X)}{\partial X}, Y \right) + \frac{1}{2} C_2 \|Y\|^2. \quad (7.5)$$

Положим

$$\bar{\alpha} = \min \left\{ \alpha, \frac{2}{C_2}, \frac{-C_1 + \sqrt{C_1^2 + C_2 \varepsilon}}{C_1 C_2} \right\}.$$

(если $C_2 = 0$, то в качестве $\bar{\alpha}$ следует взять число $\bar{\alpha} = \min \left\{ \alpha, \frac{\varepsilon}{2C_1^2} \right\}$).

Зафиксируем некоторое $\alpha_0 \in (0, \bar{\alpha})$.

Пусть уже найдено k -е приближение $X'_k \in M(X'_0)$. Если $\psi_\varepsilon(X'_k) \geq 0$, то X'_k — ε -стационарная точка, и процесс на этом заканчивается.

Если $\psi_\varepsilon(X'_k) < 0$, то полагаем $g'_k = g_\varepsilon(X'_k)$ и

$$\delta_k = -\psi_\varepsilon(X'_k) > 0.$$

В качестве $(k+1)$ -го приближения возьмем точку

$$X'_{k+1} = X'_k + \alpha_0 \delta_k g'_k.$$

Для того чтобы закончить описание метода, осталось доказать, что $X'_{k+1} \in M(X'_0)$. Докажем более точное утверждение:

$$\varphi(X'_{k+1}) < \varphi(X'_k). \quad (7.6)$$

Поскольку $\delta_k \leq C_1$ и $\alpha_0 \leq \alpha$, то в силу (7.5) имеем для всех $i \in [0 : N]$

$$f_i(X'_{k+1}) \leq f_i(X'_k) + \alpha_0 \delta_k \left(\frac{\partial f_i(X'_k)}{\partial X}, g'_k \right) + \frac{1}{2} \alpha_0^2 \delta_k^2 C_2.$$

Пусть $i \in R_\varepsilon(X'_k)$. Тогда

$$\begin{aligned} f_i(X'_{k+1}) &\leq f_i(X'_k) - \alpha_0 \delta_k^2 + \frac{1}{2} \alpha_0^2 \delta_k^2 C_2 \leq \\ &\leq \varphi(X'_k) - \frac{1}{2} \alpha_0 \delta_k^2 C_2 \left(\frac{2}{C_2} - \alpha_0 \right). \end{aligned} \quad (7.7)$$

Учитывая, что $\alpha_0 < \frac{2}{C_2}$, получим для $i \in R_\varepsilon(X'_k)$

$$f_i(X'_{k+1}) < \varphi(X'_k).$$

Для $i \notin R_\varepsilon(X'_k)$ имеем

$$f_i(X'_{k+1}) \leq \varphi(X'_k) - \varepsilon + \alpha_0 C_1^2 + \frac{1}{2} \alpha_0^2 C_1^2 C_2.$$

Поскольку $\alpha_0 < \frac{-C_1 + \sqrt{C_1^2 + C_2 \varepsilon}}{C_1 C_2}$, то

$$\frac{1}{2} \alpha_0^2 C_1^2 C_2 + \alpha_0 C_1^2 \leq \frac{1}{2} \varepsilon.$$

Таким образом, для $i \notin R_\varepsilon(X'_k)$

$$f_i(X'_{k+1}) \leq \varphi(X'_k) - \frac{1}{2} \varepsilon. \quad (7.8)$$

Объединяя (7.7) и (7.8), получим (7.6).

Утверждение доказано.

Рассмотрим последовательность $\{X'_k\}$, построенную модифицированным методом. Если эта последовательность конечна, то последний ее член по построению является ε -стационарной точкой функции $\varphi(X)$. В противном случае справедлива

Теорема 7.2. *Любая предельная точка последовательности $\{X'_k\}$ является ε -стационарной точкой функции $\varphi(X)$.*

Доказательство теоремы проводится аналогично доказательству теоремы 7.1. Пусть

$$X'_{k_p} \xrightarrow{k_p \rightarrow \infty} X^*,$$

однако X^* не является ε -стационарной точкой функции $\varphi(X)$. Тогда

$$\psi_\varepsilon(X^*) = -b < 0.$$

Для достаточно больших номеров $k_p > K$ будет

$$-\delta_k \stackrel{df}{=} \psi_\varepsilon(X'_{k_p}) \leq -\frac{b}{2}.$$

Учитывая (7.7) и (7.8), получим для $k_p > K$

$$\varphi(X'_{k_p+1}) \leq \varphi(X'_{k_p}) - \beta,$$

где $\beta = \min \left\{ \frac{1}{8} \alpha_0 b^2 \left(\frac{2}{C_2} - \alpha_0 \right), \frac{1}{2} \varepsilon \right\}$.

Последнее неравенство, так же как при доказательстве леммы 6.2, приводит нас к противоречию.

Теорема доказана.

4. Покажем, как ε -алгоритм, описанный в п. 1, может быть использован для нахождения стационарных точек функции $\varphi(X)$.

Зафиксируем некоторые $\varepsilon_0 > 0$ и $\rho_0 > 0$. Возьмем начальное приближение $X_{00} \in E_n$. Пусть $M(X_{00})$ — ограниченное множество.

Применяя ε -алгоритм с $\varepsilon = \varepsilon_0$, в конечном (в силу леммы 7.2) число шагов получим точку $X_{0l_0} \in M(X_{00})$ такую, что

$$\psi_{\varepsilon_0}(X_{0l_0}) \geq -\rho_0.$$

Положим $\varepsilon_1 = \frac{1}{2} \varepsilon_0$, $\rho_1 = \frac{1}{2} \rho_0$, $X_{10} = X_{0l_0}$.

Взяв теперь в качестве начального приближения точку X_{10} , снова применим ε -алгоритм с $\varepsilon = \varepsilon_1$. В конечном число шагов получим точку $X_{1l_1} \in M(X_{10}) \subset M(X_{00})$ такую, что

$$\psi_{\varepsilon_1}(X_{1l_1}) \geq -\rho_1.$$

Продолжая описанный процесс, получим последовательность точек $\{X_{kl_k}\}$, причем для членов этой последовательности выполняется неравенство

$$\psi_{\varepsilon_k}(X_{kl_k}) \geq -\rho_k, \quad (7.9)$$

где $\varepsilon_k = \frac{\varepsilon_0}{2^k}$, $\rho_k = \frac{\rho_0}{2^k}$, $k = 0, 1, 2, \dots$

Теорема 7.3. *Любая предельная точка последовательности $\{X_{kl_k}\}$ является стационарной точкой функции $\varphi(X)$.*

Доказательство. Будем для простоты считать, что сходится вся последовательность $\{X_{kl_k}\}$:

$$X_{kl_k} \xrightarrow{k \rightarrow \infty} X^*.$$

Требуется доказать, что $\psi(X^*) \geq 0$. Предположим противное: $\psi(X^*) = -b < 0$. В силу (6.4) найдется $\varepsilon' > 0$ такое, что

$$\psi_{\varepsilon'}(X^*) = -b < 0.$$

Отсюда следует, что для достаточно больших номеров k будет выполняться неравенство

$$\psi_{\varepsilon_k}(X_{kl_k}) = \psi_{\varepsilon'}(X_{kl_k}) \leq -\frac{b}{2},$$

что противоречит (7.9).

Теорема доказана.

§ 8. D-функция. Третий метод последовательных приближений

1. Зафиксируем $\bar{\varepsilon} > 0$ и положим для всех $X \in E_n$

$$D(X) = \inf_{\varepsilon \in [0, \bar{\varepsilon}]} \varepsilon \psi_{\varepsilon}(X). \quad (8.1)$$

Напомним (см. § 6, п. 2), что

$$\psi_{\varepsilon}(X) = \psi_{a_s(X)}(X), \quad \text{если} \quad a_s(X) \leq \varepsilon < a_{s+1}(X),$$

причем

$$0 = a_0(X) < a_1(X) < \dots < a_m(X) < a_{m+1}(X) = \infty.$$

Обозначим через $\bar{m}(X)$ наибольший из индексов s , $s \in [0; m]$, для которых

$$a_s(X) < \bar{\varepsilon}.$$

Если $\psi(X) < 0$, то получим следующее представление для $D(X)$ (рис. 23):

$$D(X) = \min \{a_1 \psi_{a_0}(X), \dots, a_{\bar{m}} \psi_{a_{\bar{m}-1}}(X), \bar{\varepsilon} \psi_{a_{\bar{m}}}(X)\}, \quad (8.2)$$

где $a_s = a_s(X)$, $s \in [0; \bar{m}]$; $\bar{m} = \bar{m}(X)$.

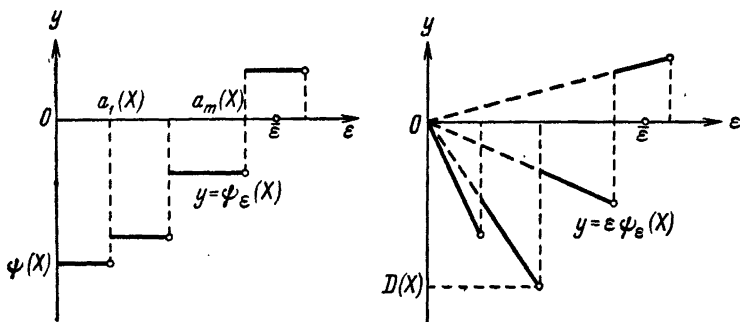


Рис. 23.

Из (8.1) следует также, что для любого $X \in E_n$

$$D(X) \leq 0. \quad (8.3)$$

Лемма 8.1. Неравенство $\psi(X) \geq 0$ эквивалентно равенству

$$D(X) = 0.$$

Доказательство. Если $\psi(X) \geq 0$, то, очевидно, инфимум в (8.1) достигается при $\varepsilon = 0$. Следовательно, $D(X) = 0$.

Пусть теперь $D(X) = 0$. Тогда при $0 < \varepsilon \leq \bar{\varepsilon}$

$$\psi_{\varepsilon}(X) \geq 0.$$

Поскольку для $0 \leq \varepsilon < a_1(X)$

$$\psi(X) = \psi_{\varepsilon}(X),$$

то получаем, что $\psi(X) \geq 0$.

Лемма доказана.

Теорема 8.1. *Для того чтобы функция максимума $\varphi(X)$ достигала минимума в точке X^* , необходимо, а в случае выпуклости $\varphi(X)$ и достаточно, чтобы выполнялось равенство*

$$D(X^*) = 0. \quad (8.4)$$

Эта теорема является очевидным следствием теоремы 3.1 и предыдущей леммы.

Учитывая (8.3) и лемму 8.1, заключаем, что задача отыскания стационарных точек функции $\varphi(X)$ сведена к задаче максимизации функции $D(X)$.

2. Нашей ближайшей целью является доказательство непрерывности функции $D(X)$ на всем E_n .

Возьмем любое γ , удовлетворяющее неравенствам

$$0 < \gamma < \frac{1}{4} \min_{s \in [0:m]} (a_{s+1}(X) - a_s(X)), \quad (8.5)$$

и пусть

$$\Gamma'_{\gamma}(X) = [0, \infty) \setminus \bigcup_{k=0}^m (a_k(X) - \gamma, a_k(X) + \gamma).$$

Лемма 8.2. *Для произвольного $X \in E_n$ и любого γ , удовлетворяющего неравенствам (8.5), найдется такое $\delta > 0$, что для любого $X' \in E_n$, $\|X' - X\| < \delta$, и любого $\varepsilon \in \Gamma'_{\gamma}(X)$ будет*

$$R_{\varepsilon}(X') = R_{\varepsilon}(X).$$

Доказательство. Пусть для определенности

$$a_{\bar{k}}(X) + \gamma \leq \varepsilon \leq a_{\bar{k}+1}(X) - \gamma.$$

Число $\delta > 0$ выберем так, чтобы из неравенства $\|X' - X\| < \delta$ следовало

$$|f_i(X') - f_i(X)| < \frac{1}{4} \gamma$$

равномерно по $i \in [0 : N]$. В этом случае выполняется также неравенство

$$|\varphi(X') - \varphi(X)| < \frac{1}{4} \gamma.$$

Пусть $i \in R_\varepsilon(X')$, причем $\|X' - X\| < \delta$. Тогда $\varphi(X) - f_i(X) =$

$$\begin{aligned} &= [\varphi(X) - \varphi(X')] + [\varphi(X') - f_i(X')] + [f_i(X') - f_i(X)] \leq \\ &\leq \frac{1}{4} \gamma + \varepsilon + \frac{1}{4} \gamma \leq a_{\bar{k}+1}(X) - \frac{1}{2} \gamma. \end{aligned}$$

Замечая, что $R_{a_{\bar{k}+1}(X) - 1/2\gamma}(X) = R_\varepsilon(X)$, получим: $i \in R_\varepsilon(X)$.

Следовательно,

$$R_\varepsilon(X') \subset R_\varepsilon(X). \quad (8.6)$$

Пусть теперь $i \in R_\varepsilon(X)$. Учитывая равенство

$$R_\varepsilon(X) = R_{a_{\bar{k}}(X) + 1/2\gamma}(X),$$

получим для $\|X' - X\| < \delta$

$$\begin{aligned} \varphi(X') - f_i(X') &= \\ &= [\varphi(X') - \varphi(X)] + [\varphi(X) - f_i(X)] + [f_i(X) - f_i(X')] \leq \\ &\leq \frac{1}{4} \gamma + a_{\bar{k}}(X) + \frac{1}{2} \gamma + \frac{1}{4} \gamma = a_{\bar{k}}(X) + \gamma \leq \varepsilon. \end{aligned}$$

Значит, $i \in R_\varepsilon(X')$, так что

$$R_\varepsilon(X) \subset R_\varepsilon(X'). \quad (8.7)$$

Объединяя (8.6) и (8.7), получим требуемое.

Лемма доказана.

Положим

$$\Gamma_\gamma(X) = \Gamma'_\gamma(X) \cap [0, \bar{\varepsilon}],$$

где γ удовлетворяет неравенствам (8.5) и $\gamma < \bar{\varepsilon}$. Очевидно, $\Gamma_\gamma(X)$ — непустое ограниченное замкнутое множество.

Введем обозначение

$$D_{\gamma}(X, X') = \inf_{\varepsilon \in \Gamma_{\gamma}(X)} \varepsilon \psi_{\varepsilon}(X').$$

Лемма 8.3. Для фиксированных X и γ по любому $\varepsilon^* > 0$ найдется такое $\delta > 0$, что для всех $X' \in E_n$, $\|X' - X\| < \delta$, будет

$$|D_{\gamma}(X, X') - D_{\gamma}(X, X)| < \varepsilon^*.$$

Доказательство. Учитывая лемму 8.2, получим для X' , $\|X' - X\| < \delta_1$ (см. также неравенство VIII из § 2 Приложения III)

$$\begin{aligned} |D_{\gamma}(X, X') - D_{\gamma}(X, X)| &\leq \sup_{\varepsilon \in \Gamma_{\gamma}(X)} \varepsilon |\psi_{\varepsilon}(X') - \psi_{\varepsilon}(X)| \leq \\ &\leq \bar{\varepsilon} \sup_{\varepsilon \in \Gamma_{\gamma}(X)} \left| \min_{\|g\|=1} \max_{i \in R_{\varepsilon}(X)} \left(\frac{\partial f_i(X')}{\partial X}, g \right) - \right. \\ &\quad \left. - \min_{\|g\|=1} \max_{i \in R_{\varepsilon}(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right) \right| \leq \\ &\leq \bar{\varepsilon} \sup_{\varepsilon \in \Gamma_{\gamma}(X)} \max_{\|g\|=1} \max_{i \in R_{\varepsilon}(X)} \left| \left(\frac{\partial f_i(X')}{\partial X} - \frac{\partial f_i(X)}{\partial X}, g \right) \right| \leq \\ &\leq \bar{\varepsilon} \max_{i \in [0:N]} \left\| \frac{\partial f_i(X')}{\partial X} - \frac{\partial f_i(X)}{\partial X} \right\|. \end{aligned}$$

Выберем теперь $\delta > 0$, $\delta \leq \delta_1$, так, чтобы для X' , $\|X' - X\| < \delta$, выполнялось неравенство

$$\bar{\varepsilon} \max_{i \in [0:N]} \left\| \frac{\partial f_i(X')}{\partial X} - \frac{\partial f_i(X)}{\partial X} \right\| < \varepsilon^*.$$

В этом случае для тех же X' будет

$$|D_{\gamma}(X, X') - D_{\gamma}(X, X)| < \varepsilon^*.$$

Лемма доказана.

Лемма 8.4. Для произвольного $X' \in E_n$ справедливо неравенство

$$0 \leq D_{\gamma}(X, X') - D(X') \leq 3\gamma C_1(X'),$$

$$\text{где } C_1(X') = \max_{i \in [0:N]} \left\| \frac{\partial f_i(X')}{\partial X} \right\|.$$

Доказательство. Для $\varepsilon \in \Gamma_\gamma(X)$ имеем

$$\varepsilon \psi_\varepsilon(X') \geq \inf_{\varepsilon \in \Gamma_\gamma(X)} \varepsilon \psi_\varepsilon(X') \stackrel{df}{=} D_\gamma(X, X'). \quad (8.8)$$

Пусть теперь $\varepsilon \in [0, \bar{\varepsilon}]$, но $\varepsilon \notin \Gamma_\gamma(X)$. Рассмотрим вначале случай

$$0 < \varepsilon < \gamma.$$

Имеем

$$\begin{aligned} \varepsilon \psi_\varepsilon(X') &= [\varepsilon \psi_\varepsilon(X') - \varepsilon \psi_\gamma(X')] + \\ &\quad + [\varepsilon \psi_\gamma(X') - \gamma \psi_\gamma(X')] + \gamma \psi_\gamma(X') \geq \\ &\geq \gamma \psi_\gamma(X') - |(\varepsilon - \gamma) \psi_\gamma(X')| - \varepsilon |\psi_\varepsilon(X') - \psi_\gamma(X')| \geq \\ &\geq \inf_{\varepsilon \in \Gamma_\gamma(X)} \varepsilon \psi_\varepsilon(X') - \gamma C_1(X') - \gamma(2C_1(X')) = \\ &= D_\gamma(X, X') - 3\gamma C_1(X'). \end{aligned} \quad (8.9)$$

Если $a_{\bar{k}}(X) - \gamma < \varepsilon < a_{\bar{k}}(X) + \gamma$, то, обозначив

$$\hat{\varepsilon} = a_{\bar{k}}(X) - \gamma \in \Gamma_\gamma(X),$$

получим

$$\begin{aligned} \varepsilon \psi_\varepsilon(X') &\geq \varepsilon \psi_\varepsilon(X') = \hat{\varepsilon} \psi_\varepsilon(X') + (\varepsilon - \hat{\varepsilon}) \psi_\varepsilon(X') \geq \\ &\geq D_\gamma(X, X') - |(\varepsilon - \hat{\varepsilon}) \psi_\varepsilon(X')| \geq D_\gamma(X, X') - 2\gamma C_1(X'). \end{aligned} \quad (8.10)$$

Итак, в силу (8.8), (8.9) и (8.10) имеем для всех $\varepsilon \in [0, \bar{\varepsilon}]$

$$\varepsilon \psi_\varepsilon(X') \geq D_\gamma(X, X') - 3\gamma C_1(X').$$

Значит, $D_\gamma(X, X') - D(X') \leq 3\gamma C_1(X')$. Неравенство $D_\gamma(X, X') - D(X') \geq 0$ очевидно.

Лемма доказана.

Теорема 8.2. *Функция $D(X)$ непрерывна на всем пространстве E_n .*

Доказательство. Зафиксируем $X \in E_n$ и возьмем произвольные $\varepsilon^* > 0$ и $\Delta > 0$.

Обозначим

$$C_1 = \max_{\|X' - X\| \leq \Delta} \max_{i \in [0: N]} \left\| \frac{\partial f_i(X')}{\partial X} \right\|.$$

Выберем теперь γ так, чтобы

$$3\gamma C_1 < \frac{1}{3} \varepsilon^*. \quad (8.11)$$

Согласно лемме 8.3 найдется δ , $0 < \delta < \Delta$, такое, что для всех X' , $\|X' - X\| < \delta$, будет

$$|D_{\gamma}(X, X') - D_{\gamma}(X, X)| < \frac{1}{3} \varepsilon^*. \quad (8.12)$$

Учитывая лемму 8.4 и неравенства (8.11) и (8.12), окончательно имеем для всех X' , $\|X' - X\| < \delta$,

$$\begin{aligned} |D(X') - D(X)| &\leq |D(X') - D_{\gamma}(X, X')| + \\ &+ |D_{\gamma}(X, X') - D_{\gamma}(X, X)| + |D_{\gamma}(X, X) - D(X)| < \\ &< \frac{1}{3} \varepsilon^* + \frac{1}{3} \varepsilon^* + \frac{1}{3} \varepsilon^* = \varepsilon^*. \end{aligned}$$

Теорема доказана.

3. Переходим к описанию метода последовательных приближений для нахождения стационарных точек функции $\varphi(X)$, использующего D -функцию.

Возьмем начальное приближение $X_0 \in E_n$. Предположим, что множество $M(X_0) = \{X | \varphi(X) \leq \varphi(X_0)\}$ ограничено.

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $D(X_k) = 0$, то в силу леммы 8.1 X_k — стационарная точка, и процесс на этом заканчивается.

Предположим, что $D(X_k) < 0$. Учитывая (8.2), запишем

$$D(X_k) = \varepsilon'_k \psi_{\varepsilon''_k}(X_k), \quad (8.13)$$

где $(\varepsilon'_k, \varepsilon''_k)$ — одна из пар

$$(a_1(X_k), a_0(X_k)), \dots, (a_{\bar{m}}(X_k), a_{\bar{m}-1}(X_k)), (\bar{e}, a_{\bar{m}}(X_k)).$$

Поскольку, по предположению, $D(X_k) < 0$, то $\psi_{\varepsilon''_k}(X_k) < 0$.

Пусть $g_k = g_{\varepsilon''_k}(X_k)$ — направление ε''_k — наискорейшего спуска функции $\varphi(X)$ в точке X_k . Рассмотрим луч

$$X = X_k(\alpha) \stackrel{\text{дф}}{=} X_k + \alpha g_k, \quad \alpha \geq 0,$$

и через $X_k(\alpha_k)$ обозначим точку, для которой

$$\varphi(X_k(\alpha_k)) = \min_{\alpha \in [0, \infty)} \varphi(X_k(\alpha)).$$

В качестве $(k+1)$ -го приближения X_{k+1} возьмем точку $X_k(\alpha_k)$:

$$X_{k+1} = X_k(\alpha_k).$$

Как обычно, показывается, что $\varphi(X_{k+1}) < \varphi(X_k)$, так что, в частности, $X_{k+1} \in M(X_0)$.

Если последовательность $\{X_k\}$, построенная описанным выше методом, конечна, то последний ее член по построению является стационарной точкой функции $\varphi(X)$. В противном случае справедлива

Теорема 8.3. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$.*

Доказательство. Пусть

$$X_{k_j} \xrightarrow{k_j \rightarrow \infty} X^*.$$

Утверждается, что $D(X^*) = 0$. Допустим противное:

$$D(X^*) = -b < 0.$$

В силу непрерывности $D(X)$ и (8.13) получим для достаточно больших номеров $k_j > K_1$

$$D(X_{k_j}) = \varepsilon'_{k_j} \psi_{\varepsilon''_{k_j}}(X_{k_j}) \leq -\frac{b}{2}.$$

Поскольку для $\varepsilon \in [\varepsilon''_{k_j}, \varepsilon'_{k_j})$ выполняется равенство

$$\psi_\varepsilon(X_{k_j}) = \psi_{\varepsilon''_{k_j}}(X_{k_j}),$$

то, положив $\tilde{\varepsilon}''_{k_j} = (\varepsilon'_{k_j} + \varepsilon''_{k_j})/2$, получим

$$\varepsilon'_{k_j} \psi_{\tilde{\varepsilon}''_{k_j}}(X_{k_j}) \leq -\frac{b}{2}, \quad (8.14)$$

причем

$$\frac{\varepsilon'_{k_j}}{2} \leq \tilde{\varepsilon}''_{k_j} \leq \varepsilon'_{k_j}. \quad (8.15)$$

Без ограничения общности можно считать, что

$$\varepsilon'_{k_j} \rightarrow \varepsilon', \quad \tilde{\varepsilon}''_{k_j} \rightarrow \tilde{\varepsilon}'',$$

причем $\varepsilon' > 0$ и в силу (8.15) $\tilde{\varepsilon}'' > 0$.

Выберем $K \geq K_1$ так, чтобы для $k_j > K$ выполнялись неравенства

$$\varepsilon'_{k_j} \leq 2\varepsilon', \quad \tilde{\varepsilon}''_{k_j} \geq \frac{\tilde{\varepsilon}''}{2}.$$

В силу (8.14) для тех же k_j получим

$$\psi_{\varepsilon''/2}(X_{k_j}) \leq -\frac{b}{4\varepsilon'}.$$

Последнее неравенство вместе с соотношениями

$$R_{\varepsilon''/2}(X_{k_j}) \subset R_{\varepsilon''_{k_j}}(X_{k_j}), \quad R_{\varepsilon''_{k_j}}(X_{k_j}) = R_{\varepsilon''_{k_j}}(X_{k_j})$$

обычным путем приводит нас к противоречию.

Теорема доказана.

З а м е ч а н и е 1. Определение стационарных точек X^* функции $\varphi(X)$ с помощью D -функции ($D(X^*)=0$) выгодно отличается от определения стационарных точек с помощью ψ -функции ($\psi(X^*) \geq 0$) тем, что функция $D(X)$ непрерывна на E_n , в то время как функция $\psi(X)$ именно в стационарных точках может быть разрывной (см. замечание в конце § 5).

З а м е ч а н и е 2. Функция $D(X)$ может быть определена в более общем виде

$$D(X) = \inf_{\varepsilon \in [0, \bar{\varepsilon}]} h(\varepsilon) \psi_{\varepsilon}(X),$$

где $h(\varepsilon)$ — непрерывная на $[0, \bar{\varepsilon}]$ строго возрастающая функция, обращающаяся в нуль при $\varepsilon=0$.

В этом случае сохраняются непрерывность функции $D(X)$ и сходимость метода последовательных приближений, построенного с помощью новой D -функции.

4. Найдем с помощью D -метода стационарную точку функции $\varphi(X)$, рассмотренной в примере 2 § 5. Напомним, что

$$X = (x_1, x_2) \in E_2,$$

$$f_1(X) = -5x_1 + x_2, \quad f_2(X) = 4x_2 + x_1^2 + x_2^2, \quad f_3(X) = 5x_1 + x_2,$$

$$\varphi(X) = \max_{i \in [1:3]} f_i(X).$$

Положим $\bar{\varepsilon}=2$. В качестве начального приближения возьмем точку $X_0=(-0,1; 0,2)$. Результаты вычислений приведены в табл. 3. Полезно заметить, что X_1 находится на дуге, начиная с которой метод наискорейшего спуска сходится к точке $\bar{X}=(0, 0)$, не являющейся стационарной.

Таблица 3

k	$X_k = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix}$	$f_1(X_k)$ $f_2(X_k)$ $f_3(X_k)$	$a_1(X_k)$ $a_2(X_k)$	$R(X_k)$ $R_{a_1}(X_k)$ $R_{a_2}(X_k)$	$\psi(X_k)$ $\psi_{a_1}(X_k)$ $\psi_{a_2}(X_k)$	$a_1\psi(X_k)$ $a_2\psi_{a_1}(X_k)$ $a_2\psi_{a_2}(X_k)$	$D(X_k)$	$R_{\varepsilon_k}''(X_k)$	$g_k = \begin{pmatrix} g_1^{(k)} \\ g_2^{(k)} \end{pmatrix}$	a_k
0	-0,10000 0,20000	0,700000 0,850000 -0,300000	0,150000 1,150000	{2} {1, 2} {1, 2, 3}	-4,404500 -3,700000 -1,000000	-0,660675 -4,255000 -2,000000	-4,255000	{1, 2}	0,578017 -0,816025	0,206321
1	0,019257 0,031648	-0,064636 0,127962 0,127932	0,000030 0,192598	{2} {2, 3} {1, 2, 3}	-4,063962 -3,477632 -1,000000	-0,090122 -0,669785 -2,000000	-2,000000	{1, 2, 3}	0,000000 -1,000000	3,063275
2	0,019257 -3,031627	-3,127911 -2,935374 -2,935343	0,000031 0,192568	{3} {2, 3} {1, 2, 3}	-5,099020 -1,775831 1,000000	-0,000158 -0,341968 2,000000	-0,341968	{2, 3}	-0,525345 0,850889	0,036909
3	-0,000133 -3,000222	-2,999557 -2,999554 -3,000885	0,000003 0,001333	{2} {1, 2} {1, 2, 3}	-2,000445 -1,715513 1,000000	-0,000006 -0,002287 2,000000	-0,002287	{1, 2}	0,514590 0,857437	0,000258
4	0,000000 -3,000000	-3,000000 -3,000000 -3,000000	-	$R(X_4) =$ $= \{1, 2, 3\}$	$\psi(X_4) =$ $= 1,000000$	-	0,000000	-	-	-

Таким образом, с помощью D -метода мы не только обошли точку $\bar{X}$, сыгравшую роковую роль в методе наискорейшего спуска, но за четыре шага получили с точностью до шести знаков после запятой стационарную точку $X^* = (0, -3)$.

Процесс поиска стационарной точки иллюстрируется на рис. 24.

5. Поясним смысл D -метода. Предположим, что функции $f_i(X)$ линейны, т. е. что

$$f_i(X) = (A_i, X) + b_i, \quad i \in [0 : N].$$

Тогда $\frac{\partial f_i(X)}{\partial X} = A_i$. Пусть

$$\max_{i \in [0 : N]} \|A_i\| = M.$$

Рассмотрим точку X_k , для которой $D(X_k) < 0$. Для каждого $a_s = a_s(X_k)$, $0 = a_0 < a_1 < \dots < a_m$, может быть построен вектор $g_k^{(s)} = g_{a_s}(X_k)$ такой, что $\|g_k^{(s)}\| = 1$ и

$$\psi_{a_s}(X_k) = \max_{i \in R_{a_s}(X_k)} (A_i, g_k^{(s)}). \quad (8.16)$$

В D -методе для спуска выбирается тот вектор $g_k^{(s)}$, для которого произведение $a_{s+1}\psi_{a_s}(X_k)$ минимально ($s \in [0 : \bar{m}]$, $a_{\bar{m}+1} = \bar{e}$). Для оправдания этого выбора приведем следующие соображения.

Поскольку

$$f_i(X_k + \alpha g_k^{(s)}) = f_i(X_k) + \alpha (A_i, g_k^{(s)}),$$

то для $i \in R(X_k)$ имеем

$$f_i(X_k + \alpha g_k^{(s)}) \geq \varphi(X_k) - \alpha M.$$

Отсюда следует, что

$$\max_{i \in R_{a_s}(X_k)} f_i(X_k + \alpha g_k^{(s)}) \geq \varphi(X_k) - \alpha M. \quad (8.17)$$

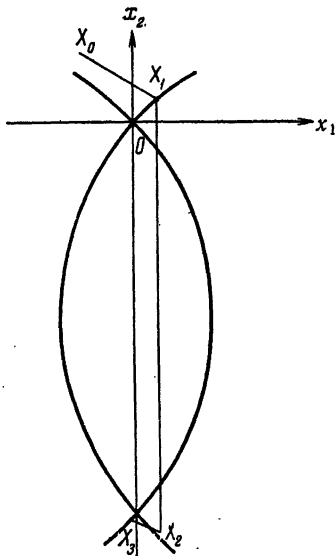


Рис. 24.

Рассмотрим $i \notin R_{a_s}(X_k)$. Учитывая, что для $\varepsilon \in [a_s, a_{s+1})$

$$R_\varepsilon(X_k) = R_{a_s}(X_k),$$

закключаем, что

$$f_i(X_k + \alpha g_k^{(s)}) \leq \varphi(X_k) - \varepsilon + \alpha M.$$

Поскольку ε может быть взято как угодно близким к a_{s+1} , то для $i \notin R_{a_s}(X_k)$ справедливо неравенство

$$f_i(X_k + \alpha g_k^{(s)}) \leq \varphi(X_k) - a_{s+1} + \alpha M.$$

Отсюда следует, что

$$\max_{i \notin R_{a_s}(X_k)} f_i(X_k + \alpha g_k^{(s)}) \leq \varphi(X_k) - (a_{s+1} - \alpha M). \quad (8.18)$$

Выражения, стоящие в правых частях неравенств (8.17) и (8.18), будут равны при $\alpha_0^{(s)} = \frac{a_{s+1}}{2M}$. Таким образом, для $\alpha \in [0, \alpha_0^{(s)}]$ гарантируется неравенство

$$\max_{i \in R_{a_s}(X_k)} f_i(X_k + \alpha g_k^{(s)}) \geq \max_{i \notin R_{a_s}(X_k)} f_i(X_k + \alpha g_k^{(s)}).$$

В частности,

$$\varphi(X_k + \alpha_0^{(s)} g_k^{(s)}) = \max_{i \in R_{a_s}(X_k)} f_i(X_k + \alpha_0^{(s)} g_k^{(s)}).$$

Далее, учитывая (8.16), получаем

$$\begin{aligned} \max_{i \in R_{a_s}(X_k)} f_i(X_k + \alpha_0^{(s)} g_k^{(s)}) &\leq \varphi(X_k) + \alpha_0^{(s)} \max_{i \in R_{a_s}(X_k)} (A_i, g_k^{(s)}) = \\ &= \varphi(X_k) + \frac{1}{2M} a_{s+1} \psi_{a_s}(X_k). \end{aligned}$$

Окончательно имеем

$$\varphi(X_k + \alpha_0^{(s)} g_k^{(s)}) \leq \varphi(X_k) + \frac{1}{2M} a_{s+1} \psi_{a_s}(X_k).$$

Величина $a_{s+1} \psi_{a_s}(X_k)$ (если она отрицательна) с точностью до множителя $1/2M$ показывает, насколько гарантируется уменьшение $\varphi(X)$, если осуществлять спуск из X_k в направлении $g_k^{(s)}$. Естественно выбрать вектор $g_k^{(s)}$

так, чтобы произведение $\alpha_{s+1}\psi_{\alpha_s}(X_k)$ было минимальным. Заметим, что в этом случае в силу (8.2) будет выполняться неравенство

$$\varphi(X_k + \alpha_0^{(s)} g_k^{(s)}) \leq \varphi(X_k) + \frac{1}{2M} D(X_k).$$

Описанная идея выбора направления спуска и реализуется в D -методе.

§ 9. Заключительные замечания

1. В §§ 6—8 были описаны три метода последовательных приближений для нахождения стационарных точек функции максимума $\varphi(X)$. Первый метод имеет в основном теоретический интерес как непосредственное обобщение метода наискорейшего спуска. Весьма эффективным при малых n и N является третий метод последовательных приближений. Однако наиболее удобным для практического использования является второй метод последовательных приближений. В связи с этим мы более подробно остановимся на основном элементе этого метода — нахождении ε -стационарных точек ($\varepsilon > 0$).

На каждом шаге вычислительного процесса, описанного в п. 1 § 7, приходится решать три вспомогательные задачи:

I. Выяснить, принадлежит ли начало координат многограннику $L_\varepsilon(X_k)$.

II. Если $0 \notin L_\varepsilon(X_k)$, то найти $g_k = g_\varepsilon(X_k)$ — направление ε -наискорейшего спуска функции $\varphi(X)$ в точке X_k .

III. Найти α_k — точку минимума функции $\varphi(X_k + \alpha g_k)$ на полуоси $\alpha \in [0, \infty)$.

Первая вспомогательная задача сводится к задаче линейного программирования. Действительно, пусть k фиксировано и $Z_1, \dots, Z_s$ — точки множества $H_\varepsilon(X_k)$. Любая точка множества $L_\varepsilon(X_k)$ может быть представлена в виде

$$V = \sum_{i=1}^s \alpha_i Z_i; \quad \alpha_i \geq 0, \quad \sum_{i=1}^s \alpha_i = 1.$$

Вопрос о том, принадлежит ли начало координат множеству $L_\varepsilon(X_k)$, равносильен вопросу о разрешимости

системы

$$\sum_{i=1}^s \alpha_i Z_i = 0$$

при дополнительном условии: $\sum_{i=1}^s \alpha_i = 1; \alpha_i \geq 0, i \in [1:s]$.

В связи с этим рассмотрим следующую задачу линейного программирования в пространстве векторов $W = (\alpha_1, \dots, \alpha_s, u) \in E_{s+1}$: найти $\min u$ при условиях

$$\begin{aligned} \sum_{i=1}^s \alpha_i z_k^{(i)} - u &\leq 0, & k \in [1:n]; \\ - \sum_{i=1}^s \alpha_i z_k^{(i)} - u &\leq 0, & k \in [1:n]; \\ \sum_{i=1}^s \alpha_i &= 1; \quad \alpha_i \geq 0, \quad i \in [1:s]; \quad u \geq 0. \end{aligned}$$

Обозначим через $W_0 = (\alpha_1^{(0)}, \dots, \alpha_s^{(0)}, u^{(0)})$ решение этой задачи. Нетрудно понять, что $0 \in L_\varepsilon(X_k)$ тогда и только тогда, когда $u^{(0)} = 0$. Если $u^{(0)} = 0$, то X_k — ε -стационарная точка функции $\varphi(X)$.

Пусть $u^{(0)} > 0$. В этом случае $0 \notin L_\varepsilon(X_k)$, и можно переходить к решению второй вспомогательной задачи. Для дальнейшего будет полезен вектор $\alpha_0 = (\alpha_1^{(0)}, \dots, \alpha_s^{(0)})$, который составлен из первых s координат вектора W_0 .

Вторая вспомогательная задача сводится к нахождению $Z_\varepsilon(X_k)$ — ближайшей к началу координат точки множества $L_\varepsilon(X_k)$, ибо (см. § 6, п. 2)

$$g_\varepsilon(X_k) = -Z_\varepsilon(X_k) / \|Z_\varepsilon(X_k)\|.$$

Для определения $Z_\varepsilon(X_k)$ можно воспользоваться каким-нибудь простым итеративным методом (см. Приложение IV), взяв в качестве начального приближения точку

$$Z_0 = \sum_{i=1}^s \alpha_i^{(0)} Z_i,$$

где $(\alpha_1^{(0)}, \dots, \alpha_s^{(0)})$ — вектор, полученный при решении первой вспомогательной задачи.

Заметим, что в качестве направления спуска g_k не обязательно брать $g_\varepsilon(X_k)$. Достаточно, зафиксировав

параметр $\theta_1 \in (0, 1)$, одинаковый для всех k , взять любую точку V_k из $L_\varepsilon(X_k)$, для которой

$$\frac{\min_{z \in H_\varepsilon(X_k)} (Z, V_k)}{(V_k, V_k)} \geq \theta_1 \quad (9.1)$$

(на рис. 25 отношение, стоящее в левой части (9.1), равно $\frac{B_k O}{V_k O}$) и положить

$$g_k = -\frac{V_k}{\|V_k\|}.$$

Можно показать, что в этом случае теорема 7.1 остается справедливой.

Переходим к третьей вспомогательной задаче. Положим при фиксированном k

$$\Phi_k(\alpha) = \varphi(X_k + \alpha g_k), \quad \alpha \geq 0.$$

Функция $\Phi_k(\alpha)$ непрерывна и имеет непрерывные производные слева и справа в любой точке $\alpha \in (0, \infty)$, причем

$$\Phi'_k(\alpha + 0) = \max_{i \in R(X_k(\alpha))} \left(\frac{\partial f_i(X_k(\alpha))}{\partial X}, g_k \right), \quad (9.2)$$

$$\Phi'_k(\alpha - 0) = \min_{i \in R(X_k(\alpha))} \left(\frac{\partial f_i(X_k(\alpha))}{\partial X}, g_k \right). \quad (9.3)$$

Здесь $X_k(\alpha) = X_k + \alpha g_k$. Поскольку $0 \notin L_\varepsilon(X_k)$, то

$$\Phi'_k(+0) = \max_{i \in R(X_k)} \left(\frac{\partial f_i(X_k)}{\partial X}, g_k \right) < 0. \quad (9.4)$$

Заметим также, что множество

$$M_k = \{\alpha \in [0, \infty) \mid \Phi_k(\alpha) \leq \Phi_k(0)\}$$

ограничено, ибо ограниченным является множество $M(X_0)$ и $X_k \in M(X_0)$.

Требуется найти α_k — точку минимума функции $\Phi_k(\alpha)$ на M_k . Заметим, однако, что для справедливости теоремы 7.1 достаточно в качестве α_k взять любую точку

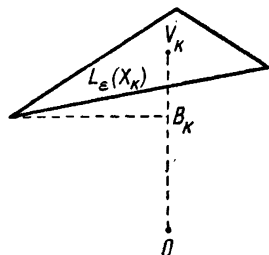


Рис. 25.

из $[0, \infty)$, для которой

$$\frac{\Phi(X_k) - \Phi(X_k(\alpha_k))}{\Phi(X_k) - \min_{\alpha \geq 0} \Phi(X_k(\alpha))} \geq \theta_2, \quad (9.5)$$

где $\theta_2 \in (0, 1)$ — фиксированный для всех k параметр. Доказательство этого утверждения предоставляется читателю.

Пусть $\Phi_k(\alpha)$ — выпуклая на $[0, \infty)$ функция. Покажем, как находить в конечном числе шагов α_k , удовлетворяющее неравенству (9.5). Для этого рассмотрим три числовые последовательности $\{\alpha_{kj}^{(1)}\}$, $\{\alpha_{kj}^{(2)}\}$, $\{\alpha_{kj}^{(3)}\}$ со свойствами:

$$\text{I. } 0 \leq \alpha_{kj}^{(1)} < \alpha_{kj}^{(2)} < \alpha_{kj}^{(3)}.$$

$$\text{II. } \Phi_k(\alpha_{kj}^{(2)}) \leq \Phi_k(\alpha_{kj}^{(1)}),$$

$$\Phi_k(\alpha_{kj}^{(2)}) \leq \Phi_k(\alpha_{kj}^{(3)}).$$

III. Все три последовательности имеют при

$j \rightarrow \infty$ один предел α_k^* (как строить такие последовательности, будет показано ниже).

Положим

$$\Delta_{kj} = \max \{ \Phi'_k(\alpha_{kj}^{(2)} + 0)(\alpha_{kj}^{(2)} - \alpha_{kj}^{(3)}), \Phi'_k(\alpha_{kj}^{(2)} - 0)(\alpha_{kj}^{(2)} - \alpha_{kj}^{(1)}), 0 \}.$$

Очевидно (рис. 26),

$$0 \leq \Phi_k(\alpha_{kj}^{(2)}) - \min_{\alpha \geq 0} \Phi_k(\alpha) \leq \Delta_{kj}. \quad (9.6)$$

В силу (9.2), (9.3) и III

$$\Delta_{kj} \xrightarrow{j \rightarrow \infty} 0. \quad (9.7)$$

Переходя в неравенстве (9.6) к пределу при $j \rightarrow \infty$, получим

$$\Phi_k(\alpha_k^*) = \min_{\alpha \geq 0} \Phi_k(\alpha).$$

В качестве α_k можно взять первый элемент последовательности $\{\alpha_{kj}^{(2)}\}$, для которого

$$(\Phi_k(0) - \Phi_k(\alpha_{kj}^{(2)}))(1 - \theta_2) \geq \theta_2 \Delta_{kj}$$

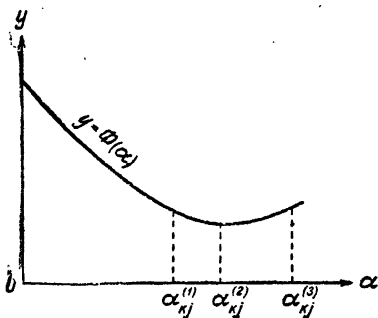


Рис. 26.

(в силу (9.4) и (9.7) такой элемент найдется). В этом случае, как нетрудно проверить, неравенство (9.5) будет выполняться.

Осталось показать, как построить три числовые последовательности $\{\alpha_{kj}^{(1)}\}$, $\{\alpha_{kj}^{(2)}\}$, $\{\alpha_{kj}^{(3)}\}$, удовлетворяющие условиям I—III.

Отметим вначале простой геометрический факт. Пусть на отрезке AC взята точка B (рис. 27) так, что

$$\frac{BC}{AC} = \frac{\sqrt{5}-1}{2} > \frac{1}{2},$$

и пусть D — точка, симметричная B относительно середины отрезка AC . Тогда

$$\frac{DC}{BC} = \frac{AB}{AD} = \frac{\sqrt{5}-1}{2}. \quad (9.8)$$

Теперь возьмем произвольное число $\alpha_0 > 0$. Возможны два случая:

$$\Phi_k(\alpha_0) < \Phi_k(0) \quad \text{и} \quad \Phi_k(\alpha_0) \geq \Phi_k(0).$$

В первом случае перебираем последовательность $\alpha_i = q^i \alpha_0$, $i = 1, 2, \dots$, где $q = \frac{3 + \sqrt{5}}{2}$, пока не дойдем до α_i , при котором $\Phi_k(\alpha_i) \geq \Phi_k(\alpha_{i-1})$ (в силу ограниченности M_k такое α_i найдется). Полагаем

$$\alpha_{k0}^{(1)} = 0, \quad \alpha_{k0}^{(2)} = \alpha_{i-1}, \quad \alpha_{k0}^{(3)} = \alpha_i.$$

Во втором случае перебираем последовательность $\alpha_{-i} = q^{-i} \alpha_0$, $i = 1, 2, \dots$, пока не дойдем до α_{-i} , при котором $\Phi_k(\alpha_{-i}) \leq \Phi_k(0)$ (в силу (9.4) такое α_{-i} найдется). Полагаем

$$\alpha_{k0}^{(1)} = 0, \quad \alpha_{k0}^{(2)} = \alpha_{-i}, \quad \alpha_{k0}^{(3)} = \alpha_{-i+1}.$$

Таким образом, в обоих случаях построены $\alpha_{k0}^{(1)}$, $\alpha_{k0}^{(2)}$, $\alpha_{k0}^{(3)}$, причем $0 = \alpha_{k0}^{(1)} < \alpha_{k0}^{(2)} < \alpha_{k0}^{(3)}$,

$$\begin{aligned} \Phi_k(\alpha_{k0}^{(2)}) &\leq \Phi_k(\alpha_{k0}^{(1)}), & \Phi_k(\alpha_{k0}^{(2)}) &\leq \Phi_k(\alpha_{k0}^{(3)}), \\ \frac{\alpha_{k0}^{(3)} - \alpha_{k0}^{(2)}}{\alpha_{k0}^{(3)} - \alpha_{k0}^{(1)}} &= \frac{\sqrt{5}-1}{2}. \end{aligned}$$

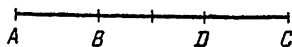


Рис. 27.

Обозначим через α_{k0} точку, симметричную $\alpha_{k0}^{(2)}$ относительно середины отрезка $[\alpha_{k0}^{(1)}, \alpha_{k0}^{(3)}]$. Если $\Phi_k(\alpha_{k0}) < \Phi_k(\alpha_{k0}^{(2)})$, то полагаем

$$\alpha_{k1}^{(1)} = \alpha_{k0}^{(2)}, \quad \alpha_{k1}^{(2)} = \alpha_{k0}, \quad \alpha_{k1}^{(3)} = \alpha_{k0}^{(3)}.$$

В противном случае

$$\alpha_{k1}^{(1)} = \alpha_{k0}^{(1)}, \quad \alpha_{k1}^{(2)} = \alpha_{k0}^{(2)}, \quad \alpha_{k1}^{(3)} = \alpha_{k0}^{(3)}.$$

Заметим, что в силу (9.8)

$$\frac{\alpha_{k1}^{(3)} - \alpha_{k1}^{(1)}}{\alpha_{k0}^{(3)} - \alpha_{k0}^{(1)}} = \frac{\sqrt{5} - 1}{2}$$

и отношение длины большего из отрезков $[\alpha_{k1}^{(1)}, \alpha_{k1}^{(2)}]$, $[\alpha_{k1}^{(2)}, \alpha_{k1}^{(3)}]$ к длине всего отрезка $[\alpha_{k1}^{(1)}, \alpha_{k1}^{(3)}]$ равно $\frac{\sqrt{5} - 1}{2}$.

Этот процесс можно продолжить, привлекая на $(j+1)$ -м шаге точку α_{kj} , симметричную $\alpha_{kj}^{(2)}$ относительно середины отрезка $[\alpha_{kj}^{(1)}, \alpha_{kj}^{(3)}]$. В результате получим три последовательности $\{\alpha_{kj}^{(1)}\}$, $\{\alpha_{kj}^{(2)}\}$, $\{\alpha_{kj}^{(3)}\}$, причем по построению выполнены условия I и II. Поскольку отрезки $[\alpha_{kj}^{(1)}, \alpha_{kj}^{(3)}]$ вложены друг в друга и

$$\alpha_{kj}^{(3)} - \alpha_{kj}^{(1)} = \left(\frac{\sqrt{5} - 1}{2} \right)^j (\alpha_{k0}^{(3)} - \alpha_{k0}^{(1)}),$$

то все три последовательности $\{\alpha_{kj}^{(1)}\}$, $\{\alpha_{kj}^{(2)}\}$, $\{\alpha_{kj}^{(3)}\}$ при $j \rightarrow \infty$ имеют одинаковый предел. Таким образом, условие III также выполнено.

Требуемые последовательности построены.

2. Рассмотренную в § 7 ε -стационарность естественно называть *абсолютной ε -стационарностью*. Введем понятие *относительной ε -стационарности*.

Пусть $\mu = \inf_{X \in E_n} \varphi(X) > 0$. Положим

$$R'_\varepsilon(X) = \{i \in [0 : N] \mid \varphi(X) - f_i(X) \leq \varepsilon \varphi(X)\}, \quad \varepsilon \geq 0.$$

Точку X^* будем называть *относительно ε -стационарной*, если

$$\min_{\|g\|=1} \max_{i \in R'_\varepsilon(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) \geq 0.$$

Если X^* — относительно ε -стационарная точка и функции $f_i(X^*)$, $i \in R'_\varepsilon(X^*)$, выпуклы на E_n , то

$$0 \leq \frac{\varphi(X^*) - \mu}{\varphi(X^*)} \leq \varepsilon.$$

Таким образом, в рассматриваемом случае $\varphi(X^*)$ есть приближенное значение для μ с относительной погрешностью, не превышающей ε .

Для нахождения относительно ε -стационарных точек ($\varepsilon > 0$) может быть развит метод, аналогичный соответствующему методу для нахождения абсолютно ε -стационарных точек (см. § 7, п. 1). При этом остаются справедливыми теоремы типа 7.1 и 7.3.

ДИСКРЕТНАЯ МИНИМАКСНАЯ ЗАДАЧА С ОГРАНИЧЕНИЯМИ НА ПАРАМЕТРЫ

§ 1. Постановка задачи

Пусть $f_i(X)$, $i \in [0 : N]$, $X = (x_1, \dots, x_n)$, — функции, заданные и непрерывно дифференцируемые на некотором открытом множестве $\Omega' \subset E_n$. Предположим также, что задано выпуклое замкнутое (не обязательно ограниченное) множество $\Omega \subset \Omega'$.

Требуется найти точку $X^* \in \Omega$, для которой

$$\max_{i \in [0 : N]} f_i(X^*) = \inf_{X \in \Omega} \max_{i \in [0 : N]} f_i(X).$$

Как и в третьей главе, введем в рассмотрение функцию

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X).$$

Эта функция задана на Ω' . Поставленная задача сводится к минимизации функции $\varphi(X)$ на множестве Ω . По доказанному в § 2 главы III функция $\varphi(X)$ является дифференцируемой по всем направлениям в любой точке множества Ω' ; в частности, везде на Ω . Это позволяет использовать технику, развитую в предыдущей главе.

В данной главе получены необходимые условия минимума функции $\varphi(X)$ на Ω и дана их геометрическая интерпретация. Устанавливаются также достаточные условия локального минимакса.

Рассмотрен один метод последовательных приближений для нахождения стационарных точек, т. е. точек, удовлетворяющих необходимому условию минимума функции $\varphi(X)$ на Ω .

§ 2. Необходимые условия минимакса

1. Предположим, что на некотором открытом множестве $\Omega' \subset E_n$ заданы непрерывно дифференцируемые функции $f_i(X)$, $i \in [0 : N]$. Как и ранее, будем использовать обозначения

$$\begin{aligned}\varphi(X) &= \max_{i \in [0 : N]} f_i(X), \\ R(X) &= \{i \in [0 : N] \mid f_i(X) = \varphi(X)\}.\end{aligned}$$

Пусть Ω — выпуклое замкнутое множество, содержащееся в Ω' . Рассматривается задача минимизации функции $\varphi(X)$ на Ω .

Теорема 2.1. Для того чтобы точка $X^ \in \Omega$ была точкой минимума функции $\varphi(X)$ на множестве Ω , необходимо, а в случае выпуклости $\varphi(X)$ на Ω и достаточно, чтобы*

$$\inf_{Z \in \Omega} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z - X^* \right) = 0. \quad (2.1)$$

Доказательство. Необходимость. Пусть $X^* \in \Omega$ — точка минимума функции $\varphi(X)$ на Ω , однако равенство (2.1) не выполняется. Тогда найдется точка $Z_1 \in \Omega$, для которой

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z_1 - X^* \right) = -\rho < 0 \quad (2.2)$$

(выражение, стоящее в левой части (2.1), положительным быть не может). Ясно, что $Z_1 \neq X^*$. Положим

$$g_1 = \frac{Z_1 - X^*}{\|Z_1 - X^*\|}; \quad \|g_1\| = 1.$$

По формуле (2.7) главы III имеем

$$\varphi(X^* + \alpha g_1) = \varphi(X^*) + \alpha \frac{\partial \varphi(X^*)}{\partial g_1} + o(g_1; \alpha). \quad (2.3)$$

Из (2.2) и формулы для производной по направлению следует, что

$$\frac{\partial \varphi(X^*)}{\partial g_1} = -\frac{\rho}{\|Z_1 - X^*\|}. \quad (2.4)$$

Таким образом, при достаточно малых α получим на основании (2.3) и (2.4)

$$\varphi(X^* + \alpha g_1) \leq \varphi(X^*) - \frac{\alpha \rho}{2 \|Z_1 - X^*\|} < \varphi(X^*). \quad (2.5)$$

Неравенство (2.5) противоречит тому, что X^* — точка минимума функции $\varphi(X)$ на Ω , ибо при всех $\alpha \in [0, \|Z_1 - X^*\|]$ точки $X^* + \alpha g_1$ принадлежат Ω .

Необходимость доказана.

Достаточность. Пусть функция $\varphi(X)$ выпукла на Ω и в точке $X^* \in \Omega$ имеет место (2.1). Докажем, что тогда X^* есть точка минимума функции $\varphi(X)$ на множестве Ω . Допустим противное. Тогда найдется точка $Z_1 \in \Omega$, для которой

$$\varphi(Z_1) < \varphi(X^*). \quad (2.6)$$

Очевидно, что $Z_1 \neq X^*$. Положим $g_1 = \frac{Z_1 - X^*}{\|Z_1 - X^*\|}$, $\|g_1\| = 1$, и найдем

$$\frac{\partial \varphi(X^*)}{\partial g_1} \stackrel{df}{=} \lim_{\alpha \rightarrow +0} \frac{1}{\alpha} [\varphi(X^* + \alpha g_1) - \varphi(X^*)]. \quad (2.7)$$

В силу выпуклости $\varphi(X)$ имеем для $\beta \in [0, 1]$

$$\begin{aligned} \varphi(X^* + \beta(Z_1 - X^*)) &= \varphi(\beta Z_1 + (1 - \beta)X^*) \leq \\ &\leq \beta \varphi(Z_1) + (1 - \beta) \varphi(X^*) = \varphi(X^*) + \beta [\varphi(Z_1) - \varphi(X^*)]. \end{aligned}$$

Отсюда получим для $\alpha \in (0, \|Z_1 - X^*\|)$

$$\begin{aligned} \frac{1}{\alpha} [\varphi(X^* + \alpha g_1) - \varphi(X^*)] &= \\ &= \frac{1}{\alpha} \left[\varphi \left(X^* + \frac{\alpha}{\|Z_1 - X^*\|} (Z_1 - X^*) \right) - \varphi(X^*) \right] \leq \\ &\leq \frac{1}{\|Z_1 - X^*\|} [\varphi(Z_1) - \varphi(X^*)]. \end{aligned} \quad (2.8)$$

Таким образом, в силу (2.6), (2.7) и (2.8)

$$\frac{\partial \varphi(X^*)}{\partial g_1} \leq \frac{\varphi(Z_1) - \varphi(X^*)}{\|Z_1 - X^*\|} < 0.$$

Остается заметить, что

$$\begin{aligned} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z_1 - X^* \right) &= \|Z_1 - X^*\| \frac{\partial \varphi(X^*)}{\partial g_1} \leq \\ &\leq \varphi(Z_1) - \varphi(X^*) < 0. \end{aligned}$$

Полученное неравенство противоречит (2.1).

Теорема доказана.

З а м е ч а н и е 1. Нетрудно проверить, что необходимое условие (2.1) эквивалентно следующему равенству;

$$\inf_{\substack{Z \in \Omega \\ \|Z - X^*\| \leq 1}} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z - X^* \right) = 0. \quad (2.1')$$

В § 3 главы III было показано, что при фиксированном X функция

$$\chi(g) = \max_{i \in R(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right)$$

является непрерывной по g функцией на всем пространстве E_n . Поскольку множество $\Omega(X)$,

$$\Omega(X) = \{g = Z - X \mid Z \in \Omega, \|Z - X\| \leq 1\},$$

при любом фиксированном X является замкнутым и ограниченным, то отсюда следует, что функция $\chi(g)$ достигает своего минимума на $\Omega(X)$.

Учитывая это замечание, заключаем, что инфимум, стоящий в левой части соотношения (2.1'), достигается, и условие (2.1') может быть переписано в виде

$$\min_{\substack{Z \in \Omega \\ \|Z - X^*\| \leq 1}} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z - X^* \right) = 0.$$

З а м е ч а н и е 2. При $N=0$ из теоремы 2.1 следует необходимое условие минимума непрерывно дифференцируемой функции $f_0(X)$ на выпуклом множестве Ω : для того чтобы непрерывно дифференцируемая функция $f_0(X)$ достигала своего минимального на Ω значения в точке $X^* \in \Omega$, необходимо, а в случае выпуклости $f_0(X)$ на Ω и достаточно, чтобы

$$\inf_{Z \in \Omega} \left(\frac{\partial f_0(X^*)}{\partial X}, Z - X^* \right) = 0. \quad (2.9)$$

О п р е д е л е н и е. Точка $X^* \in \Omega$, для которой выполняется соотношение (2.1), называется *стационарной точкой функции* $\varphi(X)$ на Ω .

2. Возьмем любую точку $X \in \Omega$ и зафиксируем ее. Рассмотрим конус $\Gamma(X)$:

$$\Gamma(X) = \{V = \lambda(Z - X) \mid \lambda > 0, Z \in \Omega\}.$$

Очевидно, $0 \in \Gamma(X)$. Замыкание конуса $\Gamma(X)$ называется *конусом возможных направлений множества Ω в точке X* и обозначается $\bar{\Gamma}(X)$. Свойства конуса $\bar{\Gamma}(X)$ подробно изложены в Приложении II.

Теорема 2.2. *Соотношение (2.1) эквивалентно неравенству*

$$\inf_{\substack{g \in \bar{\Gamma}(X^*) \\ \|g\|=1}} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) \geq 0. \quad (2.10)$$

Доказательство. Покажем вначале, что из (2.1) следует (2.10). Допустим противное, пусть имеет место (2.1) и существует вектор $g_1 \in \bar{\Gamma}(X^*)$, $\|g_1\| = 1$, такой, что

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_1 \right) = -\rho < 0.$$

По определению конуса $\bar{\Gamma}(X^*)$ найдется вектор $V_1 \in \Gamma(X^*)$, для которого

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, V_1 \right) \leq -\frac{1}{2} \rho. \quad (2.11)$$

Поскольку $V_1 \in \Gamma(X^*)$, то

$$V_1 = \lambda_1 (Z_1 - X^*); \quad \lambda_1 > 0, \quad Z_1 \in \Omega.$$

Отсюда и из (2.11) имеем

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z_1 - X^* \right) \leq -\frac{\rho}{2\lambda_1} < 0,$$

что противоречит (2.1). Итак, установлено, что из (2.1) следует (2.10).

Докажем теперь, что если в точке $X^* \in \Omega$ имеет место (2.10), то в той же точке выполняется и соотношение (2.1). Допустим противное. Тогда найдется точка $Z_1 \in \Omega$, для которой

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z_1 - X^* \right) = -\rho < 0. \quad (2.12)$$

Ясно, что $Z_1 \neq X^*$. Из (2.12) имеем

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, \frac{Z_1 - X^*}{\|Z_1 - X^*\|} \right) = -\frac{\rho}{\|Z_1 - X^*\|} < 0,$$

что противоречит (2.10), ибо вектор

$$g_1 = \frac{1}{\|Z_1 - X^*\|} (Z_1 - X^*)$$

принадлежит $\Gamma(X^*)$ и тем самым $\bar{\Gamma}(X^*)$, и $\|g_1\| = 1$.

Теорема доказана.

З а м е ч а н и е 1. Так как множество

$$\{g \mid g \in \bar{\Gamma}(X^*), \|g\| = 1\}$$

замкнуто и ограничено, то инфимум в левой части (2.10) достигается, и поэтому неравенство (2.10) может быть переписано в виде

$$\min_{\substack{g \in \bar{\Gamma}(X^*) \\ \|g\|=1}} \max_{t \in R(X^*)} \left(\frac{\partial f_t(X^*)}{\partial X}, g \right) \geq 0. \quad (2.10')$$

З а м е ч а н и е 2. Рассмотрим случай $\Omega = E_n$. Для любой точки $X \in E_n$ будем иметь

$$\bar{\Gamma}(X) = \Gamma(X) = E_n,$$

так что условие (2.10') примет вид

$$\min_{\|g\|=1} \max_{t \in R(X^*)} \left(\frac{\partial f_t(X^*)}{\partial X}, g \right) \geq 0,$$

что совпадает с условием (3.1) главы III.

§ 3. Геометрическая интерпретация необходимых условий

1. Зафиксируем $X \in \Omega$, и пусть $\bar{\Gamma}(X)$ — конус возможных направлений множества Ω в точке X . Введем в рассмотрение сопряженный конус $\Gamma^+(X)$:

$$\Gamma^+(X) = \{Z \in E_n \mid (Z, V) \geq 0 \text{ для всех } V \in \bar{\Gamma}(X)\}.$$

Свойства сопряженных конусов подробно изложены в § 2 Приложения II. В частности, там показано, что конус $\Gamma^+(X)$ является замкнутым и выпуклым множеством.

Пусть, как и ранее,

$$L(X) = \left\{ Z = \sum_{t \in R(X)} \alpha_t \frac{\partial f_t(X)}{\partial X} \mid \alpha_t \geq 0, \sum_{t \in R(X)} \alpha_t = 1 \right\}.$$

Теорема 3.1. *Соотношение (2.1) (или, что то же самое, (2.10)) эквивалентно следующему условию:*

$$\Gamma^+(X^*) \cap L(X^*) \neq \emptyset. \quad (3.1)$$

Доказательство. Покажем вначале, что если в точке $X^* \in \Omega$ выполнено условие (2.10), то в той же точке имеет место (3.1). Допустим противное. Тогда по теореме отделимости (теорема 2.1 Приложения II) найдется вектор $W_0 \in \Gamma^{++}(X^*)$ такой, что

$$\max_{Z \in L(X^*)} (W_0, Z) < 0. \quad (3.2)$$

Учитывая лемму 2.3 Приложения II, имеем

$$\Gamma^{++}(X^*) = \bar{\Gamma}(X^*),$$

так что $W_0 \in \bar{\Gamma}(X^*)$. Кроме того, в силу (3.2) $\|W_0\| > 0$. Положим $g_0 = W_0 / \|W_0\|$. Поскольку

$$\max_{Z \in L(X^*)} (g_0, Z) = \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right), \quad (3.3)$$

то из (3.2) получим

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) = \frac{1}{\|W_0\|} \max_{Z \in L(X^*)} (W_0, Z) < 0,$$

что противоречит (2.10), ибо $g_0 \in \bar{\Gamma}(X^*)$ и $\|g_0\| = 1$.

Таким образом, доказано, что из (2.10) следует (3.1).

Докажем теперь, что из (3.1) следует (2.10). Допустим противное. Тогда найдется вектор $g_0 \in \bar{\Gamma}(X^*)$, $\|g_0\| = 1$, для которого

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) < 0.$$

В силу (3.3) последнее неравенство можно переписать в виде

$$\max_{Z \in L(X^*)} (g_0, Z) < 0,$$

причем $g_0 \in \Gamma^{++}(X^*)$, ибо $\bar{\Gamma}(X^*) = \Gamma^{++}(X^*)$.

По теореме 2.1 Приложения II отсюда следует

$$\Gamma^+(X^*) \cap L(X^*) = \emptyset,$$

что противоречит (3.1).

Теорема доказана полностью.

Замечание 1. Если $\Omega = E_n$, то $\Gamma^+(X) = \{0\}$ для любого $X \in E_n$. В этом случае условие (3.1) эквивалентно следующему включению:

$$0 \in L(X^*). \quad (3.4)$$

Соотношение (3.4) совпадает с полученным в § 3 главы III необходимым условием минимакса на всем пространстве E_n .

Замечание 2. Пусть $N = 0$. В этом случае $L(X) = \left\{ \frac{\partial f_0(X)}{\partial X} \right\}$ для любого $X \in \Omega$. Учитывая (3.1), получаем следующее геометрическое условие минимума непрерывно дифференцируемой функции на множестве Ω .

Для того чтобы непрерывно дифференцируемая функция $f_0(X)$ достигала своего минимального на Ω значения в точке $X^ \in \Omega$, необходимо, а в случае выпуклости $f_0(X)$ на Ω и достаточно, чтобы*

$$\frac{\partial f_0(X^*)}{\partial X} \in \Gamma^+(X^*).$$

2. Возьмем теперь любую точку $X_0 \in \Omega$. Построим для этой точки многогранник $L(X_0)$ и конус $\Gamma^+(X_0)$. Положим

$$\rho(X_0) = \inf_{\substack{Z \in L(X_0) \\ Y \in \Gamma^+(X_0)}} \|Z - Y\|. \quad (3.5)$$

Поскольку $L(X_0)$ и $\Gamma^+(X_0)$ — замкнутые множества и одно из них ($L(X_0)$) ограничено, то инфимум в (3.5) достигается. Таким образом, существуют такие точки $Z(X_0) \in L(X_0)$ и $Y(X_0) \in \Gamma^+(X_0)$, что

$$\|Z(X_0) - Y(X_0)\| = \rho(X_0). \quad (3.6)$$

Если $\rho(X_0) = 0$, то точка X_0 является стационарной точкой функции $f(X)$ на множестве Ω , ибо в этом случае, очевидно, выполняется условие (3.1).

Пусть $\rho(X_0) > 0$, т. е. X_0 — не стационарная точка. Нетрудно показать, что в этом случае вектор $Z(X_0) - Y(X_0)$, удовлетворяющий соотношению (3.6), един-

ствен (хотя неединственными могут быть точки $Z(X_0)$ и $Y(X_0)$ — см. рис. 28). Введем обозначение

$$g(X_0) = \frac{Y(X_0) - Z(X_0)}{\|Y(X_0) - Z(X_0)\|}, \quad \|g(X_0)\| = 1.$$

Определение. Вектор $\bar{g} \in E_n$, $\|\bar{g}\| = 1$, называется направлением наискорейшего спуска функции $\varphi(X)$ на множестве Ω в точке X_0 , если

$$\frac{\partial \varphi(X_0)}{\partial \bar{g}} = \min_{\substack{g \in \bar{\Gamma}(X_0) \\ \|g\|=1}} \frac{\partial \varphi(X_0)}{\partial g}.$$

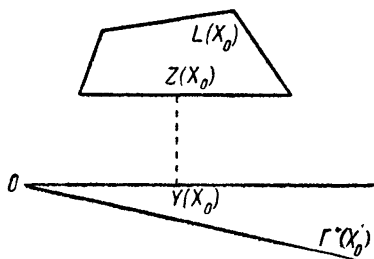


Рис. 28.

Теорема 3.2. Если $\rho(X_0) > 0$, то направление $g(X_0)$ является направлением наискорейшего спуска функции

$\varphi(X)$ на множестве Ω в точке X_0 ; при этом

$$\frac{\partial \varphi(X_0)}{\partial g(X_0)} = -\rho(X_0). \quad (3.7)$$

Для доказательства этой теоремы нам понадобится следующая

Лемма 3.1. Справедливо равенство

$$(g(X_0), Y(X_0)) = 0. \quad (3.8)$$

Доказательство. Будем использовать обозначения:

$$Y_0 = Y(X_0), \quad Z_0 = Z(X_0), \quad g_0 = g(X_0), \quad \rho_0 = \rho(X_0).$$

Поскольку $Y_0 \in \Gamma^+(X_0)$, то при всех $\lambda > 0$

$$\lambda Y_0 \in \Gamma^+(X_0).$$

Далее

$$\begin{aligned} (\lambda Y_0 - Z_0, \lambda Y_0 - Z_0) &= \\ &= ((\lambda - 1)Y_0 + (Y_0 - Z_0), (\lambda - 1)Y_0 + (Y_0 - Z_0)) = \\ &= (\lambda - 1)[(\lambda - 1)(Y_0, Y_0) + 2(Y_0 - Z_0, Y_0)] + \rho_0^2. \end{aligned}$$

Предположение о том, что $(Y_0 - Z_0, Y_0) \neq 0$, немедленно приводит к тому, что при некотором $\lambda_0 > 0$ будет

$$(\lambda_0 Y_0 - Z_0, \lambda_0 Y_0 - Z_0) < \rho_0^2.$$

Последнее неравенство противоречит (3.5). Итак, $(Y_0 - Z_0, Y_0) = 0$, что равносильно (3.8).

Лемма доказана.

Доказательство теоремы. Заметим, что в силу леммы 1.4 из Приложения II для любых $Y \in \Gamma^+(X_0)$ и $Z \in L(X_0)$ выполняется неравенство

$$(Y - Z, Y_0 - Z_0) \geq (Y_0 - Z_0, Y_0 - Z_0). \quad (3.9)$$

Докажем вначале, что $g_0 \in \bar{\Gamma}(X_0)$ или, что то же самое, что $g_0 \in \Gamma^{++}(X_0)$. Для этого достаточно установить, что для всех $Y \in \Gamma^+(X_0)$ будет

$$(Y, g_0) \geq 0. \quad (3.10)$$

В силу (3.8) и определения ρ_0 имеем

$$(Z_0, Y_0 - Z_0) = -\rho_0^2. \quad (3.11)$$

Учитывая (3.9) и (3.11), получаем

$$\begin{aligned} (Y, Y_0 - Z_0) &= \\ &= (Y - Z_0, Y_0 - Z_0) + (Z_0, Y_0 - Z_0) \geq \rho_0^2 - \rho_0^2 = 0, \end{aligned}$$

что равносильно (3.10).

Итак, установлено, что $g_0 \in \bar{\Gamma}(X_0)$.

Найдем $\frac{\partial \Phi(X_0)}{\partial g(X_0)}$. В силу (3.8) и (3.9) имеем

$$\begin{aligned} \frac{\partial \Phi(X_0)}{\partial g(X_0)} &= \max_{i \in R(X_0)} \left(\frac{\partial f_i(X_0)}{\partial X}, g_0 \right) = \max_{Z \in L(X_0)} (Z, g_0) = \\ &= \frac{1}{\|Y_0 - Z_0\|} \max_{Z \in L(X_0)} (Z - Y_0, Y_0 - Z_0) = \\ &= -\frac{1}{\|Y_0 - Z_0\|} (Y_0 - Z_0, Y_0 - Z_0) = -\rho_0. \end{aligned}$$

Тем самым установлено соотношение (3.7).

Осталось доказать, что направление g_0 есть направление наискорейшего спуска функции $\Phi(X)$ на множестве Ω в точке X_0 .

Возьмем любой вектор $\bar{g} \in \bar{\Gamma}(X_0)$, $\|\bar{g}\| = 1$, и покажем, что

$$\frac{\partial \Phi(X_0)}{\partial \bar{g}} \geq -\rho(X_0). \quad (3.12)$$

Поскольку $\bar{g} \in \Gamma^{++}(X_0)$, то для всех $Y \in \Gamma^+(X_0)$ имеем

$$(Y, \bar{g}) \geq 0.$$

Зафиксируем $Y \in \Gamma^+(X_0)$ и $Z \in L(X_0)$. Тогда

$$\begin{aligned} \|Y - Z\| \geq (Y - Z, \bar{g}) &\geq -(Z, \bar{g}) \geq - \max_{Z' \in L(X_0)} (Z', \bar{g}) = \\ &= - \frac{\partial \Phi(X_0)}{\partial \bar{g}}. \end{aligned}$$

Отсюда следует

$$\rho_0 \geq - \frac{\partial \Phi(X_0)}{\partial \bar{g}},$$

что равносильно (3.12).

Теорема доказана.

З а м е ч а н и е 1. Направление наискорейшего спуска функции $\Phi(X)$ на множестве Ω в точке X_0 в случае, если $\rho(X_0) > 0$, единственно. Доказательство этого утверждения вполне аналогично доказательству леммы 3.4 главы III.

З а м е ч а н и е 2. В теореме 3.2 утверждается лишь, что $g(X_0) \in \bar{\Gamma}(X_0)$. Может случиться, что при любых $\alpha > 0$ точки $X_0 + \alpha g(X_0) \notin \Omega$, так что движение в направлении наискорейшего спуска может оказаться невозможным.

3. В качестве примера рассмотрим функцию (см. § 5 гл. III)

$$\Phi(X) = \max_{i \in [1:3]} f_i(X),$$

где

$$\begin{aligned} f_1(X) &= -5x_1 + x_2, & f_2(X) &= 4x_2 + x_1^2 + x_2^2, \\ f_3(X) &= 5x_1 + x_2. \end{aligned}$$

Возьмем точку $X_0 = (0, 0)$. Как было показано раньше, $L(X_0)$ есть треугольник, натянутый на точки $(-5, 1)$, $(0, 4)$, $(5, 1)$. Так как $0 \notin L(X_0)$, то X_0 не является стационарной точкой функции $\Phi(X)$ на E_2 .

Введем ограничение

$$\Omega_1 = \{X \in E_2 \mid 2x_1 + x_2 \geq 0\}.$$

Нетрудно проверить, что в этом случае

$$\Gamma^+(X_0) = \{V = \alpha(2, 1) \mid \alpha \geq 0\}.$$

Так как луч $\Gamma^+(X_0)$ пересекает множество $L(X_0)$ (рис. 29), то X_0 является стационарной точкой функции $\varphi(X)$ на множестве Ω_1 , а поскольку $\varphi(X)$ — выпуклая функция, то X_0 — точка минимума функции $\varphi(X)$ на Ω_1 .

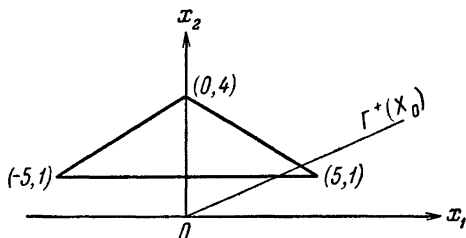


Рис. 29.

Введем теперь другое ограничение:

$$\Omega_2 = \{X \in E_2 \mid 10x_1 + x_2 \geq 0\}.$$

В этом случае $\Gamma^+(X_0) = \{V = \alpha(10, 1) \mid \alpha \geq 0\}$. Так как $L(X_0)$ и $\Gamma^+(X_0)$ не имеют общих точек (рис. 30), то $X_0 = (0, 0)$ не является стационарной точкой функции $\varphi(X)$ на

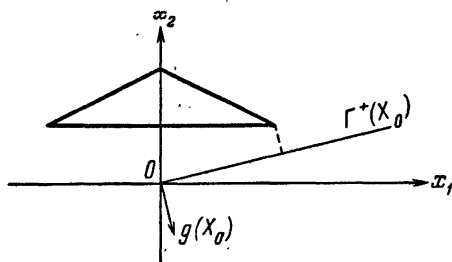


Рис. 30.

множестве Ω_2 . Нетрудно найти направление наискорейшего спуска функции $\varphi(X)$ в точке X_0 на множестве Ω_2 . Это есть вектор

$$g(X_0) = \frac{(1, -10)}{\sqrt{101}}.$$

Заметим, что $X_0 + \alpha g(X_0) \in \Omega_2$ при всех $\alpha \geq 0$.

Положим, наконец,

$$\Omega_4 = \{X \in E_2 \mid (x_1 - 1)^2 + x_2^2 \leq 1\}.$$

Нетрудно проверить, что в этом случае

$$\Gamma^+(X_0) = \{V = \alpha(1, 0) \mid \alpha \geq 0\}, \quad L(X_0) \cap \Gamma^+(X_0) = \emptyset, \\ g(X_0) = (0, -1), \quad \dot{X}_0 + \alpha g(X_0) \notin \Omega_3, \quad \text{если } \alpha > 0.$$

Таким образом, $X_0 = (0, 0)$ не является стационарной точкой функции $\varphi(X)$ на множестве Ω_3 . Вместе с тем движение вдоль направления наискорейшего спуска невозможно, ибо направление $g(X_0)$ выводит нас из области Ω_3 .

§ 4. Достаточные условия локального минимакса при наличии ограничений

1. В § 2 было показано, что если $\varphi(X)$ — выпуклая функция, то любая стационарная точка функции $\varphi(X)$ на Ω есть точка минимума $\varphi(X)$ на Ω . Ниже устанавливаются достаточные условия локального минимакса.

Пусть $X \in \Omega$. Введем в рассмотрение функцию

$$\psi(X) \stackrel{\text{def}}{=} \psi(\Omega, X) = \min_{\substack{g \in \bar{\Gamma}(X) \\ \|g\|=1}} \max_{i \in R(X)} \left(\frac{\partial f_i(X)}{\partial X}, g \right).$$

Тогда необходимое условие (2.10) может быть переписано в виде

$$\psi(X^*) \geq 0.$$

Определение. Точка $X^* \in \Omega$ называется *точкой локального минимума функции $\varphi(X)$ на множестве Ω* , если существует такое $\delta > 0$, что для $X \in \Omega$, $\|X - X^*\| < \delta$, будет

$$\varphi(X) \geq \varphi(X^*).$$

Если при тех же условиях для $X \in \Omega$, $\|X - X^*\| < \delta$, $X \neq X^*$, выполняется неравенство

$$\varphi(X) > \varphi(X^*),$$

то X^* называется *точкой строгого локального минимума функции $\varphi(X)$ на Ω* .

Теорема 4.1. Пусть $X^* \in \Omega$ — стационарная точка функции $\varphi(X)$ на Ω . Если при этом

$$\psi(X^*) > 0, \tag{4.1}$$

то X^* является точкой строгого локального минимума функции $\varphi(X)$ на Ω .

Доказательство. Допустим противное. Тогда найдется последовательность точек $\{X_k\}$, $X_k \in \Omega$, $k = 1, 2, \dots$, $X_k \neq X^*$, для которой

$$\|X_k - X^*\| \xrightarrow{k \rightarrow \infty} 0 \quad \text{и} \quad \varphi(X_k) \leq \varphi(X^*). \quad (4.2)$$

Без ограничения общности можно считать, что

$$g_k \stackrel{\text{df}}{=} \frac{X_k - X^*}{\|X_k - X^*\|} \xrightarrow{k \rightarrow \infty} g_0.$$

Ясно, что $g_0 \in \bar{G}(X^*)$, $\|g_0\| = 1$.

Возьмем любое $i \in R(X^*)$. Так как $\varphi(X^*) = f_i(X^*)$, то, учитывая (4.2), получим

$$f_i(X_k) \leq \varphi(X_k) \leq \varphi(X^*) = f_i(X^*),$$

т. е.

$$f_i(X_k) \leq f_i(X^*). \quad (4.3)$$

Далее

$$\begin{aligned} f_i(X_k) &= f_i(X^* + \|X_k - X^*\| g_k) = \\ &= f_i(X^*) + \|X_k - X^*\| \left(\frac{\partial f_i(X^* + \theta_k g_k)}{\partial X}, g_k \right), \end{aligned}$$

где $\theta_k \in (0, \|X_k - X^*\|)$.

Отсюда и из (4.3) следует, что для всех k

$$\left(\frac{\partial f_i(X^* + \theta_k g_k)}{\partial X}, g_k \right) \leq 0.$$

Учитывая, что $\theta_k \xrightarrow{k \rightarrow \infty} 0$, получим

$$\left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) \leq 0. \quad (4.4)$$

Так как (4.4) справедливо для всех $i \in R(X^*)$, то

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) \leq 0,$$

что противоречит (4.1).

Теорема доказана.

Выясним теперь геометрический смысл условия (4.1).

Теорема 4.2. *Условие (4.1) эквивалентно тому, что множество $L(X^*)$ и конус $\Gamma^+(X^*)$ не могут быть отделены,*

т. е. не существует ненулевого вектора $W \in E_n$ и числа α таких, что

$$(W, Z) \leq \alpha \text{ для всех } Z \in L(X^*), \quad (4.5)$$

$$(W, V) \geq \alpha \text{ для всех } V \in \Gamma^+(X^*). \quad (4.6)$$

Доказательство. Покажем вначале, что если множество $L(X^*)$ и конус $\Gamma^+(X^*)$ не могут быть отделены, то $\psi(X^*) > 0$. Допустим противное. Тогда найдется вектор $g_0 \in \bar{\Gamma}(X^*)$, $\|g_0\| = 1$, для которого

$$\max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, g_0 \right) = \max_{Z \in L(X^*)} (Z, g_0) \leq 0.$$

Поскольку $\bar{\Gamma}(X^*) = \Gamma^{++}(X^*)$, то $g_0 \in \Gamma^{++}(X^*)$. Значит, $(g_0, V) \geq 0$ для всех $V \in \Gamma^+(X^*)$. Но тогда $L(X^*)$ и $\Gamma^+(X^*)$ могут быть отделены, ибо для вектора $W = g_0$ и числа $\alpha = 0$ выполняются соотношения (4.5) и (4.6). Тем самым доказано, что $\psi(X^*) > 0$.

Пусть теперь $\psi(X^*) > 0$. Покажем, что тогда $L(X^*)$ и $\Gamma^+(X^*)$ не могут быть отделены. Допустим противное. Тогда найдется ненулевой вектор $W_0 \in E_n$ и число α_0 такие, что

$$(W_0, Z) \leq \alpha_0 \text{ для всех } Z \in L(X^*), \quad (4.7)$$

$$(W_0, V) \geq \alpha_0 \text{ для всех } V \in \Gamma^+(X^*). \quad (4.8)$$

Не ограничивая общности, можно считать, что $\|W_0\| = 1$. Так как $\Gamma^+(X^*)$ — конус, содержащий начало координат, то из (4.8) имеем: $\alpha_0 \leq 0$.

Отсюда и из (4.7) следует, что

$$\max_{Z \in L(X^*)} (W_0, Z) \leq 0. \quad (4.9)$$

Докажем, что

$$(W_0, V) \geq 0 \text{ для всех } V \in \Gamma^+(X^*).$$

Допустим противное, пусть для некоторого вектора $V' \in \Gamma^+(X^*)$ будет $(W_0, V') = -b < 0$. Так как $\Gamma^+(X^*)$ — конус, то для любого $\beta > 0$ должно быть $\beta V' \in \Gamma^+(X^*)$. На основании (4.8) имеем

$$(W_0, \beta V') = -\beta b \geq \alpha_0.$$

Но α_0 ограничено, а $\beta b \xrightarrow{\beta \rightarrow \infty} \infty$. Полученное противоречие и доказывает справедливость неравенства

$$(W_0, V) \geq 0 \text{ для всех } V \in \Gamma^+(X^*). \quad (4.10)$$

Из (4.10) следует, что $W_0 \in \Gamma^{++}(X^*)$, а поскольку $\Gamma^{++}(X^*) = \bar{\Gamma}(X^*)$, то $W_0 \in \bar{\Gamma}(X^*)$.

Итак, $W_0 \in \bar{\Gamma}(X^*)$ и $\|W_0\| = 1$. Учитывая (4.9), получим

$$\varphi(X^*) = \min_{\substack{g \in \bar{\Gamma}(X^*) \\ \|g\|=1}} \max_{Z \in L(X^*)} (Z, g) \leq \max_{Z \in L(X^*)} (W_0, Z) \leq 0,$$

что противоречит предположению.

Теорема доказана.

Следствие. Если $\Omega = E_n$, то $\Gamma^+(X^*) = \{0\}$, и потому условие неотделимости $L(X^*)$ и $\Gamma^+(X^*)$ означает, что 0 — внутренняя точка множества $L(X^*)$.

Это совпадает с достаточным условием, полученным в главе III.

2. Приведем пример, показывающий, что стационарная точка может и не быть точкой локального минимума, если не выполнено условие (4.1).

Пример. Пусть

$$X = (x_1, x_2) \in E_2,$$

$$\varphi(X) = \max_{i \in [1:3]} f_i(X),$$

где

$$f_1(X) = -5x_1 + x_2,$$

$$f_2(X) = 4x_2 + x_1^2 + x_2^2,$$

$$f_3(X) = 5x_1 + x_2 - x_1^2.$$

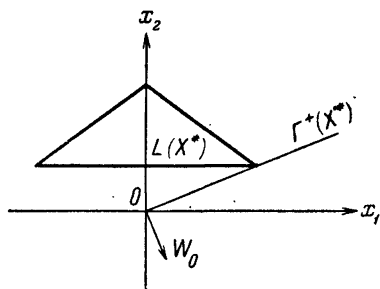


Рис. 31.

Положим $\Omega = \{X \in E_2 \mid 5x_1 + x_2 \geq 0\}$. Возьмем точку $X^* = (0, 0)$. Ясно, что $R(X^*) = \{1, 2, 3\}$; $L(X^*)$ есть треугольник с вершинами в точках $(-5, 1)$, $(0, 4)$, $(5, 1)$; $\Gamma^+(X^*) = \{V = \alpha(5, 1) \mid \alpha \geq 0\}$. Нетрудно проверить (рис. 31), что

$$L(X^*) \cap \Gamma^+(X^*) \neq \emptyset,$$

так что $X^* = (0, 0)$ является стационарной точкой функции $\varphi(X)$ на Ω .

С другой стороны, условие $\varphi(X^*) > 0$ не выполнено, поскольку $L(X^*)$ и $\Gamma^+(X^*)$ могут быть отделены прямой $(X, W_0) = 0$, где $W_0 = (1, -5)$. Действительно,

$$(W_0, Z) \leq 0 \quad \text{для всех} \quad Z \in L(X^*),$$

$$(W_0, V) = 0 \quad \text{для всех} \quad V \in \Gamma^+(X^*).$$

Таким образом, $\psi(X^*) = 0$.

Покажем, что X^* не является точкой локального минимума функции $\varphi(X)$ на Ω . Возьмем точку $X_\alpha = \alpha W_0$, $\alpha > 0$. Очевидно, $X_\alpha \in \Omega$. Имеем

$$f_1(X_\alpha) = -10\alpha, \quad f_2(X_\alpha) = -20\alpha + 26\alpha^2, \quad f_3(X_\alpha) = -\alpha^2.$$

Отсюда следует, что при малых $\alpha > 0$ будет

$$\varphi(X_\alpha) < 0 = \varphi(X^*),$$

т. е. $X^* = (0, 0)$ не является точкой локального минимума функции $\varphi(X)$ на Ω .

Заметим, что этот эффект вызван тем, что функция $f_3(X)$ не является выпуклой, вследствие чего невыпуклой оказалась и функция $\varphi(X)$.

3*. Приведем еще одно достаточное условие локального минимума функции $\varphi(X)$ на множестве Ω .

Теорема 4.3. Пусть $X^* \in \Omega$ — стационарная точка функции $\varphi(X)$ на Ω , причем $\psi(X^*) = 0$. Предположим, что функции $f_i(X)$, $i \in [0:N]$, дважды непрерывно дифференцируемы в некоторой окрестности $S_\delta(X^*)$, $\delta > 0$, точки X^* . Если при некотором $\gamma > 0$ будет

$$\min_{g \in G_\gamma} \max_{i \in R_2(X^*, g)} \left(\frac{\partial^2 f_i(X^*)}{\partial X^2} g, g \right) > 0,$$

где

$$G_\gamma = \left\{ g \mid \|g\| = 1, g \in \bar{\Gamma}(X^*), 0 \leq \frac{\partial \varphi(X^*)}{\partial g} \leq \gamma \right\},$$

$$R_2(X^*, g) = \left\{ i \mid i \in R(X^*), \left(\frac{\partial f_i(X^*)}{\partial X}, g \right) = \frac{\partial \varphi(X^*)}{\partial g} \right\},$$

то X^* является точкой строгого локального минимума функции $\varphi(X)$ на Ω .

Доказательство этой теоремы, аналогичное доказательству теоремы 4.2 гл. III, предоставляется читателю.

§ 5. Некоторые оценки

1. Теорема 5.1. Если для некоторой точки $X_0 \in \Omega$ и некоторого индексного множества $Q \subset [0:N]$ выполняется неравенство

$$\inf_{Z \in \Omega} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) \frac{df}{dZ} - a \leq 0, \quad (5.1)$$

а функции $f_i(X)$ для всех $i \in Q$ выпуклы на Ω , то

$$\min_{i \in Q} f_i(X_0) - a \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0). \quad (5.2)$$

Доказательство. Правое неравенство в (5.2) очевидно. Будем доказывать левое неравенство. Допустим противное. Тогда найдется точка $\bar{X} \in \Omega$, для которой

$$\min_{i \in Q} f_i(X_0) - \varphi(\bar{X}) \stackrel{df}{=} a_1 > a. \quad (5.3)$$

Заметим, что для всех $i \in Q$

$$f_i(X_0) - f_i(\bar{X}) \geq \min_{i \in Q} f_i(X_0) - \varphi(\bar{X}) = a_1. \quad (5.4)$$

Учитывая выпуклость функций $f_i(X)$ при $i \in Q$, лемму 3.2 предложения II и (5.4), получим

$$\left(\frac{\partial f_i(X_0)}{\partial X}, \bar{X} - X_0 \right) \leq f_i(\bar{X}) - f_i(X_0) \leq -a_1.$$

Это неравенство справедливо для всех $i \in Q$, так что

$$\max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, \bar{X} - X_0 \right) \leq -a_1.$$

Из последнего неравенства и из (5.1) следует, что $-a \leq -a_1$, т. е. что $a \geq a_1$, но это противоречит (5.3).

Теорема доказана.

Следствие 1. Если в условиях теоремы $Q \subset R(X_0)$, то

$$\varphi(X_0) - a \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0).$$

Действительно, в этом случае

$$\min_{i \in Q} f_i(X_0) = \varphi(X_0).$$

Следствие 2. Если в условиях теоремы $a = 0$, то

$$\min_{i \in Q} f_i(X_0) \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0).$$

Нетрудно проверить, что условие $a = 0$ равносильно тому, что

$$\Gamma^+(X_0) \cap L(Q, X_0) \neq \emptyset.$$

где $L(Q, X_0)$ — выпуклая оболочка, натянутая на точки $\frac{\partial f_i(X_0)}{\partial X}$, $i \in Q$ (см. доказательство теоремы 3.1).

Следствие 3. Если в условиях теоремы $Q \subset R(X_0)$ и $a = 0$, то

$$\inf_{X \in \Omega} \varphi(X) = \varphi(X_0),$$

т. е. X_0 является точкой минимума функции $\varphi(X)$ на замкнутом выпуклом множестве Ω .

З а м е ч а н и е. Допустим, что Ω — ограниченное множество, причем его диаметр не превосходит числа $d > 0$. И пусть в точке $X_0 \in \Omega$ выполняется неравенство

$$\min_{\substack{g \in \bar{\Gamma}(X_0) \\ \|g\|=1}} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) \stackrel{df}{=} -a \leq 0, \quad (5.5)$$

где $Q \subset [0:N]$. Предположим, наконец, что функции $f_i(X)$ для всех $i \in Q$ выпуклы на Ω . Тогда

$$\min_{i \in Q} f_i(X_0) - ad \leq \min_{X \in \Omega} \varphi(X) \leq \varphi(X_0). \quad (5.6)$$

Докажем вначале, что

$$\inf_{Z \in \Omega} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) \geq -ad. \quad (5.7)$$

Для этого достаточно доказать, что для всех $Z \in \Omega$

$$\max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) \geq -ad.$$

При $Z = X_0$ последнее неравенство очевидно. Пусть $Z \neq X_0$. Тогда в силу (5.5)

$$\begin{aligned} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) &= \\ &= \|Z - X_0\| \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, \frac{Z - X_0}{\|Z - X_0\|} \right) \geq \\ &\geq -a \|Z - X_0\| \geq -ad. \end{aligned}$$

Таким образом, неравенство (5.7) доказано. Теперь (5.6) следует из (5.7) и теоремы 5.1.

2. Теорема 5.2. Пусть для некоторой точки $X_0 \in \Omega$ нашлось множество $Q \subset [0:N]$ такое, что

$$\min_{\substack{g \in \bar{\Gamma}(X_0) \\ \|g\|=1}} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, g \right) \stackrel{df}{=} -a \leq 0. \quad (5.8)$$

Пусть, далее, функции $f_i(X)$ для всех $i \in Q$ являются дважды непрерывно дифференцируемыми на Ω' , причем для всех $i \in Q$, $X \in \Omega$ и $V \in E_n$

$$\left(\frac{\partial^2 f_i(X)}{\partial X^2} V, V \right) \geq m \|V\|^2, \quad (5.9)$$

где $m > 0$ не зависит от i , X и V . Тогда

$$\min_{i \in Q} f_i(X_0) - \frac{a^2}{2m} \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0). \quad (5.10)$$

Доказательство. Поскольку правое неравенство в (5.10) очевидно, будем доказывать только левое неравенство.

Для любого $Z \in \Omega$ и всех $i \in [0:N]$ имеем

$$\begin{aligned} f_i(Z) &= f_i(X_0 + (Z - X_0)) = f_i(X_0) + \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) + \\ &+ \frac{1}{2} \left(\frac{\partial^2 f_i(X_0 + \theta_i(Z - X_0))}{\partial X^2} (Z - X_0), Z - X_0 \right), \end{aligned} \quad (5.11)$$

где $\theta_i \in (0, 1)$.

По (5.8) для любого $Z \in \Omega$ найдется $i(Z) \in Q$ такое, что

$$\left(\frac{\partial f_{i(Z)}(X_0)}{\partial X}, Z - X_0 \right) \geq -\|Z - X_0\|a. \quad (5.12)$$

Учитывая (5.11), (5.12) и (5.9), получим

$$\begin{aligned} f_{i(Z)}(Z) &\geq f_{i(Z)}(X_0) - a\|Z - X_0\| + \frac{1}{2} m \|Z - X_0\|^2 \geq \\ &\geq \min_{i \in Q} f_i(X_0) - a\|Z - X_0\| + \frac{1}{2} m \|Z - X_0\|^2. \end{aligned}$$

Так как

$$\begin{aligned} \frac{1}{2} m \|Z - X_0\|^2 - a\|Z - X_0\| &\geq \\ &\geq \min_{-\infty < \beta < \infty} \left[\frac{1}{2} m \beta^2 - a\beta \right] = -\frac{a^2}{2m}, \end{aligned}$$

то

$$\varphi(Z) \geq f_{i(Z)}(Z) \geq \min_{i \in Q} f_i(X_0) - \frac{a^2}{2m}.$$

Отсюда следует $\inf_{Z \in \Omega} \varphi(Z) \geq \min_{i \in Q} f_i(X_0) - \frac{a^2}{2m}.$

Теорема доказана.

Следствие. Если в условиях теоремы $Q \subset R(X_0)$, то

$$0 \leq \varphi(X_0) - \inf_{Z \in \Omega} \varphi(Z) \leq \frac{a^2}{2m}.$$

§ 6. Метод последовательных приближений для нахождения стационарных точек

1. В этом параграфе будет систематически использоваться первая норма векторов: если $V = (v_1, \dots, v_n)$, то

$$\|V\|_I = \max_{i \in [1:n]} |v_i|.$$

Очевидно, $\|V\| \leq \sqrt{n} \|V\|_I$.

Как обычно, считаем, что Ω — замкнутое выпуклое (не обязательно ограниченное) множество в E_n . Положим при фиксированном $X \in \Omega$

$$\Omega(X) = \{Z \in \Omega \mid \|Z - X\|_I \leq 1\}.$$

Введем функции

$$\begin{aligned} \psi_1(X) &= \min_{Z \in \Omega(X)} \max_{i \in R(X)} \left(\frac{\partial f_i(X)}{\partial X}, Z - X \right), \\ \psi_{1\varepsilon}(X) &= \min_{Z \in \Omega(X)} \max_{i \in R_\varepsilon(X)} \left(\frac{\partial f_i(X)}{\partial X}, Z - X \right), \end{aligned}$$

где

$$\begin{aligned} R(X) &= \{i \in [0:N] \mid f_i(X) = \varphi(X)\}, \\ R_\varepsilon(X) &= \{i \in [0:N] \mid \varphi(X) - f_i(X) \leq \varepsilon\}. \end{aligned}$$

Нетрудно проверить, что точка $X^* \in \Omega$ тогда и только тогда будет стационарной точкой функции $\varphi(X)$ на Ω , когда $\psi_1(X^*) = 0$. Заметим также, что при фиксированном $X \in \Omega$ функция $\psi_{1\varepsilon}(X)$ как функция от ε , $0 \leq \varepsilon < \infty$, является неубывающей и кусочно-постоянной, причем существует такое $a_1(X) > 0$, что

$$\psi_{1\varepsilon}(X) = \psi_1(X) \quad \text{при} \quad \varepsilon \in [0, a_1(X)). \quad (6.1)$$

Опишем один метод последовательных приближений для нахождения стационарных точек функции $\varphi(X)$ на множестве Ω .

Зафиксируем два параметра $\varepsilon_0 > 0$ и $\rho_0 > 0$ и выберем начальное приближение $X_0 \in \Omega$. Будем считать, что множество $M(X_0) = \{X \in \Omega \mid \varphi(X) \leq \varphi(X_0)\}$ ограничено (очевидно, оно будет и замкнутым).

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $\psi_1(X_k) = 0$, то X_k — стационарная точка функции $\varphi(X)$ на Ω и процесс на этом заканчивается.

Если же $\psi_1(X_k) < 0$, то перебираем числовую последовательность

$$\varepsilon_v = \frac{\varepsilon_0}{2^v}, \quad v = 0, 1, 2, \dots,$$

до тех пор, пока впервые не выполнится неравенство

$$\psi_{1\varepsilon_v}(X_k) \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon_v \quad (6.2)$$

(конечно, это может случиться и при $v=0$). Первое v , при котором выполнится (6.2), обозначим через v_k . Число v_k конечное, ибо правая часть неравенства (6.2) стремится к нулю при $v \rightarrow \infty$, а левая часть в силу (6.1) — к $\psi_1(X_k) < 0$.

Таким образом,

$$\psi_{1\varepsilon_{v_k}}(X_k) \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon_{v_k},$$

и если $v_k > 0$, то

$$\psi_{1\varepsilon_{v_k-1}}(X_k) > -\frac{\rho_0}{\varepsilon_0} \varepsilon_{v_k-1}.$$

В дальнейшем будем использовать обозначение

$$\hat{\varepsilon}_k = \varepsilon_{v_k}.$$

Пусть точка $Z_k \in \Omega(X_k)$ такова, что

$$\psi_{1\varepsilon_k}(X_k) = \max_{i \in R_{\varepsilon_k}(X_k)} \left(\frac{\partial f_i(X_k)}{\partial X}, Z_k - X_k \right).$$

Рассмотрим отрезок

$$X = X_k(\alpha) \stackrel{\text{def}}{=} X_k + \alpha(Z_k - X_k), \quad 0 \leq \alpha \leq 1.$$

Ясно, что $X_k(\alpha) \in \Omega$ при $\alpha \in [0, 1]$. Найдем $\alpha_k \in [0, 1]$, при котором

$$\varphi(X_k(\alpha_k)) = \min_{\alpha \in [0, 1]} \varphi(X_k(\alpha)).$$

Теперь полагаем

$$X_{k+1} = X_k(\alpha_k).$$

Ясно, что $X_{k+1} \in \Omega$. Нетрудно также проверить, что

$$\varphi(X_{k+1}) < \varphi(X_k).$$

(см. доказательство неравенства (6.6) в главе III). Отсюда, в частности, следует, что $X_{k+1} \in M(X_0)$.

Продолжая описанный процесс, получим последовательность $\{X_k\}$ и связанные с нею последовательности $\{\hat{\varepsilon}_k\}$, $\{Z_k\}$, $\{\alpha_k\}$, причем $X_k \in M(X_0)$, $Z_k \in \Omega(X_k)$, $k = 0, 1, 2, \dots$, и

$$\varphi(X_0) > \varphi(X_1) > \dots > \varphi(X_k) > \dots \quad (6.3)$$

Если последовательность $\{X_k\}$ конечна, то последний ее член по построению является стационарной точкой функции $\varphi(X)$ на множестве Ω .

Рассмотрим отдельно случай, когда последовательность $\{X_k\}$ бесконечна. Поскольку множество $M(X_0)$ ограниченное и замкнутое, а функция $\varphi(X)$ непрерывна на $M(X_0)$, то $\varphi(X)$ ограничена снизу на $M(X_0)$. Отсюда и из (6.3) следует, что существует предел

$$\lim_{k \rightarrow \infty} \varphi(X_k) = \varphi^*,$$

причем при всех $k = 0, 1, 2, \dots$

$$\varphi(X_k) > \varphi^*. \quad (6.4)$$

2. Лемма 6.1. *При $k \rightarrow \infty$ имеет место предельное соотношение*

$$\hat{\varepsilon}_k \rightarrow 0.$$

Доказательство. Предположим противное. Тогда найдется число $\varepsilon^* > 0$ и подпоследовательность $\{\hat{\varepsilon}_{k_j}\}$ такие, что

$$\hat{\varepsilon}_{k_j} \geq \varepsilon^*.$$

Так как $R_{\varepsilon^*}(X_{k_j}) \subset R_{\varepsilon_{k_j}}(X_{k_j})$, то имеем для $i \in R_{\varepsilon^*}(X_{k_j})$

$$\begin{aligned} f_i(X_{k_j}(\alpha)) &= f_i(X_{k_j} + \alpha(Z_{k_j} - X_{k_j})) \leq \\ &\leq f_i(X_{k_j}) + \alpha \psi_{i k_j}(X_{k_j}) + o_{i k_j}(\alpha) \leq \\ &\leq \varphi(X_{k_j}) - \alpha \frac{\rho_0}{\varepsilon_0} \varepsilon_{k_j} + o_{i k_j}(\alpha) \leq \varphi(X_{k_j}) - \alpha \rho^* + o_{i k_j}(\alpha), \end{aligned}$$

где $\rho^* = \frac{\rho_0}{\varepsilon_0} \varepsilon^*$.

Отсюда, из условия $X_{k_j} \in M(X_0)$ и из неравенства $\|Z_{k_j} - X_{k_j}\|_I \leq 1$ следует, что найдется число α_0 , $0 < \alpha_0 < 1$, такое, что для $i \in R_{\varepsilon^*}(X_{k_j})$ и $\alpha \in [0, \alpha_0]$ равномерно по k_j будет

$$f_i(X_{k_j}(\alpha)) \leq \varphi(X_{k_j}) - \frac{1}{2} \alpha \rho^*. \quad (6.5)$$

Положим

$$C_1 = \max_{X \in M(X_0)} \max_{i \in [0: N]} \left\| \frac{\partial f_i(X)}{\partial X} \right\|.$$

Так как $X_{k_j} \in M(X_0)$, то для $i \notin R_{\varepsilon^*}(X_{k_j})$ получим

$$\begin{aligned} f_i(X_{k_j}(\alpha)) &= f_i(X_{k_j} + \alpha(Z_{k_j} - X_{k_j})) \leq f_i(X_{k_j}) + \\ &+ \sqrt{n} \alpha C_1 + o_{i k_j}(\alpha) \leq \varphi(X_{k_j}) - \varepsilon^* + \sqrt{n} \alpha C_1 + o_{i k_j}(\alpha). \end{aligned}$$

Отсюда следует, что найдется число α_1 , $0 < \alpha_1 \leq \alpha_0$, такое, что для $i \notin R_{\varepsilon^*}(X_{k_j})$ и $\alpha \in [0, \alpha_1]$ равномерно по k_j будет

$$f_i(X_{k_j}(\alpha)) \leq \varphi(X_{k_j}) - \frac{1}{2} \varepsilon^*. \quad (6.6)$$

Зафиксируем любое $\alpha' \in (0, \alpha_1)$. Учитывая (6.5) и (6.6), получаем для всех k_j

$$\varphi(X_{k_j}(\alpha')) \leq \varphi(X_{k_j}) - \beta, \quad (6.7)$$

где $\beta = \min \left\{ \frac{1}{2} \alpha' \rho^*, \frac{1}{2} \varepsilon^* \right\}$.

Так как $\varphi(X_{k_j}) \xrightarrow{k \rightarrow \infty} \varphi^*$, то при достаточно больших k_j окажется $\varphi(X_{k_j}) \leq \varphi^* + \frac{1}{2} \beta$. Из (6.7) тогда имеем

$$\varphi(X_{k_j}(\alpha')) \leq \varphi^* - \frac{1}{2} \beta.$$

Тем более

$$\varphi(X_{k_j+1}) = \min_{\alpha \in [0, 1]} \varphi(X_{k_j}(\alpha)) \leq \varphi(X_{k_j}(\alpha')) \leq \varphi^* - \frac{1}{2} \beta,$$

что противоречит (6.4).

Лемма доказана.

3. Если построенная выше последовательность $\{X_k\}$ бесконечна, то в силу того, что $X_k \in M(X_0)$, $k = 0, 1, 2, \dots$, а $M(X_0)$ — ограниченное замкнутое множество, существует по крайней мере одна предельная точка последовательности $\{X_k\}$.

Теорема 6.1. Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω .

Доказательство. Пусть $X_{k_j} \xrightarrow{k_j \rightarrow \infty} X^*$. Ясно, что $X^* \in \Omega$. Надо доказать, что $\psi_1(X^*) = 0$. Допустим противное:

$$\psi_1(X^*) \stackrel{df}{=} \min_{Z \in \Omega(X^*)} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, Z - X^* \right) = -b < 0. \quad (6.8)$$

Обозначим через $\bar{Z} \in \Omega(X^*)$ точку, для которой

$$\psi_1(X^*) = \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, \bar{Z} - X^* \right). \quad (6.9)$$

В силу определения $\Omega(X^*)$ точка $\bar{Z}$ принадлежит Ω и

$$\|\bar{Z} - X^*\|_l \leq 1. \quad (6.10)$$

По лемме 6.1 главы III найдется число $\gamma^* > 0$ такое, что для достаточно больших $k_j > K$ и $\varepsilon \in [0, \gamma^*]$ будет

$$R_\varepsilon(X_{k_j}) \subset R(X^*). \quad (6.11)$$

Увеличивая, если нужно, номер K , добьемся для $k_j > K$ выполнения еще следующих неравенств:

$$\|X_{k_j} - X^*\|_l \leq 1, \quad (6.12)$$

$$\begin{aligned} & \left| \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, \bar{Z} - X^* \right) - \right. \\ & \quad \left. - \max_{i \in R(X^*)} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, \bar{Z} - X_{k_j} \right) \right| < \frac{b}{2}. \quad (6.13) \end{aligned}$$

Докажем, что для $k_j > K$ и $\varepsilon \in [0, \gamma^*]$

$$\psi_{1\varepsilon}(X_{k_j}) \leq -\frac{b}{4}. \quad (6.14)$$

Имеем в силу (6.11) и (6.13)

$$\begin{aligned} \max_{i \in R(X^*)} \left(\frac{\partial f_i(X^*)}{\partial X}, \bar{Z} - X^* \right) &\geq \\ &\geq \max_{i \in R(X^*)} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, \bar{Z} - X_{k_j} \right) - \frac{b}{2} \geq \\ &\geq \max_{i \in R_\varepsilon(X_{k_j})} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, \bar{Z} - X_{k_j} \right) - \frac{b}{2}. \end{aligned}$$

Отсюда и из (6.9) следует, что

$$\max_{i \in R_\varepsilon(X_{k_j})} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, \bar{Z} - X_{k_j} \right) \leq -\frac{b}{2}. \quad (6.15)$$

Введем точку $\tilde{Z}_{k_j} = \frac{1}{2}(\bar{Z} + X_{k_j})$. Очевидно, $\tilde{Z}_{k_j} \in \Omega$ и

$$\tilde{Z}_{k_j} - X_{k_j} = \frac{1}{2}(\bar{Z} - X_{k_j}). \quad (6.16)$$

Учитывая (6.10) и (6.12), получаем

$$\begin{aligned} \|\tilde{Z}_{k_j} - X_{k_j}\|_I &= \frac{1}{2} \|\bar{Z} - X_{k_j}\|_I \leq \\ &\leq \frac{1}{2} (\|\bar{Z} - X^*\|_I + \|X^* - X_{k_j}\|_I) \leq 1. \end{aligned}$$

Таким образом, $\tilde{Z}_{k_j} \in \Omega(X_{k_j})$. Далее, в силу (6.16) и определения $\psi_{1\varepsilon}(X_{k_j})$ будем иметь

$$\begin{aligned} \max_{i \in R_\varepsilon(X_{k_j})} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, \bar{Z} - X_{k_j} \right) &= \\ &= 2 \max_{i \in R_\varepsilon(X_{k_j})} \left(\frac{\partial f_i(X_{k_j})}{\partial X}, \tilde{Z}_{k_j} - X_{k_j} \right) \geq 2\psi_{1\varepsilon}(X_{k_j}). \end{aligned} \quad (6.17)$$

Объединяя (6.15) и (6.17), получим (6.14). Итак, неравенство (6.14) доказано.

Положим

$$\varepsilon' = \min \left\{ \gamma^*, \frac{\varepsilon_0}{\rho_0} \cdot \frac{b}{4} \right\}.$$

Через ε^* обозначим наибольший из членов последовательности $\varepsilon_\nu = \varepsilon_0/2^\nu$, $\nu = 0, 1, 2, \dots$, меньших ε' . Тогда для всех $k_j > K$ будет

$$\psi_{1\varepsilon^*}(X_{k_j}) \leq -\frac{b}{4} \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon' \leq -\frac{\rho_0}{\varepsilon_0} \varepsilon^*.$$

Вспоминая, что $\hat{\varepsilon}_{k_j}$ — первый член последовательности $\{\varepsilon_\nu\}$, для которого выполняется аналогичное неравенство, получим для тех же $k_j > K$

$$\hat{\varepsilon}_{k_j} \geq \varepsilon^*,$$

что противоречит лемме 6.1.

Теорема доказана.

4. На каждом шаге описанного выше метода последовательных приближений приходится находить $\hat{\varepsilon}_k$, Z_k и α_k . Нахождение α_k сводится к минимизации функции одной переменной $\Phi_k(\alpha) = \varphi(X_k(\alpha))$ на отрезке $[0, 1]$. Решение этой задачи не вызывает затруднений (см. § 9 гл. III).

Успех в нахождении $\hat{\varepsilon}_k$ и Z_k зависит в конечном счете от умения решать при фиксированном X следующую минимаксную задачу:

$$\max_{i \in R_\varepsilon(X)} \left(\frac{\partial f_i(X)}{\partial X}, Z - X \right) \rightarrow \min_{Z \in \Omega(X)}. \quad (6.18)$$

Если обозначить через $L_\varepsilon(X)$ выпуклую оболочку, натянутую на точки $\partial f_i(X)/\partial X$, $i \in R_\varepsilon(X)$, то (6.18) можно переписать в виде

$$\max_{V \in L_\varepsilon(X)} (V, Z - X) \rightarrow \min_{Z \in \Omega(X)}.$$

Последняя задача не является, вообще говоря, тривиальной. Некоторые соображения о решении подобных задач будут приведены в § 5 главы VI.

Заметим лишь, что в случае, когда множество Ω задается системой линейных неравенств, задача (6.18) сводится к задаче линейного программирования. Действительно, пусть $\Omega = \{Z \in E_n \mid (A_j, Z) + b_j \leq 0 \text{ для всех } j \in [0 : N_1]\}$. Обозначим через $W^* = (z_1^*, \dots, z_n^*, u^*)$ век-

тор, минимизирующий линейную форму $\mathcal{L}(W) = u$ при ограничениях

$$\begin{aligned} \left(\frac{\partial f_i(X)}{\partial X}, Z - X \right) &\leq u, & i \in R_e(X); \\ (A_j, Z) + b_j &\leq 0, & j \in [0 : N_1]; \\ -1 &\leq z_k - x_k \leq 1, & k \in [1 : n]. \end{aligned}$$

Тогда, очевидно, $Z^* = (z_1^*, \dots, z_n^*)$ будет решением минимаксной задачи (6.18).

5. Изложенный в этом параграфе метод может быть использован и для решения минимаксных задач при отсутствии ограничений на параметры: $\Omega = E_n$. В этом случае получится метод, отличный от методов главы III и обладающий той особенностью, что для нахождения направления спуска $g_k = Z_k - X_k$ приходится решать задачу линейного программирования.

ОБОБЩЕННАЯ ЗАДАЧА НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

§ 1. Постановка задачи

Обобщенной задачей нелинейного (математического) программирования будем называть задачу минимизации функции максимума $\varphi(X)$ на множестве Ω , заданном системой неравенств

$$\Omega = \{X \in E_n \mid h_j(X) \leq 0 \quad \text{для всех} \quad j \in [0 : N_1]\}.$$

Ниже будет показано, как применить полученные в главе IV результаты к решению подобных задач. В частности, приводятся новые формы необходимого условия минимакса, учитывающие специфику задания множества Ω , и развиваются новые методы последовательных приближений для нахождения стационарных точек функции $\varphi(X)$ на Ω .

Отдельный параграф посвящен методу штрафных функций. Исследуется случай наличия ограничений типа линейных равенств.

§ 2. Свойства множеств, определяемых неравенствами

1. Пусть на E_n заданы функции $h_j(X)$, $j \in [0 : N_1]$.

Предположим, что функции $h_j(X)$ выпуклы и непрерывно дифференцируемы на E_n . Образует множество

$$\Omega = \{X \in E_n \mid h_j(X) \leq 0 \quad \text{для всех} \quad j \in [0 : N_1]\}. \quad (2.1)$$

Ясно, что множество Ω является выпуклым и замкнутым.

Введем обозначение

$$\varphi_1(X) = \max_{j \in [0 : N_1]} h_j(X).$$

Будем говорить, что выполнено *условие Слейтера*, если

$$\inf_{X \in E_n} \varphi_1(X) < 0. \quad (2.2)$$

Из (2.2) следует, что Ω имеет внутренние точки. Пусть далее

$$\partial\Omega = \{X \in E_n \mid \varphi_1(X) = 0\}, \quad Q(X) = \{j \in [0 : N_1] \mid h_j(X) = 0\}.$$

Если $X \in \partial\Omega$, то, очевидно, $Q(X) \neq \emptyset$. Обратно, если $X \in \Omega$ и $Q(X) \neq \emptyset$, то $X \in \partial\Omega$.

Лемма 2.1. Если $X_0 \in \partial\Omega$, то при выполнении условия Слейтера (2.2) имеет место неравенство

$$\min_{\|g\|=1} \max_{j \in Q(X_0)} \left(\frac{\partial h_j(X_0)}{\partial X}, g \right) < 0. \quad (2.3)$$

Доказательство. Допустим противное, т. е. что

$$\min_{\|g\|=1} \max_{j \in Q(X_0)} \left(\frac{\partial h_j(X_0)}{\partial X}, g \right) \geq 0.$$

Поскольку $X_0 \in \partial\Omega$, то

$$Q(X_0) = \{j \in [0 : N_1] \mid h_j(X_0) = \varphi_1(X_0)\}.$$

В силу теоремы 3.1 главы III и сделанных предположений о функциях $h_j(X)$ заключаем, что

$$\min_{X \in E_n} \varphi_1(X) = \varphi_1(X_0).$$

Но $\varphi_1(X_0) = 0$, значит,

$$\min_{X \in E_n} \varphi_1(X) = 0,$$

что противоречит (2.2). Лемма доказана.

З а м е ч а н и е. Неравенство (2.3) равносильно соотношению

$$0 \notin L_1(X_0),$$

где $L_1(X_0) = \text{co } H_1(X_0)$,

$$H_1(X_0) = \left\{ \frac{\partial h_j(X_0)}{\partial X} \mid j \in Q(X_0) \right\}.$$

Это утверждение следует из теоремы 3.2 главы III.

Лемма 2.2. Пусть выполнено условие Слейтера (2.2). Тогда, если для некоторой точки $X_0 \in \Omega$ и некоторого $j_0 \in [0 : N_1]$ оказалось, что $h_{j_0}(X_0) = 0$, то

$$\frac{\partial h_{j_0}(X_0)}{\partial X} \neq 0.$$

Доказательство. Прежде всего заметим, что в условиях леммы $X_0 \in \partial\Omega$ и $j_0 \in Q(X_0)$.

Допустив, что $\partial h_{j_0}(X_0)/\partial X = 0$, придем к противоречию с неравенством (2.3). Лемма доказана.

Лемма 2.3. Пусть по-прежнему выполнено условие Слейтера (2.2) и допустим, что множество Ω ограничено. Тогда существует такое $a_1 > 0$, что для всех $X \in \partial\Omega$ будет

$$\min_{j \in Q(X)} \left\| \frac{\partial h_j(X)}{\partial X} \right\| \geq a_1. \quad (2.4)$$

Доказательство. Допустим противное. Тогда найдутся последовательность $\{X_k\}$, $X_k \in \partial\Omega$, $k = 1, 2, \dots$, и соответствующая ей последовательность индексов $\{j(k)\}$, $j(k) \in [0 : N_1]$, $j(k) \in Q(X_k)$, такие, что

$$\left\| \frac{\partial h_{j(k)}(X_k)}{\partial X} \right\| \xrightarrow{k \rightarrow \infty} 0. \quad (2.5)$$

Без ограничения общности можно считать, что $j(k) = j_0$ и $X_k \xrightarrow{k \rightarrow \infty} X^*$, причем $X^* \in \partial\Omega$.

Переходя к пределу при $k \rightarrow \infty$ в равенстве $h_{j_0}(X_k) = 0$ и учитывая (2.5), получим

$$h_{j_0}(X^*) = 0, \quad \frac{\partial h_{j_0}(X^*)}{\partial X} = 0.$$

Но это противоречит лемме 2.2.

Лемма доказана.

2*. Результаты этого пункта носят вспомогательный характер и будут впервые использованы лишь в § 6.

Введем в рассмотрение множества

$$\partial_\mu \Omega = \{X \mid -\mu \leq \varphi_1(X) \leq 0\},$$

$$Q_\mu(X) = \{j \in [0 : N_1] \mid -\mu \leq h_j(X) \leq 0\}.$$

Если $X \in \partial_\mu \Omega$, то, очевидно, $Q_\mu(X) \neq \emptyset$. Обратно, если $X \in \Omega$ и $Q_\mu(X) \neq \emptyset$, то $X \in \partial_\mu \Omega$.

Лемма 2.4. Предположим, что множество Ω ограничено и выполнено условие Слейтера (2.2). Тогда для любого $a' \in (0, a_1)$ (где a_1 — число, фигурирующее

в лемме 2.3) найдется такое $\bar{\mu} > 0$, что для всех $\mu \in [0, \bar{\mu}]$ будет

$$\min_{X \in \partial_{\mu} \Omega} \min_{j \in Q_{\mu}(X)} \left\| \frac{\partial h_j(X)}{\partial X} \right\| \geq a'.$$

Доказательство. Допустим, что для некоторого $a' \in (0, a_1)$ требуемого $\bar{\mu}$ не существует. Тогда найдется последовательность $\{\mu_k\}$, $\mu_k \geq 0$, $\mu_k \xrightarrow{k \rightarrow \infty} 0$, и соответствующие ей последовательности $\{X_k\}$ и $\{j(k)\}$, $X_k \in \partial_{\mu_k} \Omega$, $j(k) \in Q_{\mu_k}(X_k)$, такие, что

$$\| \partial h_{j(k)}(X_k) / \partial X \| < a'. \quad (2.6)$$

Без ограничения общности можно считать, что $j(k) \equiv j_0$ и $X_k \xrightarrow{k \rightarrow \infty} X^*$, при этом $X^* \in \partial \Omega$ и $j_0 \in Q(X^*)$. Переходя к пределу при $k \rightarrow \infty$ в неравенстве (2.6), получим

$$\left\| \frac{\partial h_{j_0}(X^*)}{\partial X} \right\| \leq a' < a_1,$$

что противоречит (2.4).

Лемма доказана.

Теперь рассмотрим множества

$$\partial'_{\mu} \Omega = \{X \mid -\mu \leq \max_{j \in [0 : N_1]} t_j(X) h_j(X) \leq 0\},$$

$$Q'_{\mu}(X) = \{j \in [0 : N_1] \mid -\mu \leq t_j(X) h_j(X) \leq 0\},$$

где $t_j(X) = \| \partial h_j(X) / \partial X \|^{-1}$ (по определению $\frac{-a}{0} = -\infty$, если $a > 0$).

Лемма 2.5. Пусть выполнено условие Слейтера (2.2), и пусть множество Ω ограничено. Для любого $a' \in (0, a_1)$ найдется $\mu' > 0$ такое, что для всех $\mu \in [0, \mu']$ будет

$$\min_{X \in \partial'_{\mu} \Omega} \min_{j \in Q'_{\mu}(X)} \left\| \frac{\partial h_j(X)}{\partial X} \right\| \geq a'. \quad (2.7)$$

Доказательство. Справедливость этой леммы непосредственно следует из леммы 2.4. При этом в качестве μ' можно взять $\mu' = \bar{\mu}/a'$. Действительно, пусть $\mu \in [0, \mu']$, $X \in \partial'_{\mu} \Omega$, $j \in Q'_{\mu}(X)$. В этом случае

$$-\mu \leq t_j(X) h_j(X) \leq 0.$$

Учитывая, что $\mu \leq \bar{\mu}/a'$, запишем

$$\frac{-\bar{\mu}}{a'} \leq t_j(X) h_j(X) \leq 0. \quad (2.8)$$

Если $-\bar{\mu} \leq h_j(X) \leq 0$, то $X \in \partial_{\bar{\mu}}\Omega$, $j \in Q_{\bar{\mu}}(X)$, и потому по лемме 2.4

$$\left\| \frac{\partial h_j(X)}{\partial X} \right\| \geq a'. \quad (2.9)$$

Если же $h_j(X) < -\bar{\mu}$, то в силу (2.8) должно быть

$$t_j(X) < \frac{1}{a'},$$

или, что то же самое,

$$\left\| \frac{\partial h_j(X)}{\partial X} \right\| > a'. \quad (2.10)$$

Из (2.9) и (2.10) следует (2.7).

Лемма доказана.

З а м е ч а н и е. В леммах 2.3—2.5 требовалось, чтобы множество Ω было ограниченным. Предположим теперь, что Ω не является ограниченным множеством, а $M \subset \Omega$ — ограниченное замкнутое подмножество. Тогда леммы 2.3—2.5 остаются справедливыми, если в них множества $\partial\Omega$, $\partial_{\mu}\Omega$, $\partial'_{\mu}\Omega$ заменить соответственно на $(\partial\Omega) \cap M$, $(\partial_{\mu}\Omega) \cap M$, $(\partial'_{\mu}\Omega) \cap M$.

Л е м м а 2.6. Пусть M — ограниченное замкнутое множество, содержащееся в Ω , и μ — некоторое положительное число. Тогда найдется такое $\delta > 0$, что равномерно по $X \in M$ и $j \notin Q'_{\mu}(X)$ будет

$$h_j(X) \leq -\delta.$$

Д о к а з а т е л ь с т в о. Множество M_j тех X из M , для которых $t_j(X) h_j(X) \leq -\mu$, ограничено и замкнуто, и потому $h_j(X)$ достигает на нем своего максимального значения $-\delta_j < 0$. Остается положить $\delta = \min \delta_j$ и заметить, что из условий $X \in M$, $j \notin Q'_{\mu}(X)$ следует $X \in M_j$.

Лемма доказана.

3. Пусть по-прежнему

$$\Omega = \{X \mid h_j(X) \leq 0, \quad j \in [0 : N_1]\},$$

и выполнено условие Слейтера (2.2), т. е. существует точка $\bar{X} \in \Omega$, для которой

$$\max_{j \in [0 : N_1]} h_j(\bar{X}) < 0. \quad (2.11)$$

Зафиксируем $X_0 \in \Omega$ и рассмотрим множество

$$\Gamma(X_0) = \{V \in E_n \mid h_j(X_0 + \alpha V) \leq 0 \text{ для всех } j \in [0 : N_1] \text{ и всех } \alpha \in [0, \alpha_0(V)], \text{ где } \alpha_0(V) > 0\}.$$

Через $\bar{\Gamma}(X_0)$ обозначим замыкание множества $\Gamma(X_0)$. Очевидно, $\bar{\Gamma}(X_0)$ является конусом возможных направлений множества Ω в точке X_0 (см. Приложение II, § 2).

Введем в рассмотрение множество

$$B(X_0) = \left\{ V \in E_n \mid \left(V, \frac{\partial h_j(X_0)}{\partial X} \right) \leq 0 \text{ для всех } j \in Q(X_0) \right\},$$

где $Q(X_0) = \{j \in [0 : N_1] \mid h_j(X_0) = 0\}$. Если $Q(X_0) = \emptyset$, то по определению полагаем $B(X_0) = E_n$.

Лемма 2.7. При выполнении условия Слейтера (2.11) справедливо равенство

$$\bar{\Gamma}(X_0) = B(X_0). \quad (2.12)$$

Доказательство. Утверждение очевидно, если $Q(X_0) = \emptyset$. Действительно, в этом случае X_0 — внутренняя точка множества Ω , поэтому $\Gamma(X_0) = \bar{\Gamma}(X_0) = E_n$. С другой стороны, по определению $B(X_0) = E_n$, и равенство (2.12) доказано.

Поэтому пусть $Q(X_0) \neq \emptyset$. Тогда $X_0 \in \partial\Omega$. Прежде всего установим, что

$$\bar{\Gamma}(X_0) \subset B(X_0). \quad (2.13)$$

Возьмем $V_0 \in \bar{\Gamma}(X_0)$. Тогда найдется последовательность векторов $\{V_i\}$ такая, что

$$V_0 = \lim_{i \rightarrow \infty} V_i; \quad V_i \in \Gamma(X_0), \quad i = 1, 2, \dots$$

Зафиксируем любое i . Для всех $j \in [0 : N_1]$ и всех $\alpha \in (0, \alpha_0(V_i))$ по определению $\Gamma(X_0)$ будет

$$h_j(X_0 + \alpha V_i) \leq 0.$$

Отсюда следует

$$h_j(X_0 + \alpha V_i) = h_j(X_0) + \alpha \left(\frac{\partial h_j(X_0 + \tau_i V_i)}{\partial X}, V_i \right) \leq 0,$$

где $\tau_i \stackrel{\text{df}}{=} \tau_i(\alpha, j) \in (0, \alpha)$.

Если $j \in Q(X_0)$, то $h_j(X_0) = 0$, и поэтому

$$\left(\frac{\partial h_j(X_0 + \tau_i(\alpha, j) V_i)}{\partial X}, V_i \right) \leq 0.$$

Последнее неравенство справедливо для всех $\alpha \in (0, \alpha_0(V_i))$. Устремляя α к нулю, получим

$$\left(\frac{\partial h_j(X_0)}{\partial X}, V_i \right) \leq 0.$$

В этом неравенстве перейдем к пределу при $i \rightarrow \infty$. Это дает для всех $j \in Q(X_0)$

$$\left(\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) \leq 0,$$

т. е. $V_0 \in B(X_0)$, и включение (2.13) доказано.

Докажем теперь, что

$$B(X_0) \in \bar{\Gamma}(X_0). \quad (2.14)$$

Пусть $V_0 \in B(X_0)$. Предположим вначале, что

$$\left(\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) < 0 \quad \text{для всех } j \in Q(X_0).$$

Поскольку

$$h_j(X_0 + \alpha V_0) = h_j(X_0) + \alpha \left(\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) + o_j(\alpha),$$

то найдется такое $\alpha_0 > 0$, что для всех $j \in Q(X_0)$ и $\alpha \in [0, \alpha_0]$ будет

$$h_j(X_0 + \alpha V_0) \leq h_j(X_0) = 0.$$

Далее $h_j(X_0) < 0$ при $j \notin Q(X_0)$. По непрерывности функций $h_j(X)$ найдется $\alpha_1 > 0$, $\alpha_1 \leq \alpha_0$, такое, что при $j \notin Q(X_0)$ и $\alpha \in [0, \alpha_1]$ будет

$$h_j(X_0 + \alpha V_0) < 0.$$

Итак, $X_0 + \alpha V_0 \in \Omega$ для всех $\alpha \in [0, \alpha_1]$, и потому $V_0 \in \Gamma(X_0)$.

Пусть теперь V_0 — произвольный вектор из $B(X_0)$. Поскольку $X_0 \in \partial\Omega$, то в силу леммы 2.1 найдется вектор $g_0 \in E_n$, $\|g_0\| = 1$, такой, что

$$\left(\frac{\partial h_j(X_0)}{\partial X}, g_0 \right) < 0 \quad \text{для всех } j \in Q(X_0).$$

При любых $\varepsilon > 0$ и $j \in Q(X_0)$ будем иметь

$$\left(V_0 + \varepsilon g_0, \frac{\partial h_j(X_0)}{\partial X} \right) < 0,$$

так что по доказанному $V_0 + \varepsilon g_0 \in \Gamma(X_0)$. Устремляя ε к нулю, получим $V_0 \in \bar{\Gamma}(X_0)$. Соотношение (2.14), а с ним и лемма, доказаны.

Итак,

$$\bar{\Gamma}(X_0) = \begin{cases} \left\{ V \mid \left(V, \frac{\partial h_j(X_0)}{\partial X} \right) \leq 0 \text{ для всех } j \in Q(X_0) \right\}, & \text{если } Q(X_0) \neq \emptyset; \\ E_n, & \text{если } Q(X_0) = \emptyset. \end{cases}$$

4. Через $K(X_0)$ обозначим выпуклую коническую оболочку множества $\tilde{H}_1(X_0) = \left\{ V = - \frac{\partial h_j(X_0)}{\partial X} \mid j \in Q(X_0) \right\}$:

$$K(X_0) = \left\{ V = - \sum_{j \in Q(X_0)} \lambda_j \frac{\partial h_j(X_0)}{\partial X} \mid \lambda_j \geq 0 \right\}.$$

Если $Q(X_0) = \emptyset$, то по определению полагаем $K(X_0) = \{0\}$. Пусть $Q(X_0) \neq \emptyset$, т. е. $X_0 \in \partial\Omega$. В силу леммы 2.6 Приложения II заключаем, что $K(X_0)$ является замкнутым выпуклым конусом.

Лемма 2.8. При выполнении условия Слейтера (2.2) справедливо равенство

$$\Gamma^+(X_0) = K(X_0). \quad (2.15)$$

Доказательство. В случае $Q(X_0) = \emptyset$ утверждение леммы очевидно. Поэтому пусть $Q(X_0) \neq \emptyset$.

Докажем, что

$$B(X_0) = K^+(X_0). \quad (2.16)$$

Возьмем $V_0 \in B(X_0)$. Имеем

$$\left(- \frac{\partial h_j(X_0)}{\partial X}, V_0 \right) \geq 0 \quad \text{для всех } j \in Q(X_0).$$

Отсюда следует

$$(Z, V_0) \geq 0 \quad \text{для всех } Z \in K(X_0),$$

т. е. по определению сопряженного конуса $V_0 \in K^+(X_0)$. Таким образом,

$$B(X_0) \subset K^+(X_0). \quad (2.17)$$

Пусть теперь $V_0 \in K^+(X_0)$. Это означает, что

$$(Z, V_0) \geq 0 \quad \text{для всех } Z \in K(X_0).$$

В частности, для всех $j \in Q(X_0)$

$$\left(-\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) \geq 0,$$

т. е. $V_0 \in B(X_0)$. Таким образом,

$$K^+(X_0) \subset B(X_0). \quad (2.18)$$

Объединяя (2.17) и (2.18), получим (2.16). В равенстве (2.16) перейдем к сопряженным конусам

$$B^+(X_0) = K^{++}(X_0). \quad (2.19)$$

В силу леммы 2.7

$$B^+(X_0) = \Gamma^+(X_0). \quad (2.20)$$

С другой стороны, поскольку $K(X_0)$ — выпуклый замкнутый конус, то в силу леммы 2.3 Приложения II

$$K^{++}(X_0) = K(X_0). \quad (2.21)$$

Объединяя (2.19), (2.20) и (2.21), получим (2.15).

Лемма доказана.

Из леммы 2.8 следует

Теорема 2.1. Если $X_0 \in \Omega$ и выполнено условие Слейтера (2.2), то

$$\Gamma^+(X_0) = \begin{cases} \left\{ V = - \sum_{j \in Q(X_0)} \lambda_j \frac{\partial h_j(X_0)}{\partial X} \mid \lambda_j \geq 0 \right\}, \\ \text{если } Q(X_0) \neq \emptyset; \\ \{0\}, \text{ если } Q(X_0) = \emptyset. \end{cases}$$

5. Допустим, что среди функций $h_j(X)$, $j \in [0 : N_1]$, есть линейные функции

$$h_j(X) = (A_j, X) + b_j, \quad j \in [N' + 1 : N_1]. \quad (2.22)$$

Тогда условие Слейтера (2.2), выполнение которого требовалось в теореме 2.1, может быть заменено более слабым условием. Достаточно потребовать существования точки $\bar{X}$ такой, что

$$\left. \begin{aligned} h_j(\bar{X}) &< 0 \quad \text{для } j \in [0 : N'], \\ h_j(\bar{X}) &\leq 0 \quad \text{для } j \in [N' + 1 : N_1]. \end{aligned} \right\} \quad (2.23)$$

Условие (2.23) будем называть *обобщенным условием Слейтера*.

Докажем теперь, что теорема 2.1 справедлива при выполнении обобщенного условия Слейтера (2.23). По существу нужно пересмотреть лишь доказательство соотношения $B(X_0) \subset \bar{\Gamma}(X_0)$ в лемме 2.7, ибо только здесь мы пользовались условием Слейтера.

Прежде всего заметим, что из (2.23) следует существование вектора W_0 такого, что

$$\left. \begin{aligned} \left(\frac{\partial h_j(X_0)}{\partial X}, W_0 \right) &< 0 \quad \text{при } j \in [0 : N'] \cap Q(X_0), \\ \left(\frac{\partial h_j(X_0)}{\partial X}, W_0 \right) &\leq 0 \quad \text{при } j \in [N' + 1 : N_1] \cap Q(X_0). \end{aligned} \right\} \quad (2.24)$$

Действительно, положим $W_0 = \bar{X} - X_0$, где $\bar{X}$ — точка из (2.23). В силу леммы 3.2 Приложения II при $j \in Q(X_0)$ будем иметь

$$\left(\frac{\partial h_j(X_0)}{\partial X}, \bar{X} - X_0 \right) \leq h_j(\bar{X}) - h_j(X_0) = h_j(\bar{X}).$$

Отсюда и из (2.23) очевидным образом следует (2.24).

Далее, если вектор $V_0 \in B(X_0)$ таков, что

$$\left. \begin{aligned} \left(\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) &< 0 \quad \text{при } j \in [0 : N'] \cap Q(X_0), \\ \left(\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) &\leq 0 \quad \text{при } j \in [N' + 1 : N_1] \cap Q(X_0), \end{aligned} \right\} \quad (2.25)$$

то $V_0 \in \Gamma(X_0)$. Действительно, при $j \in [0 : N'] \cap Q(X_0)$ и малых $\alpha \geq 0$, так же, как в основном тексте, имеем

$$h_j(X_0 + \alpha V_0) \leq h_j(X_0) = 0.$$

Для $j \in [N' + 1 : N_1] \cap Q(X_0)$ получим в силу (2.22) и (2.25) при всех $\alpha \geq 0$

$$h_j(X_0 + \alpha V_0) = h_j(X_0) + \alpha \left(\frac{\partial h_j(X_0)}{\partial X}, V_0 \right) \leq 0.$$

Итак, при достаточно малых $\alpha \geq 0$ и $j \in Q(X_0)$ будет

$$h_j(X_0 + \alpha V_0) \leq 0.$$

С другой стороны, при малых $\alpha \geq 0$ и $j \notin Q(X_0)$

$$h_j(X_0 + \alpha V_0) < 0.$$

Получили, что $h_j(X_0 + \alpha V_0) \leq 0$ для всех $j \in [0 : N_1]$ и $\alpha \in [0, \alpha_0(V_0)]$, где $\alpha_0(V_0) > 0$. Но это равносильно включению $V_0 \in \Gamma(X_0)$.

После сделанных замечаний доказательство соотношения $B(X_0) \subset \bar{\Gamma}(X_0)$ проводится так же, как в лемме 2.7.

Утверждение доказано.

Замечание 1. Если все функции $h_j(X)$ линейны, то (2.23) переписывается в виде

$$h_j(\bar{X}) \leq 0 \quad \text{для} \quad j \in [0 : N_1].$$

Таким образом, в этом случае обобщенное условие Слейтера заменяется требованием непустоты множества Ω .

Замечание 2. Поскольку каждое равенство вида

$$(A_j, X) + b_j = 0$$

может быть заменено системой неравенств

$$\begin{aligned} (A_j, X) + b_j &\leq 0, \\ -(A_j, X) - b_j &\leq 0, \end{aligned}$$

то в определение множества Ω могут входить линейные равенства. Однако в этом случае условие Слейтера никогда не выполняется, а обобщенное условие Слейтера может и выполняться.

§ 3. Необходимые условия минимакса

1. Пусть выпуклое замкнутое множество Ω задается в виде (2.1), и пусть Ω' — некоторое открытое множество, содержащее внутри себя Ω . Положим

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X),$$

где $f_i(X)$, $i \in [0 : N]$, — непрерывно дифференцируемые на Ω' функции.

Рассматривается задача минимизации функции $\varphi(X)$ на Ω . Общей задаче такого рода была посвящена глава IV. Результаты настоящего параграфа учитывают специфику задания множества Ω .

Напомним, что имеет место следующая теорема (см. § 3 главы IV): для того чтобы в точке $X^* \in \Omega$ функция $\varphi(X)$ достигала своего минимального на Ω значения, необходимо, а в случае выпуклости $\varphi(X)$ на Ω и достаточно, чтобы

$$\Gamma^+(X^*) \cap L(X^*) \neq \emptyset, \quad (3.1)$$

где

$$L(X^*) = \text{co } H(X^*), \quad H(X^*) = \left\{ \frac{\partial f_i(X^*)}{\partial X} \mid i \in R(X^*) \right\},$$

а $\Gamma^+(X^*)$ — конус, сопряженный к конусу возможных направлений множества Ω в точке X^* .

В случае, если множество Ω задается соотношением (2.1) и выполнено условие Слейтера (2.2), вид конуса $\Gamma^+(X^*)$ установлен теоремой 2.1.

При построении алгоритмов для минимизации функции $\varphi(X)$ оказывается полезной следующая

Теорема 3.1. Пусть множество Ω задано в виде (2.1) и выполнено условие Слейтера (2.2). Тогда соотношение (3.1) эквивалентно включению

$$0 \in \tilde{L}(X^*), \quad (3.2)$$

где

$$\tilde{L}(X^*) = \text{co } \tilde{H}(X^*), \quad \tilde{H}(X^*) = H(X^*) \cup H_1(X^*),$$

$$H(X^*) = \left\{ \frac{\partial f_i(X^*)}{\partial X} \mid i \in R(X^*) \right\},$$

$$H_1(X^*) = \left\{ \frac{\partial h_j(X^*)}{\partial X} \mid j \in Q(X^*) \right\},$$

$$R(X^*) = \{i \in [0 : N] \mid f_i(X^*) = \varphi(X^*)\},$$

$$Q(X^*) = \{j \in [0 : N_1] \mid h_j(X^*) = 0\}.$$

Доказательство. Утверждение теоремы очевидно, если $Q(X^*) = \emptyset$. Действительно, в этом случае $\tilde{L}(X^*) = L(X^*)$, и в силу теоремы 2.1

$$\Gamma^+(X^*) = \{0\}.$$

Таким образом, и (3.1) и (3.2) сводятся к включению

$$0 \in L(X^*).$$

Это дает нам право в дальнейшем считать, что $Q(X^*) \neq \emptyset$, т. е. $X^* \in \partial\Omega$.

Пусть в точке X^* выполнено условие (3.2). Покажем, что тогда имеет место и (3.1).

Из (3.2) следует

$$0 = \sum_{i \in R(X^*)} \lambda_{1i} \frac{\partial f_i(X^*)}{\partial X} + \sum_{j \in Q(X^*)} \lambda_{2j} \frac{\partial h_j(X^*)}{\partial X}, \quad (3.3)$$

где $\lambda_{1i} \geq 0$, $\lambda_{2j} \geq 0$,

$$\sum_{i \in R(X^*)} \lambda_{1i} + \sum_{j \in Q(X^*)} \lambda_{2j} = 1.$$

При этом

$$a = \sum_{i \in R(X^*)} \lambda_{1i} > 0,$$

ибо в противном случае оказалось бы, что

$$\sum_{j \in Q(X^*)} \lambda_{2j} \frac{\partial h_j(X^*)}{\partial X} = 0; \quad \lambda_{2j} \geq 0, \quad \sum_{j \in Q(X^*)} \lambda_{2j} = 1,$$

т. е. $0 \in L_1(X^*)$, где $L_1(X^*) = \text{co } H_1(X^*)$, но это противоречит лемме 2.1.

Итак, $a > 0$. Перепишем (3.3) в виде

$$0 = \sum_{i \in R(X^*)} \lambda'_{1i} \frac{\partial f_i(X^*)}{\partial X} + \sum_{j \in Q(X^*)} \lambda'_{2j} \frac{\partial h_j(X^*)}{\partial X}, \quad (3.4)$$

где

$$\lambda'_{1i} = \frac{\lambda_{1i}}{a}, \quad \lambda'_{2j} = \frac{\lambda_{2j}}{a}.$$

Ясно, что

$$\lambda'_{1i} \geq 0, \quad \lambda'_{2j} \geq 0, \quad \sum_{i \in R(X^*)} \lambda'_{1i} = 1.$$

Теперь имеем

$$\sum_{i \in R(X^*)} \lambda'_{1i} \frac{\partial f_i(X^*)}{\partial X} \in L(X^*), \quad (3.5)$$

и в силу (3.4) и теоремы 2.1

$$\sum_{i \in R(X^*)} \lambda'_{1i} \frac{\partial f_i(X^*)}{\partial X} = - \sum_{j \in Q(X^*)} \lambda'_{2j} \frac{\partial h_j(X^*)}{\partial X} \in \Gamma^+(X^*). \quad (3.6)$$

Из (3.5) и (3.6) следует (3.1).

Докажем теперь обратное. Предположим, что выполнено соотношение (3.1). Требуется доказать, что $0 \in \tilde{L}(X^*)$.

Пусть $Z_0 \in \Gamma^+(X^*) \cap L(X^*)$, т. е.

$$Z_0 = - \sum_{j \in Q(X^*)} \alpha_j \frac{\partial h_j(X^*)}{\partial X}, \quad \alpha_j \geq 0;$$

$$Z_0 = \sum_{i \in R(X^*)} \lambda_i \frac{\partial f_i(X^*)}{\partial X}, \quad \lambda_i \geq 0, \quad \sum_{i \in R(X^*)} \lambda_i = 1.$$

Имеем

$$\sum_{i \in R(X^*)} \lambda_i \frac{\partial f_i(X^*)}{\partial X} + \sum_{j \in Q(X^*)} \alpha_j \frac{\partial h_j(X^*)}{\partial X} = 0. \quad (3.7)$$

Введем обозначения:

$$c = \sum_{j \in Q(X^*)} \sigma_j; \quad \lambda_{1i} = \frac{\lambda_i}{1+c}, \quad i \in R(X^*);$$

$$\lambda_{2j} = \frac{\alpha_j}{1+c}, \quad j \in Q(X^*).$$

Поделив равенство (3.7) на $1+c$, получим

$$\sum_{i \in R(X^*)} \lambda_{1i} \frac{\partial f_i(X^*)}{\partial X} + \sum_{j \in Q(X^*)} \lambda_{2j} \frac{\partial h_j(X^*)}{\partial X} = 0,$$

причем

$$\lambda_{1i} \geq 0, \quad \lambda_{2j} \geq 0, \quad \sum_{i \in R(X^*)} \lambda_{1i} + \sum_{j \in Q(X^*)} \lambda_{2j} = 1.$$

Но это и означает, что $0 \in \tilde{L}(X^*)$.

Теорема доказана.

З а м е ч а н и е 1. Условие Слейтера в формулировке теоремы 3.1 не может быть заменено обобщенным условием Слейтера.

П р и м е р. Пусть

$$\Omega = \{X = (x_1, x_2) \mid x_1 + x_2 = 0\},$$

$$\varphi(X) = f_0(X) = (X, X).$$

В этом случае обобщенное условие Слейтера выполнено. Положив $A = (1, 1)$, перепишем представление для Ω в виде

$$\Omega = \{X \in E_2 \mid (A, X) \leq 0, -(A, X) \leq 0\}.$$

Нетрудно доказать, что для всех $X \in \Omega$ будет

$$0 \in \tilde{L}(X).$$

Действительно, $H_1(X)$ состоит из двух векторов: A и $-A$, так что $0 \in \text{co } H_1(X)$, и тем более $0 \in \tilde{L}(X)$.

С другой стороны, $\Gamma^+(X) = \{\lambda A \mid \lambda \text{ произвольно}\}$, $L(X) = \{2X\}$. Для $X_0 = (1, -1) \in \Omega$, например, будем иметь

$$L(X_0) \cap \Gamma^+(X_0) = \emptyset.$$

Итак, в рассматриваемом случае соотношения (3.1) и (3.2) неравносильны.

З а м е ч а н и е 2. Множество индексов $[0 : N_1]$ произвольным образом разобьем на два множества $[0 : N']$ и $[N' + 1 : N_1]$. Введем обозначения

$$Q_1(X) = \{j \in [0 : N'] \mid h_j(X) = 0\},$$

$$Q_2(X) = \{j \in [N' + 1 : N_1] \mid h_j(X) = 0\}.$$

Очевидно, $Q_1(X) \cup Q_2(X) = Q(X)$.

Имеет место более общее, чем теорема 3.1, утверждение:

Пусть выполнено условие Слейтера (2.2) и $X^ \in \Omega$; тогда соотношение (3.1) эквивалентно следующему соотношению:*

$$\Gamma'^+(X^*) \cap \tilde{L}'(X^*) \neq \emptyset,$$

где

$$\begin{aligned}\tilde{L}'(X^*) &= \text{co } \tilde{H}'(X^*), \quad \tilde{H}'(X^*) = H(X^*) \cup H_1'(X^*); \\ H_1'(X^*) &= \left\{ \frac{\partial h_j(X^*)}{\partial X} \mid j \in Q_1(X^*) \right\}; \\ \Gamma'^+(X^*) &\stackrel{\text{df}}{=} \begin{cases} \left\{ V = - \sum_{j \in Q_2(X^*)} \lambda_j \frac{\partial h_j(X^*)}{\partial X} \mid \lambda_j \geq 0 \right\}, \\ \text{если } Q_2(X^*) \neq \emptyset; \\ \{0\}, \text{ если } Q_2(X^*) = \emptyset. \end{cases}\end{aligned}$$

Доказательство этого утверждения вполне аналогично доказательству теоремы 3.1.

2. Определение. Точка $X^* \in \Omega$ называется *стационарной точкой функции* $\varphi(X)$ на множестве Ω , если имеет место (3.1).

Пусть точка $X_0 \in \Omega$ не является стационарной. Тогда

$$d(X_0) \stackrel{\text{df}}{=} \min_{\substack{Z \in L(X_0) \\ V \in \Gamma^+(X_0)}} \|Z - V\| = \|Z(X_0) - V(X_0)\| > 0,$$

где $Z(X_0) \in L(X_0)$, $V(X_0) \in \Gamma^+(X_0)$.

В § 3 главы IV было показано, что направление

$$g(X_0) = \frac{V(X_0) - Z(X_0)}{\|V(X_0) - Z(X_0)\|}, \quad \|g(X_0)\| = 1,$$

является направлением наискорейшего спуска функции $\varphi(X)$ в точке X_0 на множестве Ω . Там же отмечалось, что $g(X_0) \in \bar{\Gamma}(X_0)$.

Может оказаться, что $g(X_0) \notin \Gamma(X_0)$, т. е. ни одна точка луча $X = X_0(\alpha) \stackrel{\text{df}}{=} X_0 + \alpha g(X_0)$, $\alpha > 0$, не принадлежит Ω . Вот почему мы не можем непосредственно использовать направление наискорейшего спуска для разработки методов последовательных приближений при нахождении стационарных точек.

Предположим, что выполнено условие Слейтера (2.2), и пусть $Q(X_0) \neq \emptyset$. Найдем

$$\tilde{d}(X_0) \stackrel{\text{df}}{=} \min_{Z \in \tilde{L}(X_0)} \|Z\| = \|\tilde{Z}(X_0)\|. \quad (3.8)$$

Так как X_0 не является стационарной точкой, то в силу теоремы 3.1 $\tilde{d}(X_0) > 0$ и $\tilde{Z}(X_0) \neq 0$. Из (3.8) следует, что

направление

$$\tilde{g}(X_0) = -\frac{\tilde{Z}(X_0)}{\|\tilde{Z}(X_0)\|}, \quad \|\tilde{g}(X_0)\| = 1, \quad (3.9)$$

обладает тем свойством, что

$$\left(\frac{\partial f_i(X_0)}{\partial X}, \tilde{g}(X_0)\right) \leq -\tilde{d}(X_0) < 0, \quad i \in R(X_0); \quad (3.10)$$

$$\left(\frac{\partial h_j(X_0)}{\partial X}, \tilde{g}(X_0)\right) \leq -\tilde{d}(X_0) < 0, \quad j \in Q(X_0). \quad (3.11)$$

Рассмотрим луч $X = X_0(\alpha) \stackrel{\text{def}}{=} X_0 + \alpha \tilde{g}(X_0)$, $\alpha \geq 0$. В силу (3.11) имеем при достаточно малых $\alpha > 0$

$$h_j(X_0 + \alpha \tilde{g}(X_0)) < h_j(X_0) = 0 \quad \text{для всех } j \in Q(X_0).$$

В силу непрерывности функций $h_j(X)$ для $j \notin Q(X_0)$ при малых $\alpha > 0$ будет

$$h_j(X_0 + \alpha \tilde{g}(X_0)) < 0.$$

Таким образом, $X_0 + \alpha \tilde{g}(X_0) \in \Omega$ при достаточно малых $\alpha > 0$.

С другой стороны, в силу (3.10) при малых $\alpha > 0$

$$\varphi(X_0 + \alpha \tilde{g}(X_0)) < \varphi(X_0).$$

Получили, что направление $\tilde{g}(X_0)$ является направлением убывания функции $\varphi(X)$ в точке X_0 и, кроме того, в этом направлении можно двигаться, оставаясь в множестве Ω . Это обстоятельство и будет использовано нами в дальнейшем.

Отметим, однако, что направление $\tilde{g}(X_0)$ не является, вообще говоря, направлением наискорейшего спуска функции $\varphi(X)$ в точке X_0 на множестве Ω .

§ 4*. Зависимость направления спуска от способа задания множества Ω

Установим зависимость полученного по формуле (3.9) направления спуска для функции $\varphi(X)$ от способа задания множества Ω . Множество Ω было определено соотношением (2.1). Ясно, что его можно задать и так:

$$\Omega = \{X \mid t_j h_j(X) \leq 0 \quad \text{для всех } j \in [0 : N_1]\}, \quad (4.1)$$

где t_j — любые положительные числа. Тогда вместо многогранника $\tilde{L}(X)$ (см. теорему 3.1) появится многогранник

$$\tilde{L}_t(X) = \text{co } \tilde{H}_t(X),$$

где $t = (t_0, \dots, t_{N_1})$,

$$\tilde{H}_t(X) = H(X) \cup H_t(X),$$

$$H_t(X) = \left\{ t_j \frac{\partial h_j(X)}{\partial X} \mid j \in Q(X) \right\},$$

и необходимое условие (3.2) заменится условием

$$0 \in \tilde{L}_t(X^*). \quad (4.2)$$

Конечно, условия (3.2) и (4.2) эквивалентны, поэтому неважно, какое из них использовать для проверки точки на стационарность.

Пусть $X_0 \in \Omega$ не является стационарной точкой функции $\varphi(X)$ на Ω . Тогда направление

$$\tilde{g}_t(X_0) = - \frac{\tilde{Z}_t(X_0)}{\|\tilde{Z}_t(X_0)\|}, \quad \|\tilde{g}_t(X_0)\| = 1,$$

где

$$\tilde{d}_t(X_0) \stackrel{\text{def}}{=} \min_{Z \in \tilde{L}_t(X_0)} \|Z\| = \|\tilde{Z}_t(X_0)\|,$$

существенно зависит от вектора коэффициентов t . При этом при любом положительном t (это означает, что все координаты вектора t положительны) направление $\tilde{g}_t(X_0)$ является направлением убывания функции $\varphi(X)$ и существует такое $\alpha_0(t, X_0) > 0$, что $X_0 + \alpha \tilde{g}_t(X_0) \in \Omega$ при всех $\alpha \in [0, \alpha_0(t, X_0)]$.

Введем обозначение

$$m(t) = \min_{j \in [0 : N_1]} t_j.$$

Лемма 4.1. Если $X_0 \in \Omega$ не является стационарной точкой функции $\varphi(X)$ на Ω и если выполнено условие Слейтера (2.2), то

$$\lim_{m(t) \rightarrow \infty} \tilde{d}_t(X_0) = d(X_0),$$

где (см. п. 2 § 3)

$$d(X_0) = \min_{\substack{Z \in L(X_0) \\ V \in \Gamma^+(X_0)}} \|Z - V\|.$$

Доказательство. Если $Q(X_0) = \emptyset$, то

$$\Gamma^+(X_0) = \{0\}, \quad \tilde{L}_t(X_0) = L(X_0),$$

так что $\tilde{d}_t(X_0) = d(X_0)$ при всех положительных t . Поэтому будем считать, что $Q(X_0) \neq \emptyset$, т. е. $X_0 \in \partial\Omega$. Из леммы 2.1 следует

$$\psi_1(X_0) \stackrel{df}{=} \min_{\|g\|=1} \max_{j \in Q(X_0)} \left(\frac{\partial h_j(X_0)}{\partial X}, g \right) = -a < 0. \quad (4.3)$$

Покажем, что для любых $\gamma_j \geq 0$, $j \in [0:N_1]$, будет

$$\left\| \sum_{j \in Q(X_0)} \gamma_j \frac{\partial h_j(X_0)}{\partial X} \right\| \geq a \sum_{j \in Q(X_0)} \gamma_j. \quad (4.4)$$

Действительно, если $\sum_{j \in Q(X_0)} \gamma_j = 0$, то неравенство (4.4) очевидно. Пусть

$$\sum_{j \in Q(X_0)} \gamma_j = b > 0.$$

Тогда

$$\sum_{j \in Q(X_0)} \gamma_j \frac{\partial h_j(X_0)}{\partial X} = b \left[\sum_{j \in Q(X_0)} \gamma'_j \frac{\partial h_j(X_0)}{\partial X} \right],$$

где

$$\gamma'_j = \frac{\gamma_j}{b}, \quad \sum_{j \in Q(X_0)} \gamma'_j = 1.$$

Так как (см. замечание к лемме 2.1)

$$W_0 \stackrel{df}{=} \sum_{j \in Q(X_0)} \gamma'_j \frac{\partial h_j(X_0)}{\partial X} \in L_1(X_0),$$

то на основании (4.3) имеем

$$\|W_0\| \geq -\psi_1(X_0) = a, \quad (4.5)$$

ибо

$$-\psi_1(X_0) = \min_{Z \in L_1(X_0)} \|Z\|.$$

Из (4.5) и следует (4.4).

Напомним теперь, что

$$\tilde{d}_t(X_0) = \min_{\lambda \in \Lambda(t)} \left\| \sum_{i \in R(X_0)} \lambda_{1i} \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \lambda_{2j} t_j \frac{\partial h_j(X_0)}{\partial X} \right\|,$$

где

$$\Lambda(t) = \left\{ \lambda \mid \lambda_{1i} \geq 0, \lambda_{2j} \geq 0, \sum_{i \in R(X_0)} \lambda_{1i} + \sum_{j \in Q(X_0)} \lambda_{2j} = 1 \right\},$$

и

$$d(X_0) = \min_{\lambda \in \Lambda} \left\| \sum_{i \in R(X_0)} \lambda_{1i} \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \lambda_{2j} \frac{\partial h_j(X_0)}{\partial X} \right\|,$$

где

$$\Lambda = \left\{ \lambda \mid \lambda_{1i} \geq 0, \lambda_{2j} \geq 0, \sum_{i \in R(X_0)} \lambda_{1i} = 1 \right\}.$$

Положив

$$c = \max_{i \in R(X_0)} \left\| \frac{\partial f_i(X_0)}{\partial X} \right\|,$$

получим на основании (4.4)

$$\begin{aligned} \tilde{d}_t(X_0) &= \left\| \sum_{i \in R(X_0)} \lambda_{1i}(t) \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \lambda_{2j}(t) t_j \frac{\partial h_j(X_0)}{\partial X} \right\| \geq \\ &\geq a \sum_{j \in Q(X_0)} \lambda_{2j}(t) t_j - c \geq a m(t) \sum_{j \in Q(X_0)} \lambda_{2j}(t) - c. \end{aligned} \quad (4.6)$$

Так как, с другой стороны,

$$\tilde{d}_t(X_0) \leq \max_{i \in R(X_0)} \left\| \frac{\partial f_i(X_0)}{\partial X} \right\| = c,$$

то из (4.6) следует

$$\begin{aligned} \sum_{j \in Q(X_0)} \lambda_{2j}(t) \xrightarrow{m(t) \rightarrow \infty} 0, \\ \mu(t) \stackrel{df}{=} \sum_{i \in R(X_0)} \lambda_{1i}(t) \xrightarrow{m(t) \rightarrow \infty} 1. \end{aligned} \quad (4.7)$$

Далее, при достаточно больших $m(t)$ имеем

$$\tilde{d}_t(X_0) = \mu(t) \left\| \sum_{i \in R(X_0)} \lambda'_{1i}(t) \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \lambda'_{2j}(t) \frac{\partial h_j(X_0)}{\partial X} \right\|,$$

где

$$\lambda'_{1i}(t) = \frac{\lambda_{1i}(t)}{\mu(t)}, \quad \lambda'_{2j}(t) = \frac{\lambda_{2j}(t) t_j}{\mu(t)}.$$

Поскольку

$$\lambda'_{1i}(t) \geq 0, \quad \lambda'_{2j}(t) \geq 0, \quad \sum_{i \in R(X_0)} \lambda'_{1i}(t) = 1,$$

то по определению $d(X_0)$ получим следующее неравенство:

$$\tilde{d}_t(X_0) \geq \mu(t) d(X_0). \quad (4.8)$$

Докажем теперь, что $d(X_0) \leq \tilde{d}_t(X_0)$ при всех положительных t . Действительно, запишем цепочку равенств

$$\begin{aligned} d(X_0) &= \left\| \sum_{i \in R(X_0)} \tilde{\lambda}_{1i} \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \tilde{\lambda}_{2j} \frac{\partial h_j(X_0)}{\partial X} \right\| = \\ &= \left\| \sum_{i \in R(X_0)} \tilde{\lambda}_{1i} \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \tilde{\lambda}'_{2j} t_j \frac{\partial h_j(X_0)}{\partial X} \right\| = \\ &= b \left\| \sum_{i \in R(X_0)} \lambda'_{1i} \frac{\partial f_i(X_0)}{\partial X} + \sum_{j \in Q(X_0)} \lambda'_{2j} t_j \frac{\partial h_j(X_0)}{\partial X} \right\|. \end{aligned}$$

Здесь

$$\begin{aligned} \tilde{\lambda}'_{2j} &= \tilde{\lambda}_{2j}/t_j, \\ b &= \sum_{i \in R(X_0)} \tilde{\lambda}_{1i} + \sum_{j \in Q(X_0)} \tilde{\lambda}'_{2j} t_j = 1 + \sum_{j \in Q(X_0)} \tilde{\lambda}_{2j} \geq 1, \\ \lambda'_{1i} &= \tilde{\lambda}_{1i}/b, \quad \lambda'_{2j} = \tilde{\lambda}'_{2j}/b. \end{aligned}$$

Ясно, что $\lambda'_{1i} \geq 0$, $\lambda'_{2j} \geq 0$, $\sum_{i \in R(X_0)} \lambda'_{1i} + \sum_{j \in Q(X_0)} \lambda'_{2j} = 1$.

Поэтому, учитывая определение $\tilde{d}_t(X_0)$, получим

$$d(X_0) \geq b \tilde{d}_t(X_0) \geq \tilde{d}_t(X_0). \quad (4.9)$$

Из (4.8) и (4.9) следует

$$d(X_0) \geq \tilde{d}_t(X_0) \geq \mu(t) d(X_0).$$

Осталось перейти к пределу при $m(t) \rightarrow \infty$ в последних неравенствах и воспользоваться соотношением (4.7).

Лемма доказана.

Теорема 4.1. Если $X_0 \in \Omega$ не является стационарной точкой функции $\Phi(X)$ на множестве Ω и если выполнено условие Слейтера (2.2), то

$$\tilde{g}_t(X_0) \xrightarrow{m(t) \rightarrow \infty} g(X_0),$$

где $g(X_0)$ — направление наискорейшего спуска функции $\Phi(X)$ в точке X_0 на множестве Ω .

Доказательство. На основании леммы 4.1 имеем

$$\tilde{d}_t(X_0) \xrightarrow{m(t) \rightarrow \infty} d(X_0). \quad (4.10)$$

Но

$$\tilde{d}_t(X_0) = \min_{Z \in \tilde{L}_t(X_0)} \|Z\| = \|Z(t)\|,$$

$$d(X_0) = \min_{Z \in G(X_0)} \|Z\| = \|Z_0\|,$$

где

$$G(X_0) = L(X_0) - \Gamma^+(X_0) = \{Z - V \mid Z \in L(X_0), V \in \Gamma^+(X_0)\}.$$

Заметим, что множество $G(X_0)$ замкнутое, ибо замкнутыми являются множества $L(X_0)$ и $\Gamma^+(X_0)$ и $L(X_0)$ ограничено.

Из доказательства леммы 4.1 следует, что справедливо представление

$$Z(t) = \mu(t) Z'(t), \quad (4.11)$$

где $Z'(t) \in G(X_0)$ и $\mu(t) \xrightarrow{m(t) \rightarrow \infty} 1$. В силу (4.10) можно считать, что множество векторов $\{Z'(t)\}$ для достаточно больших $m(t)$ ограничено.

Докажем следующее предельное соотношение:

$$Z'(t) \xrightarrow{m(t) \rightarrow \infty} Z_0. \quad (4.12)$$

Допустим противное. Тогда найдется такая последовательность $\{t_s\}$, что $Z'(t_s) \xrightarrow{m(t_s) \rightarrow \infty} Z' \neq Z_0$. Поскольку $Z'(t_s) \in G(X_0)$ и множество $G(X_0)$ замкнутое, то $Z' \in G(X_0)$. Учитывая, что Z_0 — единственная ближайшая к началу координат точка множества $G(X_0)$, получим

$$\|Z'\| > \|Z_0\|.$$

С другой стороны, в силу (4.10) и (4.11) имеем

$$\|Z'\| = \|Z_0\|.$$

Полученное противоречие убеждает нас в справедливости соотношения (4.12).

Теперь заметим, что

$$\dot{g}_i(X_0) = - \frac{Z(i)}{\|Z(i)\|} = - \frac{Z'(i)}{\|Z'(i)\|}, \quad g(X_0) = - \frac{Z_0}{\|Z_0\|}.$$

Отсюда и из (4.12) следует требуемое.

Теорема доказана.

§ 5. Множители Лагранжа и теорема Куна — Таккера

1. Пусть по-прежнему

$$\varphi(X) = \max_{i \in [0:N]} f_i(X),$$

где $f_i(X)$, $i \in [0:N]$, — непрерывно дифференцируемые на $\Omega' \subset E_n$ функции, и пусть множество $\Omega \subset \Omega'$ задано в виде

$$\Omega = \{X \mid \max_{j \in [0:N_1]} h_j(X) \leq 0\},$$

где функции $h_j(X)$, $j \in [0:N_1]$, — выпуклые и непрерывно дифференцируемые на E_n . Предположим также, что выполнено условие Слейтера

$$\inf_{X \in E_n} \max_{j \in [0:N_1]} h_j(X) < 0$$

(или обобщенное условие Слейтера (2.23)). В этом случае справедлива следующая

Теорема 5.1. Для того чтобы функция $\varphi(X)$ достигала в точке $X^ \in \Omega$ своего минимального на Ω значения, необходимо, а в случае выпуклости $\varphi(X)$ на Ω и достаточно, чтобы нашелся вектор*

$$\lambda^* = (\lambda_{10}^*, \dots, \lambda_{1N}^*, \lambda_{20}^*, \dots, \lambda_{2N_1}^*) \in E_{N+N_1+2}$$

такой, что

$$\text{I. } \sum_{i=0}^N \lambda_{1i}^* \frac{\partial f_i(X^*)}{\partial X} + \sum_{j=0}^{N_1} \lambda_{2j}^* \frac{\partial h_j(X^*)}{\partial X} = 0.$$

$$\text{II. } \lambda_{1i}^* \geq 0, \quad \sum_{i=0}^N \lambda_{1i}^* = 1; \quad \lambda_{1i}^* = 0, \quad \text{если } f_i(X^*) < \varphi(X^*).$$

$$\text{III. } \lambda_{2j}^* \geq 0; \quad \lambda_{2j}^* h_j(X^*) = 0, \quad j \in [0:N_1].$$

Доказательство. Необходимость. Пусть X^* — точка минимума функции $\varphi(X)$ на Ω . Будем считать для определенности, что множество

$$Q(X^*) = \{j \in [0 : N_1] \mid h_j(X^*) = 0\}$$

не пусто (в противном случае доказательство лишь упрощается). На основании теоремы 2.1 и 3.1 главы IV заключаем, что найдется точка Z_0 , общая для $L(X^*)$ и $\Gamma^+(X^*)$. Таким образом,

$$\begin{aligned} Z_0 &= \sum_{i \in R(X^*)} \alpha_i \frac{\partial f_i(X^*)}{\partial X}; & \alpha_i &\geq 0, & \sum_{i \in R(X^*)} \alpha_i &= 1; \\ Z_0 &= - \sum_{j \in Q(X^*)} \beta_j \frac{\partial h_j(X^*)}{\partial X}; & \beta_j &\geq 0, & j &\in Q(X^*). \end{aligned}$$

Отсюда следует

$$\sum_{i \in R(X^*)} \alpha_i \frac{\partial f_i(X^*)}{\partial X} + \sum_{j \in Q(X^*)} \beta_j \frac{\partial h_j(X^*)}{\partial X} = 0.$$

Компоненты вектора λ^* определим соотношениями:

$$\begin{aligned} \lambda_{1i}^* &= \alpha_i & \text{для } i \in R(X^*), & \lambda_{1i}^* = 0 & \text{для } i \notin R(X^*), \\ \lambda_{2j}^* &= \beta_j & \text{для } j \in Q(X^*), & \lambda_{2j}^* = 0 & \text{для } j \notin Q(X^*). \end{aligned}$$

Получим

$$\sum_{i=1}^N \lambda_{1i}^* \frac{\partial f_i(X^*)}{\partial X} + \sum_{j=0}^{N_1} \lambda_{2j}^* \frac{\partial h_j(X^*)}{\partial X} = 0.$$

Условия II и III также выполняются. Необходимость доказана.

Достаточность. Пусть $\varphi(X)$ выпуклая на Ω функция, $X^* \in \Omega$ и выполнены условия I—III. Для определенности будем считать, что $Q(X^*) \neq \emptyset$. Имеем в силу II и III

$$\begin{aligned} Z_0 &\stackrel{df}{=} \sum_{i=0}^N \lambda_{1i}^* \frac{\partial f_i(X^*)}{\partial X} = \sum_{i \in R(X^*)} \lambda_{1i}^* \frac{\partial f_i(X^*)}{\partial X} \in L(X^*), \\ Z_0' &\stackrel{df}{=} - \sum_{j=0}^{N_1} \lambda_{2j}^* \frac{\partial h_j(X^*)}{\partial X} = - \sum_{j \in Q(X^*)} \lambda_{2j}^* \frac{\partial h_j(X^*)}{\partial X} \in \Gamma^+(X^*). \end{aligned}$$

Поскольку $Z_0 = Z'_0$, то

$$L(X^*) \cap \Gamma^+(X^*) \neq \emptyset.$$

Учитывая теоремы 2.1 и 3.1 главы IV, заключаем, что в точке X^* функция $\varphi(X)$ достигает своего минимального на Ω значения.

Теорема доказана.

2. Введем в рассмотрение функцию

$$F(X, \lambda) = \sum_{i=0}^N \lambda_{1i} f_i(X) + \sum_{j=0}^{N_1} \lambda_{2j} h_j(X).$$

Функция $F(X, \lambda)$ называется функцией Лагранжа, а коэффициенты λ_{1i} , λ_{2j} — множителями Лагранжа. Предположим, что $f_i(X)$, $i \in [0 : N]$, и $h_j(X)$, $j \in [0 : N_1]$, заданы на всем E_n , и пусть

$$\Lambda = \{\lambda = (\lambda_{10}, \dots, \lambda_{1N}, \lambda_{20}, \dots, \lambda_{2N_1}) \mid \lambda_{1i} \geq 0,$$

$$\sum_{i=0}^N \lambda_{1i} = 1; \lambda_{2j} \geq 0, j \in [0 : N_1]\}.$$

Определение. Точка $[X^*, \lambda^*]$ называется седловой точкой функции $F(X, \lambda)$ на $E_n \times \Lambda$, если

$$F(X^*, \lambda) \leq F(X^*, \lambda^*) \leq F(X, \lambda^*) \quad (5.1)$$

для всех $X \in E_n$ и $\lambda \in \Lambda$.

Справедлива

Теорема 5.2 (Куна—Таккера). Пусть функции $f_i(X)$, $i \in [0 : N]$, и $h_j(X)$, $j \in [0 : N_1]$, выпуклы и непрерывно дифференцируемы на E_n и выполнено условие Слейтера (или обобщенное условие Слейтера). Для того чтобы функция $\varphi(X)$ достигала в точке X^* своего минимального на Ω значения, необходимо и достаточно, чтобы нашелся такой вектор $\lambda^* \in \Lambda$, что точка $[X^*, \lambda^*]$ является седловой точкой функции $F(X, \lambda)$ на $E_n \times \Lambda$.

Доказательство. Необходимость. Пусть X^* — точка минимума $\varphi(X)$ на Ω . Тогда в силу теоремы 5.1 найдется вектор $\lambda^* \in \Lambda$, для которого выполнены условия I—III. В частности,

$$\sum_{i=0}^N \lambda_{1i}^* \frac{\partial f_i(X^*)}{\partial X} + \sum_{j=0}^{N_1} \lambda_{2j}^* \frac{\partial h_j(X^*)}{\partial X} = 0. \quad (5.2)$$

Но слева в (5.2) стоит $\partial F(X^*, \lambda^*)/\partial X$, а так как функция $F(X, \lambda)$ выпуклая по X на E_n при любом $\lambda \in \Lambda$, то

$$F(X^*, \lambda^*) = \min_{X \in E_n} F(X, \lambda^*).$$

Отсюда следует правое неравенство из (5.1).

Далее в силу II, III и определения $F(X, \lambda)$ имеем

$$F(X^*, \lambda^*) = \sum_{i \in R(X^*)} \lambda_{1i}^* f_i(X^*) = \varphi(X^*) \quad \sum_{i \in R(X^*)} \lambda_{1i}^* = \varphi(X^*). \quad (5.3)$$

Возьмем любое $\lambda \in \Lambda$. Поскольку $X^* \in \Omega$, то

$$\begin{aligned} F(X^*, \lambda) &= \sum_{i=0}^N \lambda_{1i} f_i(X^*) + \sum_{j=0}^{N_1} \lambda_{2j} h_j(X^*) \leq \\ &\leq \sum_{i=0}^N \lambda_{1i} f_i(X^*) \leq \varphi(X^*) \sum_{i=0}^N \lambda_{1i} = \varphi(X^*). \end{aligned} \quad (5.4)$$

Объединяя (5.3) и (5.4), получим левое неравенство из (5.1). Необходимость доказана.

Достаточность. Пусть $[X^*, \lambda^*] \in E_n \times \Lambda$ — седловая точка функции $F(X, \lambda)$ на $E_n \times \Lambda$. Докажем, что

$$\varphi(X^*) = \min_{X \in \Omega} \varphi(X).$$

Заметим прежде всего, что $X^* \in \Omega$. Действительно, допустим противное. Тогда для некоторого $j_0 \in [0 : N_1]$ будет $h_{j_0}(X^*) > 0$. Положим

$$\lambda' = (\lambda_{10}^*, \dots, \lambda_{1N}^*, \lambda_{20}^*, \dots, \lambda_{2, j_0-1}^*, \lambda_{2j_0}^* + a, \lambda_{2, j_0+1}^*, \dots, \lambda_{2N_1}^*) \in \Lambda,$$

где $a > 0$. В этом случае

$$F(X^*, \lambda') = F(X^*, \lambda^*) + ah_{j_0}(X^*) > F(X^*, \lambda^*),$$

что противоречит (5.1). Итак, $X^* \in \Omega$.

Заметим далее, что

$$\lambda_{2j}^* h_j(X^*) = 0 \quad \text{для всех } j \in [0 : N_1]. \quad (5.5)$$

В противном случае для некоторого j_0 будет

$$\lambda_{2j_0}^* h_{j_0}(X^*) < 0.$$

Положив

$$\lambda' = (\lambda_{10}^*, \dots, \lambda_{1N}^*, \lambda_{20}^*, \dots, \lambda_{2, j_0-1}^*, 0, \lambda_{2, j_0+1}^*, \dots, \lambda_{2N_1}^*) \in \Lambda,$$

получим

$$F(X^*, \lambda') = F(X^*, \lambda^*) - \lambda_{2i_0}^* h_{i_0}(X^*) > F(X^*, \lambda^*),$$

что противоречит (5.1). Тем самым доказана справедливость соотношений (5.5).

Покажем теперь, что

$$\lambda_{1i}^* = 0, \quad \text{если } i \notin R(X^*). \quad (5.6)$$

Допустим противное. Тогда найдется $\lambda_{1i_0}^* > 0$, $i_0 \notin R(X^*)$. Пусть $i_1 \in R(X^*)$. Положим

$$\lambda' = (\lambda'_{10}, \dots, \lambda'_{1N}, \lambda'_{20}, \dots, \lambda'_{2N_1}) \in \Lambda,$$

где

$$\lambda'_{1i} = \begin{cases} \lambda_{1i}^* & \text{для } i \neq i_0, i \neq i_1; \\ 0 & \text{для } i = i_0; \\ \lambda_{1i_1}^* + \lambda_{1i_0}^* & \text{для } i = i_1; \end{cases}$$

$$\lambda'_{2j} = \lambda_{2j}^*, \quad j \in [0 : N_1].$$

В этом случае будем иметь

$$F(X^*, \lambda') = F(X^*, \lambda^*) - \lambda_{1i_0}^* [f_{i_1}(X^*) - f_{i_0}(X^*)] > F(X^*, \lambda^*),$$

что противоречит (5.1). Итак, соотношения (5.6) также справедливы. Из (5.5) и (5.6) следует

$$F(X^*, \lambda^*) = \varphi(X^*). \quad (5.7)$$

Теперь уже нетрудно доказать равенство

$$\varphi(X^*) = \min_{X \in \Omega} \varphi(X).$$

Действительно, допустим противное. Тогда найдется $X' \in \Omega$, для которого $\varphi(X') < \varphi(X^*)$. Имеем

$$\begin{aligned} F(X', \lambda^*) &= \sum_{i=0}^N \lambda_{1i}^* f_i(X') + \sum_{j=0}^{N_1} \lambda_{2j}^* h_j(X') \leq \\ &\leq \sum_{i=0}^N \lambda_{1i}^* f_i(X') \leq \varphi(X') < \varphi(X^*). \end{aligned} \quad (5.8)$$

Учитывая (5.7) и (5.8), получим

$$F(X', \lambda^*) < F(X^*, \lambda^*),$$

что противоречит (5.1).

Теорема доказана полностью.

З а м е ч а н и е. При доказательстве достаточности мы не пользовались никакими специальными свойствами функций $f_i(X)$ и $h_j(X)$. Таким образом, наличие седловой точки $[X^*, \lambda^*]$ функции Лагранжа $F(X, \lambda)$ на $E_n \times \Lambda$ является достаточным условием равенства

$$\max_{i \in [0 : N]} f_i(X^*) = \min_{X \in \Omega} \max_{i \in [0 : N]} f_i(X),$$

где $\Omega = \{X | h_j(X) \leq 0, j \in [0 : N_1]\}$, для произвольных функций $f_i(X)$ и $h_j(X)$, заданных на E_n .

§ 6. Первый метод последовательных приближений

1. Результаты этого параграфа представляют собой естественное обобщение результатов § 6 главы III.

Пусть

$$\varphi(X) = \max_{i \in [0 : N]} f_i(X), \quad \varphi_1(X) = \max_{j \in [0 : N_1]} h_j(X),$$

$$\Omega = \{X \in E_n | h_j(X) \leq 0 \text{ для всех } j \in [0 : N_1]\}.$$

Как обычно, считаем, что функции $f_i(X)$, $i \in [0 : N]$, непрерывно дифференцируемы на некотором открытом множестве Ω' , содержащем Ω , функции $h_j(X)$, $j \in [0 : N_1]$, выпуклы и непрерывно дифференцируемы на E_n , и что выполнено условие Слейтера (2.2).

Положим при фиксированном $X \in \Omega$

$$R_\varepsilon(X) = \{i \in [0 : N] | \varphi(X) - f_i(X) \leq \varepsilon\}, \quad \varepsilon \geq 0;$$

$$Q_\mu(X) = \{j \in [0 : N_1] | -\mu \leq h_j(X) \leq 0\}, \quad \mu \geq 0.$$

Очевидно, $R_0(X) = R(X)$, и если $\varepsilon' \geq \varepsilon$, то

$$R_{\varepsilon'}(X) \supset R_\varepsilon(X). \quad (6.1)$$

Ясно также, что $Q_0(X) = Q(X)$, и если $\mu' \geq \mu$, то

$$Q_{\mu'}(X) \supset Q_\mu(X). \quad (6.2)$$

В § 6 главы III было указано явное представление для $R_\varepsilon(X)$ при различных ε . В частности, там отмечалось, что при некотором $a_1(X) > 0$ (конечном или бесконечном)

$$R_\varepsilon(X) = R(X), \quad \text{если } 0 \leq \varepsilon < a_1(X). \quad (6.3)$$

Аналогичное представление может быть получено и для $Q_\mu(X)$. Перенумеруем числа $h_j \stackrel{df}{=} h_j(X)$ ($X \in \Omega$ фиксировано) в порядке убывания:

$$h_{j_{01}} = \dots = h_{j_{0q_0}} > h_{j_{11}} = \dots = h_{j_{1q_1}} > \dots > h_{j_{r1}} = \dots = h_{j_{rq_r}}.$$

Очевидно, $0 \leq r \leq N_1$, причем $r \stackrel{df}{=} r(X)$ зависит от X . Обозначим

$$J_k(X) = \{j_{k1}, \dots, j_{kq_k}\}, \quad k \in [0:r];$$

$$b_k(X) = -h_j(X), \quad j \in J_k(X).$$

Ясно, что

$$0 \leq b_0(X) < b_1(X) < \dots < b_r(X).$$

По определению полагаем $b_{r+1}(X) = \infty$. Пусть

$$Q^s(X) = \bigcup_{k=0}^s J_k(X), \quad s \in [0:r].$$

Теперь можно записать представление для $Q_\mu(X)$:

$$Q_\mu(X) = \begin{cases} \emptyset, & \text{если } \mu \in [0, b_0(X)), \\ Q^s(X), & \text{если } b_s(X) \leq \mu < b_{s+1}(X), \quad s \in [0:r]. \end{cases}$$

Отсюда, в частности, следует, что

$$\left. \begin{aligned} Q_\mu(X) &= Q(X) \text{ для } 0 \leq \mu < b_0(X), \text{ если } b_0(X) > 0; \\ Q_\mu(X) &= Q(X) \text{ для } 0 \leq \mu < b_1(X), \text{ если } b_0(X) = 0. \end{aligned} \right\} \quad (6.4)$$

Лемма 6.1. Для любого фиксированного $X_0 \in \Omega$ существуют такие числа $\gamma_0 > 0$, $\delta_0 > 0$, $\gamma_1 > 0$, $\delta_1 > 0$, что

I. Для любых $\varepsilon \in [0, \gamma_0]$ и $X' \in \Omega$, $\|X' - X_0\| < \delta_0$, будет

$$R_\varepsilon(X') \subset R(X_0).$$

II. Для любых $\mu \in [0, \gamma_1]$ и $X' \in \Omega$, $\|X' - X_0\| < \delta_1$ будет

$$Q_\mu(X') \subset Q(X_0).$$

Доказательство этой леммы, аналогичное доказательству леммы 6.1 главы III, мы опускаем.

2. Введем обозначение

$$\tilde{L}_{\varepsilon\mu}(X) = \text{co } \tilde{H}_{\varepsilon\mu}(X),$$

где

$$X \in \Omega, \quad \tilde{H}_{\varepsilon\mu}(X) = H_{\varepsilon}(X) \cup H_{1\mu}(X),$$

$$H_{\varepsilon}(X) = \left\{ \frac{\partial f_i(X)}{\partial X} \mid i \in R_{\varepsilon}(X) \right\},$$

$$H_{1\mu}(X) = \left\{ \frac{\partial h_j(X)}{\partial X} \mid j \in Q_{\mu}(X) \right\}.$$

Ясно, что $\tilde{L}_{00}(X) = \tilde{L}(X)$ (см. § 3).

В силу (6.3) и (6.4) имеем

$$\tilde{L}_{\varepsilon\mu}(X) = \tilde{L}(X) \quad (6.5)$$

при $\varepsilon \in [0, a_1(X))$ и

$$\mu \in \begin{cases} [0, b_0(X)), & \text{если } b_0(X) > 0, \\ [0, b_1(X)), & \text{если } b_0(X) = 0. \end{cases}$$

Пусть

$$\tilde{d}_{\varepsilon\mu}(X) = \min_{Z \in \tilde{L}_{\varepsilon\mu}(X)} \|Z\|, \quad \tilde{d}(X) = \min_{Z \in \tilde{L}_{00}(X)} \|Z\|.$$

Определение 1. Точка $X^* \in \Omega$, для которой

$$0 \in \tilde{L}_{\varepsilon\mu}(X^*),$$

называется (ε, μ) -квазистационарной точкой функции $\varphi(X)$ на множестве Ω .

Определение 2. Пусть $0 \notin \tilde{L}_{\varepsilon\mu}(\bar{X})$. Вектор

$$\tilde{g}_{\varepsilon\mu}(\bar{X}) = -\tilde{Z}_{\varepsilon\mu}(\bar{X}) / \|\tilde{Z}_{\varepsilon\mu}(\bar{X})\|,$$

где $\|\tilde{Z}_{\varepsilon\mu}(\bar{X})\| = \min_{Z \in \tilde{L}_{\varepsilon\mu}(\bar{X})} \|Z\|$, называется направлением

(ε, μ) -квазинаискорейшего спуска функции $\varphi(X)$ в точке $\bar{X}$.

3. Приступим к описанию первого метода последовательных приближений для нахождения стационарных точек функции $\varphi(X)$ на Ω .

Зафиксируем три параметра $\varepsilon_0 > 0$, $\mu_0 > 0$ и $\rho_0 > 0$. Возьмем начальное приближение $X_0 \in \Omega$.

Предположим, что множество

$$M(X_0) = \{X \in \Omega \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено. В силу непрерывности $\varphi(X)$ и замкнутости Ω множество $M(X_0)$ является также замкнутым.

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $\tilde{d}(X_k) = 0$ (т. е. $0 \in \tilde{L}(X_k)$), то X_k — стационарная точка функции $\varphi(X)$ на Ω , и процесс на этом заканчивается.

Предположим, что $\tilde{d}(X_k) > 0$. Обозначим через ν_k наименьшее $\nu \in \{0, 1, 2, \dots\}$, при котором

$$\tilde{d}_{\varepsilon_{\nu} \mu_{\nu}}(X_k) \geq \rho_{\nu}, \quad (6.6)$$

где $\varepsilon_{\nu} = \varepsilon_0/2^{\nu}$, $\mu_{\nu} = \mu_0/2^{\nu}$, $\rho_{\nu} = \rho_0/2^{\nu}$. Заметим, что ν_k — конечное число. Это следует из того, что $\rho_{\nu} \rightarrow 0$ при $\nu \rightarrow \infty$, а левая часть неравенства (6.6) стремится в силу (6.5) к $\tilde{d}(X_k) > 0$.

В дальнейшем будем использовать обозначения

$$\hat{\varepsilon}_k = \varepsilon_{\nu_k}, \quad \hat{\mu}_k = \mu_{\nu_k}, \quad \hat{\rho}_k = \rho_{\nu_k}, \quad \tilde{d}_k = \tilde{d}_{\varepsilon_{\nu_k} \mu_{\nu_k}}(X_k).$$

Положим

$$\tilde{g}_k = \tilde{g}_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k),$$

где $\tilde{g}_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k)$ — направление $(\hat{\varepsilon}_k, \hat{\mu}_k)$ -квазинискорейшего спуска функции $\varphi(X)$ в точке X_k . Рассмотрим луч

$$X = X_k(\alpha) \stackrel{\text{df}}{=} X_k + \alpha \tilde{g}_k, \quad \alpha \geq 0.$$

Найдем $\alpha_k \in [0, \infty)$ такое, что

$$\varphi(X_k(\alpha_k)) = \min_{\substack{\alpha \in [0, \infty) \\ X_k(\alpha) \in \Omega}} \varphi(X_k(\alpha))$$

(минимум здесь достигается в силу ограниченности и замкнутости множества $M(X_0)$).

Теперь положим $X_{k+1} = X_k(\alpha_k)$. Очевидно, $X_{k+1} \in M(X_0)$.

Докажем прежде всего, что

$$\tilde{g}_k \in \Gamma(X_k). \quad (6.7)$$

Если $\varphi_1(X_k) < 0$, то это очевидно. Пусть $\varphi_1(X_k) = 0$.

Поскольку $\max_{Z \in \tilde{L}_{\varepsilon\mu}(X)} (Z, \tilde{g}_{\varepsilon\mu}(X)) = -\tilde{d}_{\varepsilon\mu}(X)$, то

$$\max_{j \in Q_{\mu}(X)} \left(\frac{\partial h_j(X)}{\partial X}, \tilde{g}_{\varepsilon\mu}(X) \right) \leq \max_{Z \in \tilde{L}_{\varepsilon\mu}(X)} (Z, \tilde{g}_{\varepsilon\mu}(X)) = -\tilde{d}_{\varepsilon\mu}(X). \quad (6.8)$$

Учитывая (6.8), получим для достаточно малых $\alpha > 0$

$$\begin{aligned}\varphi_1(X_k + \alpha \tilde{g}_k) &= \varphi_1(X_k) + \alpha \frac{\partial \varphi_1(X_k)}{\partial \tilde{g}_k} + o_k(\alpha) \leq \\ &\leq \alpha \max_{j \in Q_{\hat{\mu}_k}(X_k)} \left(\frac{\partial h_j(X_k)}{\partial X}, \tilde{g}_{\hat{\mu}_k}(X_k) \right) + o_k(\alpha) \leq \\ &\leq -\alpha \tilde{d}_k + o_k(\alpha) \leq -\frac{1}{2} \alpha \hat{\rho}_k < 0,\end{aligned}$$

откуда и следует (6.7).

Аналогично доказывается, что при малых $\alpha > 0$

$$\varphi(X_k + \alpha \tilde{g}_k) \leq \varphi(X_k) - \frac{1}{2} \alpha \hat{\rho}_k < \varphi(X_k). \quad (6.9)$$

Из (6.7) и (6.9) следует неравенство

$$\varphi(X_{k+1}) < \varphi(X_k).$$

Продолжая описанные выше построения, получим последовательность $\{X_k\}$ и связанные с нею последовательности

$$\{\hat{e}_k\}, \quad \{\hat{\mu}_k\}, \quad \{\hat{\rho}_k\}, \quad \{\tilde{g}_k\}, \quad \{\alpha_k\},$$

причем $X_k \in M(X_0)$, $k = 0, 1, 2, \dots$, и

$$\varphi(X_0) > \varphi(X_1) > \dots > \varphi(X_k) > \dots \quad (6.10)$$

Если последовательность $\{X_k\}$ конечна, то последний ее член по построению является стационарной точкой функции $\varphi(X)$ на Ω .

Пусть теперь последовательность $\{X_k\}$ бесконечна.

Лемма 6.2. *При $k \rightarrow \infty$ выполняются предельные соотношения*

$$\hat{e}_k \rightarrow 0, \quad \hat{\mu}_k \rightarrow 0, \quad \hat{\rho}_k \rightarrow 0. \quad (6.11)$$

Доказательство. Прежде всего заметим, что по построению соотношения (6.11) выполняются или не выполняются одновременно. Допустим, что (6.11) не выполняются. Тогда найдутся числа $\varepsilon^* > 0$, $\mu^* > 0$, $\rho^* > 0$ и подпоследовательность индексов $\{k_s\}$, $k_s \rightarrow \infty$, такие, что

$$\hat{e}_{k_s} \geq \varepsilon^*, \quad \hat{\mu}_{k_s} \geq \mu^*, \quad \hat{\rho}_{k_s} \geq \rho^*.$$

Поскольку

$$R_{\varepsilon^*}(X_{k_s}) \subset R_{\hat{e}_{k_s}}(X_{k_s}), \quad Q_{\mu^*}(X_{k_s}) \subset Q_{\hat{\mu}_{k_s}}(X_{k_s}),$$

то для $i \in R_{\varepsilon^*}(X_{k_s})$ получим

$$\begin{aligned} f_i(X_{k_s} + \alpha \tilde{g}_{k_s}) &= f_i(X_{k_s}) + \alpha \left(\frac{\partial f_i(X_{k_s})}{\partial X}, \tilde{g}_{k_s} \right) + o_{ik_s}(\alpha) \leq \\ &\leq f_i(X_{k_s}) - \alpha \tilde{d}_{k_s} + o_{ik_s}(\alpha) \leq f_i(X_{k_s}) - \alpha \rho^* + o_{ik_s}(\alpha). \end{aligned}$$

Аналогично для $j \in Q_{\mu^*}(X_{k_s})$ будем иметь

$$h_j(X_{k_s} + \alpha \tilde{g}_{k_s}) \leq h_j(X_{k_s}) - \alpha \rho^* + o_{jk_s}(\alpha).$$

Выберем теперь $\alpha_0 > 0$, не зависящее от i, j и k_s , такое, что для $i \in R_{\varepsilon^*}(X_{k_s})$, $j \in Q_{\mu^*}(X_{k_s})$, $\alpha \in [0, \alpha_0]$ будет

$$f_i(X_{k_s} + \alpha \tilde{g}_{k_s}) \leq \varphi(X_{k_s}) - \frac{1}{2} \alpha \rho^* \quad (6.12)$$

$$h_j(X_{k_s} + \alpha \tilde{g}_{k_s}) \leq -\frac{1}{2} \alpha \rho^*. \quad (6.13)$$

Пусть далее

$$C_1 = \max_{X \in M(X_0)} \max_{i \in [0 : N]} \left\| \frac{\partial f_i(X)}{\partial X} \right\|,$$

$$C_2 = \max_{X \in M(X_0)} \max_{j \in [0 : N]} \left\| \frac{\partial h_j(X)}{\partial X} \right\|.$$

Так как $X_{k_s} \in M(X_0)$, то для $i \notin R_{\varepsilon^*}(X_{k_s})$ и $j \notin Q_{\mu^*}(X_{k_s})$ будем иметь

$$\begin{aligned} f_i(X_{k_s} + \alpha \tilde{g}_{k_s}) &\leq f_i(X_{k_s}) + \alpha C_1 + o_{ik_s}(\alpha) \leq \\ &\leq \varphi(X_{k_s}) - \varepsilon^* + \alpha C_1 + o_{ik_s}(\alpha); \\ h_j(X_{k_s} + \alpha \tilde{g}_{k_s}) &\leq h_j(X_{k_s}) + \alpha C_2 + o_{jk_s}(\alpha) \leq \\ &\leq -\mu^* + \alpha C_2 + o_{jk_s}(\alpha). \end{aligned}$$

Очевидно, найдется такое $\alpha_1 > 0$, $0 < \alpha_1 \leq \alpha_0$, не зависящее от i, j и k_s , что для $i \notin R_{\varepsilon^*}(X_{k_s})$, $j \notin Q_{\mu^*}(X_{k_s})$ и $\alpha \in [0, \alpha_1]$ будет

$$f_i(X_{k_s} + \alpha \tilde{g}_{k_s}) \leq \varphi(X_{k_s}) - \frac{1}{2} \varepsilon^*; \quad (6.14)$$

$$h_j(X_{k_s} + \alpha \tilde{g}_{k_s}) \leq -\frac{1}{2} \mu^*. \quad (6.15)$$

Из (6.12)—(6.15) следует, в частности, что равномерно по k_s

$$\varphi(X_{k_s} + \alpha_1 \tilde{g}_{k_s}) \leq \varphi(X_{k_s}) - \beta_1; \quad (6.16)$$

$$\varphi_1(X_{k_s} + \alpha_1 \tilde{g}_{k_s}) \leq -\beta_2, \quad (6.17)$$

где

$$\beta_1 = \min \left\{ \frac{1}{2} \alpha_1 \rho^*, \frac{1}{2} \varepsilon^* \right\},$$

$$\beta_2 = \min \left\{ \frac{1}{2} \alpha_1 \rho^*, \frac{1}{2} \mu^* \right\}.$$

Теперь уже нетрудно получить противоречие. Из (6.10) следует, что последовательность $\{\varphi(X_k)\}$ имеет предел

$$\varphi(X_k) \xrightarrow{k \rightarrow \infty} \varphi^*, \quad (6.18)$$

причем для всех $k = 0, 1, 2, \dots$

$$\varphi(X_k) > \varphi^*. \quad (6.19)$$

На основании (6.17)

$$X_{k_s}(\alpha_1) \stackrel{df}{=} X_{k_s} + \alpha_1 \tilde{g}_{k_s} \in \Omega,$$

а на основании (6.16)

$$\varphi(X_{k_s}(\alpha_1)) \leq \varphi(X_{k_s}) - \beta_1.$$

При достаточно больших k_s будем иметь в силу (6.16) и (6.18)

$$\varphi(X_{k_s}(\alpha_1)) \leq \varphi^* - \frac{1}{2} \beta_1.$$

Но

$$\varphi(X_{k_s+1}) = \min_{\substack{\alpha \geq 0 \\ X_{k_s}(\alpha) \in \Omega}} \varphi(X_{k_s}(\alpha)) \leq \varphi(X_{k_s}(\alpha_1)) \leq \varphi^* - \frac{1}{2} \beta_1,$$

что противоречит (6.19).

Лемма доказана.

Отметим, что в приведенных выше рассуждениях могло быть

$$Q_{\mu^*}(X_{k_s}) = \emptyset.$$

4. Если построенная выше последовательность $\{X_k\}$ бесконечна, то она имеет хотя бы одну предельную точку, ибо $X_k \in M(X_0)$, $k = 0, 1, 2, \dots$, а, по предположению, множество $M(X_0)$ ограничено и замкнуто.

Теорема 6.1. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω .*

Доказательство. Пусть

$$X_{k_s} \xrightarrow{k_s \rightarrow \infty} X^*.$$

Ясно, что $X^* \in M(X_0)$ и, в частности, $X^* \in \Omega$. Утверждается, что X^* — стационарная точка функции $\varphi(X)$ на Ω , т. е. что $\tilde{d}(X^*) = 0$.

Допустим противное:

$$\tilde{d}(X^*) \stackrel{\text{def}}{=} \min_{Z \in \tilde{L}(X^*)} \|Z\| = d^* > 0.$$

По лемме 6.1 найдутся числа $\gamma^* > 0$, $\delta^* > 0$ и $K_1 > 0$ такие, что при $k_s \geq K_1$, $\varepsilon \in [0, \gamma^*]$, $\mu \in [0, \delta^*]$ будет

$$R_\varepsilon(X_{k_s}) \subset R(X^*), \quad Q_\mu(X_{k_s}) \subset Q(X^*).$$

Отсюда следует

$$\tilde{d}_{\varepsilon\mu}(X_{k_s}) \geq \min_{Z \in L'(X_{k_s})} \|Z\|,$$

где

$$L'(X_{k_s}) = \text{co } H'(X_{k_s}),$$

$$H'(X_{k_s}) = \left\{ \frac{\partial f_i(X_{k_s})}{\partial X} \mid i \in R(X^*) \right\} \cup \left\{ \frac{\partial h_j(X_{k_s})}{\partial X} \mid j \in Q(X^*) \right\}.$$

Очевидно, существует столь большое $K \geq K_1$, что при $k_s \geq K$ будет

$$\min_{Z \in L'(X_{k_s})} \|Z\| \geq \frac{1}{2} \tilde{d}(X^*) = \frac{d^*}{2}.$$

Итак, для $\varepsilon \in [0, \gamma^*]$, $\mu \in [0, \delta^*]$ и $k_s \geq K$ выполняется неравенство

$$\tilde{d}_{\varepsilon\mu}(X_{k_s}) \geq \frac{d^*}{2}. \quad (6.20)$$

Через v_1 обозначим наименьшее целое неотрицательное число, для которого

$$\frac{\varepsilon_0}{2^{v_1}} \leq \gamma^*, \quad \frac{\mu_0}{2^{v_1}} \leq \delta^*, \quad \frac{\rho_0}{2^{v_1}} \leq \frac{d^*}{2}. \quad (6.21)$$

Тогда для всех $k_s \geq K$ на основании (6.20) и (6.21) будет

$$\tilde{d}_{e_{v_1} \mu_{v_1}}(X_{k_s}) \geq \frac{d^*}{2} \geq \rho_{v_1}.$$

Отсюда следует

$$\hat{e}_{k_s} \geq e_{v_1}, \quad \hat{\mu}_{k_s} \geq \mu_{v_1}, \quad \hat{\rho}_{k_s} \geq \rho_{v_1},$$

что противоречит лемме 6.2.

Теорема доказана.

З а м е ч а н и е. На каждом шаге описанного выше метода последовательных приближений приходится находить $\hat{e}_k$, $\hat{\mu}_k$, $\hat{\rho}_k$, $\tilde{g}_k$ и α_k . По лемме 6.2 $\hat{e}_k \rightarrow 0$, $\hat{\mu}_k \rightarrow 0$, $\hat{\rho}_k \rightarrow 0$, поэтому число испытаний (сравнений $\tilde{d}_{e_v \mu_v}(X_k)$ с ρ_v , $v = 0, 1, 2, \dots$, — см. (6.6)) для выбора $\hat{e}_k$, $\hat{\mu}_k$ и $\hat{\rho}_k$ возрастает и стремится к бесконечности при $k \rightarrow \infty$.

Как уже отмечалось в главе III, при реальных расчетах можно использовать следующее соображение: испытания начинать не с e_0 , μ_0 , ρ_0 , а с $\hat{e}_{k-1}$, $\hat{\mu}_{k-1}$, $\hat{\rho}_{k-1}$, найденных на предыдущем шаге. Если

$$\tilde{d}_{\hat{e}_{k-1} \hat{\mu}_{k-1}}(X_k) < \hat{\rho}_{k-1},$$

то перебираем последовательности

$$\hat{e}_{k-1}/2^v, \quad \hat{\mu}_{k-1}/2^v, \quad \hat{\rho}_{k-1}/2^v, \quad v = 1, 2, \dots,$$

пока не дойдем до $\hat{e}_k$, $\hat{\mu}_k$ и $\hat{\rho}_k$.

Если же

$$\tilde{d}_{\hat{e}_{k-1} \hat{\mu}_{k-1}}(X_k) \geq \hat{\rho}_{k-1},$$

то ищем наибольшее $v \in [0 : \log_2 \frac{e_0}{\hat{e}_{k-1}}]$, для которого

$$\tilde{d}_{e_{kv} \mu_{kv}}(X_k) \geq \rho_{kv},$$

где

$$e_{kv} = 2^v \hat{e}_{k-1}, \quad \mu_{kv} = 2^v \hat{\mu}_{k-1}, \quad \rho_{kv} = 2^v \hat{\rho}_{k-1}.$$

Соответствующие наибольшему v величины e_{kv} , μ_{kv} , ρ_{kv} и являются требуемыми $\hat{e}_k$, $\hat{\mu}_k$, $\hat{\rho}_k$.

5*. Опишем одну модификацию первого метода последовательных приближений,

Положим $\bar{e}_{-1} = e_0$, $\bar{\mu}_{-1} = \mu_0$, $\bar{\rho}_{-1} = \rho_0$ и возьмем начальное приближение $X'_0 \in \Omega$. Будем считать, что множество $M(X'_0) = \{X \in \Omega \mid \varphi(X) \leq \varphi(X'_0)\}$ ограничено.

Пусть уже найдены $X'_k \in M(X'_0)$, $\bar{e}_{k-1}$, $\bar{\mu}_{k-1}$, $\bar{\rho}_{k-1}$, и пусть $\tilde{d}(X'_k) > 0$. Для построения X'_{k+1} найдем наименьшее $v \in \{0, 1, 2, \dots\}$, при котором

$$\tilde{d}_{\bar{e}'_v \bar{\mu}'_v}(X'_k) \geq \rho'_v,$$

где $\bar{e}'_v = \bar{e}_{k-1}/2^v$, $\bar{\mu}'_v = \bar{\mu}_{k-1}/2^v$, $\rho'_v = \bar{\rho}_{k-1}/2^v$.

Обозначим его через v'_k . Положим

$$\bar{e}_k = \bar{e}'_{v'_k}, \quad \bar{\mu}_k = \bar{\mu}'_{v'_k}, \quad \bar{\rho}_k = \rho'_{v'_k}.$$

Далее поступаем, как в основном методе: находим $g'_k = \tilde{g}_{\bar{e}_k \bar{\mu}_k}(X'_k)$ и в качестве X'_{k+1} берем точку минимума функции $\varphi(X)$ на отрезке $X = X'_k(\alpha) \stackrel{\text{df}}{=} X'_k + \alpha g'_k$, $\alpha \geq 0$, $X'_k(\alpha) \in \Omega$.

Ясно, что $X'_{k+1} \in M(X'_0)$, $\bar{e}_k \leq \bar{e}_{k-1}$, $\bar{\mu}_k \leq \bar{\mu}_{k-1}$, $\bar{\rho}_k \leq \bar{\rho}_{k-1}$ и что $\varphi(X'_{k+1}) < \varphi(X'_k)$.

Если построенная последовательность $\{X'_k\}$ конечна, то последний ее элемент является стационарной точкой функций $\varphi(X)$ на Ω .

В противном случае справедлива

Теорема 6.2. Пусть $\{X'_{k_p}\}$ — такая подпоследовательность последовательности $\{X'_k\}$, что

$$\bar{e}_{k_p} < \bar{e}_{k_p-1}, \quad \bar{\mu}_{k_p} < \bar{\mu}_{k_p-1}.$$

Если множество $M(X'_0)$ ограничено, то любая предельная точка подпоследовательности $\{X'_{k_p}\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω .

Доказательство этой теоремы аналогично доказательству теоремы 6.2 главы III. Наличие ограничений необходимо учитывать так же, как это делалось при доказательстве теоремы 6.1.

6*. В описанном выше первом методе последовательных приближений могут быть сделаны следующие

изменения: вместо $Q_\mu(X)$ можно рассматривать множество

$$Q'_\mu(X) = \{j \in [0 : N_1] \mid -\mu \leq t_j(X) h_j(X) \leq 0\},$$

а вместо $H_{1\mu}(X)$ —

$$H'_{1\mu}(X) = \left\{ b_0 t_j(X) \frac{\partial h_j(X)}{\partial X} \mid j \in Q'_\mu(X) \right\}.$$

Здесь $t_j(X) = \left\| \frac{\partial h_j(X)}{\partial X} \right\|^{-1}$, а b_0 — произвольная положительная константа.

Множество $\tilde{L}'_{e\mu}(X)$ будет иметь вид

$$\tilde{L}'_{e\mu}(X) = \text{co}(H_e(X) \cup H'_{1\mu}(X)),$$

причем условие $0 \in \tilde{L}'_{00}(X)$ по-прежнему определяет стационарные точки функции $\varphi(X)$ на множестве Ω . Формальное описание метода в этом случае сохраняется. При доказательстве же сходимости нужно дополнительно использовать следующие факты (см. замечание к лемме 2.5 и лемму 2.6):

I. Существуют такие $a > 0$ и $\bar{\mu} > 0$, что для всех $\mu \in [0, \bar{\mu}]$, $X \in (\partial_\mu \Omega) \cap M(X_0)$ и $j \in Q'_\mu(X)$ будет

$$\left\| \frac{\partial h_j(X)}{\partial X} \right\| \geq a.$$

II. Для каждого $\mu > 0$ найдется такое $c > 0$, что для всех $X \in M(X_0)$ и $j \notin Q'_\mu(X)$ будет

$$h_j(X) \leq -c.$$

Параметр метода μ_0 следует брать не большим, чем $\bar{\mu}$. Подробное доказательство сходимости предоставляется читателю.

7*. Приведем соображение, по которому целесообразно использовать $Q'_\mu(X)$. Предположим, что все ограничения, задающие множество Ω , линейны:

$$\Omega = \{Z \in E_n \mid h_j(Z) \stackrel{\text{def}}{=} (A_j, Z) + b_j \leq 0 \text{ для всех } j \in [0 : N_1]\}.$$

Покажем, что для любой точки $X \in \Omega$ индекс j принадлежит $Q'_\mu(X)$ тогда и только тогда, когда расстояние точки X до плоскости $h_j(Z) = 0$ не больше μ (рис. 32).

Действительно, расстояние ρ_j точки X до плоскости $(A_j, Z) + b_j = 0$ с точностью до знака есть

$$\beta_j(X) = \|A_j\|^{-1} ((A_j, X) + b_j).$$

Учитывая, что $X \in \Omega$, т. е. $\beta_j(X) \leq 0$, будем иметь

$$\rho_j = -\beta_j(X).$$

Значит, расстояние точки X до плоскости $h_j(Z) = 0$ не больше μ тогда и только тогда, когда

$$0 \geq \|A_j\|^{-1} ((A_j, X) + b_j) \geq -\mu,$$

или, что то же самое,

$$-\mu \leq t_j(X) h_j(X) \leq 0, \quad (6.22)$$

$$\text{где } t_j(X) \stackrel{\text{дф}}{=} \left\| \frac{\partial h_j(X)}{\partial X} \right\|^{-1} = \|A_j\|^{-1}.$$

С другой стороны, неравенства (6.22) есть по определению условие принадлежности индекса j множеству $Q'_\mu(X)$.

Рис. 32.

Утверждение доказано.

Что касается коэффициента b_0 в определении множества $H'_\mu(X)$, то следует напомнить (см. § 4), что при $\varepsilon = 0$, $\mu = 0$ и больших b_0 вектор $\tilde{g}_{00}(\bar{X})$ близок к направлению наискорейшего спуска функции $\varphi(X)$ в точке $\bar{X}$ на множестве Ω (если, разумеется, $\tilde{d}(\bar{X}) > 0$).

§ 7. Нахождение (ε, μ) -квазистационарных точек. Второй метод последовательных приближений

1. Определение (ε, μ) -квазистационарных точек было дано в предыдущем параграфе. Опишем теперь метод последовательных приближений для нахождения (ε, μ) -квазистационарных точек функции $\varphi(X)$ на множестве Ω .

Зафиксируем $\varepsilon > 0$ и $\mu > 0$. Возьмем начальное приближение $X_0 \in \Omega$. Будем предполагать, что множество

$$M(X_0) = \{X \in \Omega \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено (кроме того, $M(X_0)$ — замкнутое множество).

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $\tilde{d}_{e\mu}(X_k) = 0$ (т. е. $0 \in \tilde{L}_{e\mu}(X_k)$), то X_k — (e, μ) -квазистационарная точка функции $\varphi(X)$ на Ω , и процесс на этом заканчивается.

Если же $\tilde{d}_{e\mu}(X_k) > 0$, то строим следующее приближение X_{k+1} . Положим

$$g_k = \tilde{g}_{e\mu}(X_k),$$

где $\tilde{g}_{e\mu}(X_k)$ — направление (e, μ) -квазинаискорейшего спуска функции $\varphi(X)$ в точке X_k , и рассмотрим луч $X = X_k(\alpha) \stackrel{\text{дф}}{=} X_k + \alpha g_k$, $\alpha \geq 0$. Найдем $\alpha_k \in [0, \infty)$ такое, что

$$\varphi(X_k(\alpha_k)) = \min_{\substack{\alpha \geq 0 \\ X_k(\alpha) \in \Omega}} \varphi(X_k(\alpha))$$

(инфимум здесь достигается, ибо $M(X_0)$ замкнуто и ограничено). В качестве $(k+1)$ -го приближения X_{k+1} возьмем $X_k(\alpha_k)$:

$$X_{k+1} = X_k(\alpha_k).$$

Ясно, что $X_{k+1} \in M(X_0)$. Как и в § 6, доказывается неравенство

$$\varphi(X_{k+1}) < \varphi(X_k). \quad (7.1)$$

Теперь можно считать построенной последовательность $\{X_k\}$, причем все ее элементы принадлежат $M(X_0)$. Если эта последовательность конечна, то последний ее элемент по построению является (e, μ) -квазистационарной точкой функции $\varphi(X)$ на Ω . В противном случае справедлива

Теорема 7.1. *Любая предельная точка последовательности $\{X_k\}$ является (e, μ) -квазистационарной точкой функции $\varphi(X)$ на множестве Ω .*

Предварительно докажем одну лемму.

Лемма 7.1. *Имеет место предельное соотношение*

$$\lim_{k \rightarrow \infty} \tilde{d}_{e\mu}(X_k) = 0. \quad (7.2)$$

Доказательство. Допустим противное. Тогда найдется подпоследовательность $\{X_{k_s}\}$ и число $b > 0$ такие, что

$$\tilde{d}_{e\mu}(X_{k_s}) \geq b.$$

Поскольку для $j \in Q_\mu(X_{k_s})$ и малых α равномерно по k_s выполняется неравенство

$$h_j(X_{k_s} + \alpha g_{k_s}) \leq h_j(X_{k_s}) - \frac{1}{2} \alpha b,$$

а для $j \notin Q_\mu(X_{k_s})$ и малых α равномерно по k_s

$$h_j(X_{k_s} + \alpha g_{k_s}) \leq -\frac{1}{2} \mu,$$

то найдется $\alpha_1 > 0$ такое, что для всех $\alpha \in [0, \alpha_1]$ и всех k_s будет

$$\varphi_1(X_{k_s} + \alpha g_{k_s}) \leq 0,$$

т. е. $X_{k_s} + \alpha g_{k_s} \in \Omega$.

Аналогично показывается, что найдется столь малое $\alpha' \in (0, \alpha_1)$, для которого

$$\varphi(X_{k_s} + \alpha' g_{k_s}) \leq \varphi(X_{k_s}) - \beta, \quad (7.3)$$

где $\beta = \min \left\{ \frac{1}{2} \alpha' b, \frac{1}{2} \varepsilon \right\}$.

Значит, точка $X_{k_s}(\alpha') = X_{k_s} + \alpha' g_{k_s}$ принадлежит Ω , и выполняется неравенство (7.3).

Теперь уже нетрудно получить противоречие. Из (7.1) следует

$$\varphi(X_k) \xrightarrow{k \rightarrow \infty} \varphi^* \geq \min_{X \in M(X_0)} \varphi(X),$$

причем для всех $k = 0, 1, 2, \dots$

$$\varphi(X_k) > \varphi^*. \quad (7.4)$$

При достаточно больших k_s будет

$$\varphi(X_{k_s}) \leq \varphi^* + \frac{1}{2} \beta,$$

так что в силу (7.3)

$$\varphi(X_{k_s+1}) \stackrel{df}{=} \min_{\substack{\alpha \geq 0 \\ X_{k_s}(\alpha) \in \Omega}} \varphi(X_{k_s}(\alpha)) \leq \varphi(X_{k_s}(\alpha')) \leq \varphi^* - \frac{1}{2} \beta,$$

что противоречит (7.4).

Лемма доказана.

Доказательство теоремы. Пусть

$$X_{k_s} \xrightarrow{k_s \rightarrow \infty} X^*.$$

Ясно, что $X^* \in M(X_0)$, в частности, $X^* \in \Omega$. Требуется доказать, что $\tilde{d}_{\varepsilon\mu}(X^*) = 0$. Допустим противное:

$$\tilde{d}_{\varepsilon\mu}(X^*) \stackrel{\text{df}}{=} \min_{Z \in \tilde{L}_{\varepsilon\mu}(X^*)} \|Z\| = d^* > 0.$$

В силу леммы 7.1 главы III найдется такое $\delta > 0$, что для всех $X \in \Omega$, $\|X - X^*\| < \delta$, будет

$$R_\varepsilon(X) \subset R_\varepsilon(X^*), \quad Q_\mu(X) \subset Q_\mu(X^*).$$

Отсюда следует, что при больших k_s

$$\tilde{d}_{\varepsilon\mu}(X_{k_s}) \geq \frac{d^*}{2}$$

(см. доказательство леммы 7.2 главы III). Однако это противоречит (7.2).

Теорема доказана.

2. На основе метода для нахождения (ε, μ) -квазистационарных точек можно разработать метод для нахождения стационарных точек функции $\varphi(X)$ на множестве Ω .

Возьмем произвольные $\varepsilon_0 > 0$, $\mu_0 > 0$, $\rho_0 > 0$ и начальное приближение $X_0 \in \Omega$. Предполагаем, что множество $M(X_0)$ ограничено.

Если $\tilde{d}(X_0) = 0$, то X_0 — стационарная точка функции $\varphi(X)$ на Ω . Если же $\tilde{d}(X_0) > 0$, то с помощью описанного выше метода для нахождения (ε, μ) -квазистационарных точек в конечное число шагов получим точку $X_1 \in M(X_0)$, для которой

$$\tilde{d}_{\varepsilon_0\mu_0}(X_1) \leq \rho_0.$$

При этом $\varphi(X_1) < \varphi(X_0)$. Теперь, имея X_1 , полагаем

$$\varepsilon_1 = \frac{1}{2} \varepsilon_0, \quad \mu_1 = \frac{1}{2} \mu_0, \quad \rho_1 = \frac{1}{2} \rho_0$$

и повторяем описанную выше процедуру.

В результате построим последовательность $\{X_k\}$ такую, что

$$\tilde{d}_{\varepsilon_k \mu_k}(X_{k+1}) \leq \rho_k, \quad (7.5)$$

где $\varepsilon_k = \varepsilon_0/2^k$, $\mu_k = \mu_0/2^k$, $\rho_k = \rho_0/2^k$, $k = 0, 1, 2, \dots$

Если последовательность $\{X_k\}$ конечна, то последний ее элемент является по построению стационарной точкой функции $\varphi(X)$ на Ω . В противном случае справедлива

Теорема 7.2. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω .*

Доказательство. Пусть $X_{k_s} \xrightarrow{k_s \rightarrow \infty} X^*$. Очевидно, $X^* \in \Omega$. Докажем, что $\tilde{d}(X^*) = 0$.

Допустим противное: $\tilde{d}(X^*) > 0$. В силу (6.3) и (6.4) найдутся такие $\varepsilon > 0$ и $\mu > 0$, что

$$R_\varepsilon(X^*) = R(X^*), \quad Q_\mu(X^*) = Q(X^*).$$

Значит,

$$\tilde{d}_{\varepsilon \mu}(X^*) = \tilde{d}(X^*) > 0.$$

При достаточно больших k_s будет

$$\tilde{d}_{\varepsilon_{k_s} \mu_{k_s}}(X_{k_s+1}) \geq \frac{1}{2} \tilde{d}(X^*),$$

что противоречит (7.5).

Теорема доказана.

Замечание. Напомним, что первый и второй методы последовательных приближений могут быть использованы только при выполнении условия Слейтера (2.2), ибо только в этом случае равенство $\tilde{d}(X^*) = 0$ эквивалентно соотношению (3.1), определяющему стационарные точки.

§ 8. Метод наискорейшего спуска. Случай линейных ограничений

1. В §§ 6 и 7 для построения методов последовательных приближений использовалось необходимое условие минимакса в форме (3.2). Ниже приводятся алгоритмы, использующие необходимое условие минимакса в форме (3.1).

Пусть точка $X_0 \in \Omega$ не является стационарной. Тогда

$$d(X_0) \stackrel{df}{=} \min_{\substack{Z \in L(X_0) \\ V \in \Gamma^+(X_0)}} \|Z - V\| = \|Z(X_0) - V(X_0)\| > 0.$$

Направление

$$g(X_0) = \frac{V(X_0) - Z(X_0)}{\|V(X_0) - Z(X_0)\|}, \quad \|g(X_0)\| = 1,$$

является направлением наискорейшего спуска функции $\varphi(X)$ в точке X_0 на множестве Ω .

В настоящем параграфе рассмотрен случай, когда все функции $h_j(X)$ линейны:

$$h_j(X) = (A_j, X) + b_j, \quad j = [0 : N_1].$$

Будем предполагать, что $\|A_j\| \neq 0$. Без ограничения общности можно считать, что $\|A_j\| = 1$, $j \in [0 : N_1]$. Кроме того, предполагается, что множество

$$\Omega = \{X \in E_n \mid (A_j, X) + b_j \leq 0 \text{ для всех } j \in [0 : N_1]\}$$

не пусто. В этом случае в силу замечания к теореме 2.1 для любой точки $X_0 \in \Omega$ будет

$$\Gamma^+(X_0) = \left\{ V = - \sum_{j \in Q(X_0)} \alpha_j A_j \mid \alpha_j \geq 0 \right\}.$$

Введем в рассмотрение множества

$$L_\varepsilon(X_0) = \text{co } H_\varepsilon(X_0), \quad H_\varepsilon(X_0) = \left\{ \frac{\partial f_i(X_0)}{\partial X} \mid i \in R_\varepsilon(X_0) \right\};$$

$$R_\varepsilon(X_0) = \{i \in [0 : N] \mid \varphi(X_0) - f_i(X_0) \leq \varepsilon\}, \quad \varepsilon \geq 0;$$

$$\Gamma_\mu^+(X_0) = \left\{ V = - \sum_{j \in Q_\mu(X_0)} \alpha_j A_j \mid \alpha_j \geq 0 \right\};$$

$$Q_\mu(X_0) = \{j \in [0 : N_1] \mid -\mu \leq h_j(X_0) \leq 0\}, \quad \mu \geq 0.$$

Если $Q_\mu(X_0) = \emptyset$, то полагаем $\Gamma_\mu^+(X_0) = \{0\}$.

Определение 1. Пусть $\varepsilon \geq 0$, $\mu \geq 0$. Точка $X_0 \in \Omega$, удовлетворяющая условию

$$d_{\varepsilon\mu}(X_0) \stackrel{df}{=} \min_{\substack{Z \in L_\varepsilon(X_0) \\ V \in \Gamma_\mu^+(X_0)}} \|Z - V\| = 0,$$

называется (ε, μ) -стационарной точкой функции $\varphi(X)$ на множестве Ω ,

Определение 2. Пусть $X_0 \in \Omega$. Если

$$d_{\varepsilon\mu}(X_0) = \|Z_{\varepsilon\mu}(X_0) - V_{\varepsilon\mu}(X_0)\| > 0,$$

где $Z_{\varepsilon\mu}(X_0) \in L_\varepsilon(X_0)$, $V_{\varepsilon\mu}(X_0) \in \Gamma_\mu^+(X_0)$, то направление

$$g_{\varepsilon\mu}(X_0) = \frac{V_{\varepsilon\mu}(X_0) - Z_{\varepsilon\mu}(X_0)}{\|Z_{\varepsilon\mu}(X_0) - Z_{\varepsilon\mu}(X_0)\|} = \frac{V_{\varepsilon\mu}(X_0) - Z_{\varepsilon\mu}(X_0)}{d_{\varepsilon\mu}(X_0)}$$

называется направлением (ε, μ) -наискорейшего спуска функции $\varphi(X)$ в точке X_0 на множестве Ω .

Ясно, что $\|g_{\varepsilon\mu}(X_0)\| = 1$. Как и в главе IV, нетрудно установить, что направление $g_{\varepsilon\mu}(X_0)$ единственно, хотя пара точек $Z_{\varepsilon\mu}(X_0)$ и $V_{\varepsilon\mu}(X_0)$ может оказаться и неединственной.

Из (6.3) и (6.4) следует, что

$$L_\varepsilon(X_0) = L(X_0) \quad \text{при } \varepsilon \in [0, a_1(X_0)],$$

$$\Gamma_\mu^+(X_0) = \Gamma^+(X_0) \quad \text{при } \mu \in \begin{cases} [0, b_0(X_0)], & \text{если } b_0(X_0) > 0, \\ [0, b_1(X_0)], & \text{если } b_0(X_0) = 0. \end{cases}$$

Поэтому для таких ε и μ будет

$$d_{\varepsilon\mu}(X_0) = d(X_0). \quad (8.1)$$

Лемма 8.1. Если $d_{\varepsilon\mu}(X_0) > 0$, то

$$\left(\frac{\partial f_i(X_0)}{\partial X}, g_{\varepsilon\mu}(X_0) \right) \leq -d_{\varepsilon\mu}(X_0), \quad i \in R_\varepsilon(X_0);$$

$$\left(\frac{\partial h_j(X_0)}{\partial X}, g_{\varepsilon\mu}(X_0) \right) = (A_j, g_{\varepsilon\mu}(X_0)) \leq 0, \quad j \in Q_\mu(X_0).$$

Доказательство немедленно следует из определения $d_{\varepsilon\mu}(X_0)$ и $g_{\varepsilon\mu}(X_0)$ (см. доказательство теоремы 2.1 Приложения II).

2. Переходим к описанию метода наискорейшего спуска для нахождения стационарных точек функции $\varphi(X)$ на Ω .

Прежде всего зафиксируем три параметра $\varepsilon_0 > 0$, $\mu_0 > 0$, $\rho_0 > 0$. Возьмем начальное приближение $X_0 \in \Omega$. Будем предполагать, что множество

$$M(X_0) = \{X \in \Omega \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено.

Пусть уже найдено k -е приближение $X_k \in M(X_0)$. Если $d(X_k) = 0$, то X_k — стационарная точка, и процесс на этом заканчивается.

Пусть $d(X_k) > 0$. Тогда найдем наименьшее $v \in \{0, 1, 2, \dots\}$, при котором

$$d_{\varepsilon_v \mu_v}(X_k) \geq \rho_v, \quad (8.2)$$

где $\varepsilon_v = \varepsilon_0/2^v$, $\mu_v = \mu_0/2^v$, $\rho_v = \rho_0/2^v$. Обозначим его через v_k . Очевидно, v_k — конечное число, ибо $\rho_v \rightarrow 0$ при $v \rightarrow \infty$, а левая часть неравенства (8.2) в силу (8.1) стремится к $d(X_k) > 0$.

Введем обозначения

$$\begin{aligned} \hat{\varepsilon}_k &= \varepsilon_{v_k}, \quad \hat{\mu}_k = \mu_{v_k}, \quad \hat{\rho}_k = \rho_{v_k}, \\ d_k &= d_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k), \quad g_k = g_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k), \end{aligned}$$

где $g_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k)$ — направление $(\hat{\varepsilon}_k, \hat{\mu}_k)$ -наискорейшего спуска функции $\varphi(X)$ в точке X_k .

Покажем, что

$$X_k + \alpha g_k \in \Omega \quad \text{при всех } \alpha \in [0, \hat{\mu}_k]. \quad (8.3)$$

Действительно, для $j \in Q_{\hat{\mu}_k}(X_k)$ и всех $\alpha \geq 0$ имеем в силу леммы 8.1

$$h_j(X_k + \alpha g_k) = h_j(X_k) + \alpha(A_j, g_k) \leq 0.$$

Пусть $j \notin Q_{\hat{\mu}_k}(X_k)$. Учитывая, что $\|A_j\| = 1$, получим

$$h_j(X_k + \alpha g_k) = h_j(X_k) + \alpha(A_j, g_k) \leq -\hat{\mu}_k + \alpha.$$

Итак, для $\alpha \in [0, \hat{\mu}_k]$ и всех $j \in [0 : N_1]$ будет

$$h_j(X_k + \alpha g_k) \leq 0,$$

что равносильно (8.3). Отсюда, в частности, следует, что g_k — допустимое направление: $g_k \in \Gamma(X_k)$.

Рассмотрим теперь луч

$$X = X_k(\alpha) \stackrel{\text{df}}{=} X_k + \alpha g_k, \quad \alpha \geq 0,$$

и найдем $\alpha_k \in [0, \infty)$, при котором

$$\varphi(X_k(\alpha_k)) = \min_{\substack{\alpha \geq 0 \\ X_k(\alpha) \in \Omega}} \varphi(X_k(\alpha)).$$

В качестве $(k+1)$ -го приближения возьмем $X_k(a_k)$:

$$X_{k+1} = X_k(a_k).$$

Ясно, что

$$X_{k+1} \in M(X_0), \quad \varphi(X_{k+1}) < \varphi(X_k).$$

Далее продолжаем аналогично. В результате получим последовательность $\{X_k\}$. Если эта последовательность конечна, то последний ее элемент по построению является стационарной точкой функции $\varphi(X)$ на Ω .

Рассмотрим случай, когда последовательность $\{X_k\}$ бесконечна.

Лемма 8.2. *При $k \rightarrow \infty$ выполняются предельные соотношения*

$$\hat{\varepsilon}_k \rightarrow 0, \quad \hat{\mu}_k \rightarrow 0, \quad \hat{\rho}_k \rightarrow 0. \quad (8.4)$$

Доказательство. Допустим противное. Тогда найдутся числа $\varepsilon^* > 0$, $\mu^* > 0$, $\rho^* > 0$ и подпоследовательность индексов $\{k_s\}$, $k_s \xrightarrow{s \rightarrow \infty} \infty$, такие, что

$$\hat{\varepsilon}_{k_s} \geq \varepsilon^*, \quad \hat{\mu}_{k_s} \geq \mu^*, \quad \hat{\rho}_{k_s} \geq \rho^*.$$

В силу (8.3) при $\alpha \in [0, \mu^*]$ будем иметь

$$X_{k_s} + \alpha g_{k_s} \in \Omega.$$

Учитывая лемму 8.1, нетрудно показать (см. доказательство леммы 6.2), что найдется α' , $0 < \alpha' \leq \mu^*$, при котором равномерно по k_s будет

$$\varphi(X_{k_s} + \alpha' g_{k_s}) \leq \varphi(X_{k_s}) - \beta, \quad (8.5)$$

где $\beta > 0$. Неравенство (8.5) обычным путем приводит нас к противоречию.

Лемма доказана.

Теорема 8.1. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω .*

Доказательство. Существование предельных точек последовательности $\{X_k\}$ следует из ее ограниченности. Пусть $X_{k_s} \xrightarrow{k_s \rightarrow \infty} X^*$. Ясно, что $X^* \in \Omega$. Утверждается, что $d(X^*) = 0$.

Допустим противное:

$$d(X^*) = \min_{\substack{Z \in L(X^*) \\ V \in \Gamma^+(X^*)}} \|Z - V\| = d^* > 0.$$

По лемме 6.1 найдутся числа $\gamma^* > 0$, $\delta^* > 0$, $K_1 > 0$ такие, что при $k_s \geq K_1$, $\varepsilon \in [0, \gamma^*]$, $\mu \in [0, \delta^*]$ будет

$$R_\varepsilon(X_{k_s}) \subset R(X^*), \quad Q_\mu(X_{k_s}) \subset Q(X^*).$$

Поэтому

$$d_{\varepsilon\mu}(X_{k_s}) \geq \min_{\substack{Z \in L'(X_{k_s}) \\ V \in \Gamma'^+(X_{k_s})}} \|Z - V\|,$$

где

$$L'(X_{k_s}) = \text{co } H'(X_{k_s}), \quad H'(X_{k_s}) = \left\{ \frac{\partial f_i(X_{k_s})}{\partial X} \mid i \in R(X^*) \right\},$$

$$\Gamma'^+(X_{k_s}) = \left\{ V = - \sum_{i \in Q(X^*)} \alpha_i A_i \mid \alpha_i \geq 0 \right\}.$$

Очевидно,

$$\Gamma'^+(X_{k_s}) = \Gamma^+(X^*).$$

Нетрудно показать, что при достаточно больших $k_s \geq K \geq K_1$ будет

$$\min_{\substack{Z \in L'(X_{k_s}) \\ V \in \Gamma'^+(X_{k_s})}} \|Z - V\| \geq \frac{1}{2} d^*.$$

Значит,

$$d_{\varepsilon\mu}(X_{k_s}) \geq \frac{1}{2} d^* \quad (8.6)$$

для $\varepsilon \in [0, \gamma^*]$, $\mu \in [0, \delta^*]$ и $k_s \geq K$.

Обозначим через v_1 наименьшее неотрицательное целое число, при котором

$$\frac{\varepsilon_0}{2^{v_1}} \leq \gamma^*, \quad \frac{\mu_0}{2^{v_1}} \leq \delta^*, \quad \frac{\rho_0}{2^{v_1}} \leq \frac{d^*}{2}.$$

В силу (8.6) будем иметь для $k_s \geq K$

$$d_{\varepsilon_{v_1}\mu_{v_1}}(X_{k_s}) \geq \frac{d^*}{2} \geq \rho_{v_1}.$$

Следовательно,

$$\hat{\varepsilon}_{k_s} \geq \varepsilon_{v_1}, \quad \hat{\mu}_{k_s} \geq \mu_{v_1}, \quad \hat{\rho}_{k_s} \geq \rho_{v_1},$$

что противоречит (8.4).

Теорема доказана.

3. Аналогично тому, как это делалось в § 7, может быть описан метод последовательных приближений для нахождения (ε, μ) -стационарных точек и на его основе развит метод для определения стационарных точек функции $\varphi(X)$ на множестве Ω . Напомним, что

$$\Omega = \{X \in E_n \mid (A_j, X) + b_j \leq 0 \text{ для всех } j \in [0 : N_1]\}.$$

Подробное описание соответствующих алгоритмов и доказательство их сходимости предоставляется читателю.

§ 9. Случай нелинейных ограничений. Способы корректировки направлений

1. Пусть

$$\Omega = \{X \in E_n \mid h_j(X) \leq 0 \text{ для всех } j \in [0 : N_1]\},$$

где $h_j(X)$ — выпуклые и непрерывно дифференцируемые функции, и пусть выполнено условие Слейтера, т. е. существует точка $\bar{X} \in \Omega$, для которой

$$\max_{j \in [0 : N_1]} h_j(\bar{X}) < 0. \quad (9.1)$$

Для любой точки $X_0 \in \Omega$ положим

$$d'_{\varepsilon\mu}(X_0) = \min_{\substack{Z \in L_{\varepsilon}(X_0) \\ V \in \Gamma_{\mu}^{+}(X_0)}} \|Z - V\|,$$

где (см. § 2, п. 2)

$$\Gamma_{\mu}^{+}(X_0) = \left\{ V = - \sum_{j \in Q'_{\mu}(X_0)} \alpha_j \frac{\partial h_j(X_0)}{\partial X} \mid \alpha_j \geq 0 \right\},$$

$$Q'_{\mu}(X_0) = \{j \in [0 : N_1] \mid -\mu \leq t_j(X_0) h_j(X_0) \leq 0\},$$

$$t_j(X_0) = \left\| \frac{\partial h_j(X_0)}{\partial X} \right\|^{-1}.$$

Очевидно, $d'_{00}(X_0) = 0$ тогда и только тогда, когда $d(X_0) = 0$. Таким образом, равенство $d'_{00}(X_0) = 0$ определяет стационарные точки функции $\varphi(X)$ на Ω . Если $d'_{\varepsilon\mu}(X_0) = \|Z_{\varepsilon\mu}(X_0) - V_{\varepsilon\mu}(X_0)\| > 0$, то положим

$$g'_{\varepsilon\mu}(X_0) = \frac{V_{\varepsilon\mu}(X_0) - Z_{\varepsilon\mu}(X_0)}{\|V_{\varepsilon\mu}(X_0) - Z_{\varepsilon\mu}(X_0)\|}.$$

Направление $g'_{\varepsilon\mu}(X_0)$ обладает следующими свойствами:

$$\left(\frac{\partial f_i(X_0)}{\partial X}, g'_{\varepsilon\mu}(X_0)\right) \leq -d'_{\varepsilon\mu}(X_0) \quad \text{для } i \in R_\varepsilon(X_0), \quad (9.2)$$

$$\left(\frac{\partial h_j(X_0)}{\partial X}, g'_{\varepsilon\mu}(X_0)\right) \leq 0 \quad \text{для } j \in Q'_\mu(X_0). \quad (9.3)$$

Лемма 9.1. Пусть $X_0 \in \Omega$, множество $M(X_0) = \{X \in \Omega \mid \varphi(X) < \varphi(X_0)\}$ ограничено и выполнено условие Слейтера (9.1). Тогда существуют такие $a' > 0$ и $\mu' > 0$, что для всех $\mu \in [0, \mu']$ и всех $X \in (\partial'_\mu \Omega) \cap M(X_0)$ будет

$$\max_{j \in Q'_\mu(X)} \left(\frac{\partial h_j(X)}{\partial X}, q(X)\right) \leq -a',$$

где $q(X) = \frac{\bar{X} - X}{\|\bar{X} - X\|}$ и $\bar{X}$ — точка из (9.1).

Доказательство. Допустим противное. Тогда найдутся две монотонно убывающие и стремящиеся к нулю числовые последовательности $\{a_k\}$ и $\{\mu_k\}$, последовательность векторов $\{X_k\}$, $X_k \in (\partial'_{\mu_k} \Omega) \cap M(X_0)$, $k = 1, 2, \dots$, и индекс $j_0 \in Q'_{\mu_k}(X_k)$ такие, что

$$\left(\frac{\partial h_{j_0}(X_k)}{\partial X}, q(X_k)\right) > -a_k.$$

Не ограничивая общности, будем считать, что $X_k \xrightarrow{k \rightarrow \infty} X^*$, причем $X^* \in \Omega$. Поскольку $j_0 \in Q'_{\mu_k}(X_k)$ при всех k , то

$$-\mu_k \leq t_{j_0}(X_k) h_{j_0}(X_k) \leq 0.$$

Отсюда следует, что $h_{j_0}(X^*) = 0$, т. е. $X^* \in \partial\Omega$. В частности, $X^* \neq \bar{X}$. Теперь имеем

$$-a_k < \left(\frac{\partial h_{j_0}(X_k)}{\partial X}, q(X_k)\right) \leq \frac{h_{j_0}(\bar{X}) - h_{j_0}(X_k)}{\|\bar{X} - X_k\|}.$$

Переходя в этом неравенстве к пределу при $k \rightarrow \infty$, получим: $h_{j_0}(\bar{X}) \geq 0$, что противоречит (9.1).

Лемма доказана.

2. Переходим к описанию метода последовательных приближений для нахождения стационарных точек функции $\varphi(X)$ на Ω . Зафиксируем $\varepsilon_0 > 0$, $\mu_0 > 0$, $\rho_0 > 0$

и выберем начальное приближение $X_0 \in \Omega$. Будем считать, что множество $M(X_0)$ ограничено и что выполняется неравенство $\mu_0 \leq \mu'$, где μ' — число, фигурирующее в лемме 9.1.

Пусть уже найдено $X_k \in M(X_0)$. Если $d'_{00}(X_k) = 0$, то X_k — стационарная точка, и процесс на этом заканчивается.

Предположим, что $d'_{00}(X_k) > 0$. Найдем наименьшее $v \in \{0, 1, 2, \dots\}$, при котором

$$d'_{\varepsilon_v \mu_v}(X_k) \geq \rho_v, \quad (9.4)$$

где $\varepsilon_v = \varepsilon_0/2^v$, $\mu_v = \mu_0/2^v$, $\rho_v = \rho_0/2^v$. Обозначим его через v_k . Положим также

$$\hat{\varepsilon}_k = \varepsilon_{v_k}, \quad \hat{\mu}_k = \mu_{v_k}, \quad \hat{\rho}_k = \rho_{v_k}, \\ d'_k = d'_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k), \quad g'_k = g'_{\hat{\varepsilon}_k \hat{\mu}_k}(X_k).$$

Если $Q'_{\hat{\mu}_k}(X_k) = \emptyset$, то направление g'_k является допустимым. В этом случае в качестве X_{k+1} можно взять точку минимума функции $\varphi(X)$ на множестве

$$\{X = X_k + \alpha g'_k \mid \alpha \geq 0, X \in \Omega\}.$$

Пусть теперь $Q'_{\hat{\mu}_k}(X_k) \neq \emptyset$. Направление g'_k может оказаться недопустимым, поэтому его необходимо скорректировать. Положим

$$c_k = \max_{i \in [0: N]} \left\| \frac{\partial f_i(X_k)}{\partial X} \right\|, \\ \beta_k = \min \left\{ \frac{1}{3}, \frac{d'_k}{3c_k} \right\}, \\ p_k = \frac{g'_k + \beta_k q(X_k)}{\|g'_k + \beta_k q(X_k)\|}.$$

Очевидно, $\beta_k > 0$ и $2/3 \leq \|g'_k + \beta_k q(X_k)\| \leq 4/3$. Кроме того, имеем в силу (9.2), (9.3) и леммы 9.1

$$\left(\frac{\partial f_i(X_k)}{\partial X}, p_k \right) \leq -d'_k/2 \quad \text{для } i \in R_{\hat{\mu}_k}(X_k), \quad (9.5)$$

$$\left(\frac{\partial h_j(X_k)}{\partial X}, p_k \right) \leq -3/4 \alpha' \beta_k \quad \text{для } j \in Q'_{\hat{\mu}_k}(X_k). \quad (9.6)$$

Из последнего неравенства следует существование такого $\gamma_k > 0$, что при $\alpha \in [0, \gamma_k]$ будет

$$X'_k(\alpha) = X_k + \alpha p_k \in \Omega.$$

Найдем теперь $\alpha_k \geq 0$, при котором

$$\varphi(X'_k(\alpha_k)) = \min_{\substack{\alpha \geq 0 \\ X'_k(\alpha) \in \Omega}} \varphi(X'_k(\alpha)),$$

и положим $X_{k+1} = X'_k(\alpha_k)$. Очевидно,

$$X_{k+1} \in M(X_0) \quad \text{и} \quad \varphi(X_{k+1}) < \varphi(X_k).$$

Продолжая аналогично, получим последовательность $\{X_k\}$. Если эта последовательность конечна, то последний ее элемент по построению является стационарной точкой функции $\varphi(X)$ на Ω . В противном случае справедлива

Теорема 9.1. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на Ω .*

Доказательство этой теоремы аналогично доказательству теоремы 8.1. При этом необходимо дополнительно учесть соотношения (9.5) и (9.6), определение β_k и лемму 2.6.

Замечание. Описанный метод может быть применен и при выполнении обобщенного условия Слейтера (2.23). При этом следует предположить, что линейные ограничения записаны в нормальной форме

$$h_j(X) \stackrel{\text{def}}{=} (A_j, X) + b_j \leq 0, \quad \|A_j\| = 1, \quad j \in [N' + 1 : N_1].$$

Точка $\bar{X}$ будет удовлетворять неравенствам

$$\begin{aligned} h_j(\bar{X}) &< 0, & j \in [0 : N'], \\ h_j(\bar{X}) &\leq 0, & j \in [N' + 1 : N_1]. \end{aligned}$$

3*. Изложенный в предыдущем пункте метод использовал для корректировки g'_k направление $q(X_k)$, ведущее строго внутрь множества Ω . Однако корректировку g'_k можно производить и иначе.

Будем считать, что выполнено условие Слейтера. Поскольку множество $M(X_0)$ ограничено, то при

некотором $a'' > 0$ имеет место неравенство

$$\max_{j \in Q'_\mu(X)} t_j(X) \left(\frac{\partial h_j(X)}{\partial X}, q(X) \right) \leq -a'' \quad (9.7)$$

для всех $\mu \in [0, \mu']$ и всех $X \in (\partial'_\mu \Omega) \cap M(X_0)$. Это непосредственно следует из леммы 9.1.

Введем множества

$$l_{j\xi}(X) = \left\{ Z \left\| Z + t_j(X) \frac{\partial h_j(X)}{\partial X} \right\| \leq \xi \right\},$$

$$H_{\mu\xi}(X) = \text{co} \bigcup_{j \in Q'_\mu(X)} l_{j\xi}(X).$$

Нетрудно проверить, опираясь на (9.7), что при $\xi \in [0, a'']$, $\mu \in [0, \mu']$ и $X \in (\partial'_\mu \Omega) \cap M(X_0)$ будет

$$0 \notin H_{\mu\xi}(X).$$

В силу леммы 2.5 Приложения II конус $\Gamma_{\mu\xi}^+(X)$, натянутый на множество $\bigcup_{j \in Q'_\mu(X)} l_{j\xi}(X)$, замкнут. Если $Q'_\mu(X) = \emptyset$,

то по определению полагаем $\Gamma_{\mu\xi}^+(X) = \{0\}$ для всех $\xi \geq 0$.

Заметим, что если для некоторого вектора $p \in E_n$, $\|p\| = 1$, оказалось

$$(p, g) \geq 0 \quad \text{для всех} \quad g \in \Gamma_{\mu\xi}^+(X), \quad (9.8)$$

то для любого $j \in Q'_\mu(X)$ будет

$$\left(p, t_j(X) \frac{\partial h_j(X)}{\partial X} \right) \leq -\xi. \quad (9.9)$$

Действительно, достаточно в неравенстве (9.8) положить

$$g = -\xi p - t_j(X) \frac{\partial h_j(X)}{\partial X}.$$

Теперь зафиксируем $\xi_0 \in (0, a'')$. Пусть уже найдено $X_k \in \Omega$. Если $d'_{00}(X_k) = 0$, то точка X_k стационарная. В противном случае найдем $\hat{\varepsilon}_k, \hat{\mu}_k, \hat{\rho}_k, d'_k$, как было описано выше. Допустим, что $Q'_{\hat{\mu}_k}(X_k) \neq \emptyset$.

Для каждого $\xi_s = (1/2)^s \xi_0$, $s = 0, 1, 2, \dots$, положим

$$\bar{d}_{ks} = \min_{\substack{Z \in L_{\hat{\varepsilon}_k}(X_k) \\ V \in \Gamma_{\hat{\mu}_k \xi_s}^+(X_k)}} \|Z - V\| = \|Z_{ks} - V_{ks}\|.$$

Обозначим через s_k первое s , при котором

$$\bar{d}_{ks} \geq \frac{1}{2} d'_k.$$

Очевидно, s_k — конечное число. Введем обозначения

$$\bar{d}_k = \bar{d}_{ks_k}, \quad \bar{\xi}_k = \xi_{s_k}, \quad \bar{g}_k = \bar{d}_k^{-1} (V_{ks_k} - Z_{ks_k}).$$

Имеем $\bar{d}_k \geq \frac{1}{2} d'_k$ и в силу (9.8) и (9.9)

$$\left(\bar{g}_k, t_j(X_k) \frac{\partial h_j(X_k)}{\partial X} \right) \leq -\bar{\xi}_k$$

для $j \in Q'_{\mu_k}(X_k)$.

Теперь рассмотрим луч $X = X_k(\alpha) \stackrel{df}{=} X_k + \alpha \bar{g}_k$, $\alpha \geq 0$. Найдем α_k , при котором

$$\varphi(X_k(\alpha_k)) = \min_{\substack{\alpha \geq 0 \\ X_k(\alpha) \in \Omega}} \varphi(X_k(\alpha))$$

и положим $X_{k+1} = X_k(\alpha_k)$.

Для построенной таким образом последовательности $\{X_k\}$ справедлива теорема, аналогичная теореме 9.1.

§ 10. Метод штрафных функций

1. В этом параграфе будет показано, как задача минимизации функции максимума $\varphi(X)$ на множестве Ω , заданном системой неравенств, сводится к решению последовательности минимаксных задач на всем пространстве E_n .

Пусть

$$\varphi(X) = \max_{i \in [0 : M]} f_i(X),$$

$$\Omega = \{X \in E_n \mid h_j(X) \leq 0 \text{ для всех } j \in [0 : N_1]\}.$$

Функции $f_i(X)$ и $h_j(X)$ предполагаются непрерывными на E_n . Предположим также, что существует точка $\bar{X} \in \Omega$, для которой

$$\varphi(\bar{X}) = \min_{X \in \Omega} \varphi(X),$$

и что множество

$$M_0(\bar{X}) = \{X \in E_n \mid \varphi(X) \leq \varphi(\bar{X})\}$$

ограничено. В силу непрерывности $\varphi(X)$ множество $M_0(\bar{X})$ является и замкнутым.

Образует функцию

$$\varphi_\lambda(X) = \varphi(X) + \sum_{j=0}^{N_1} \lambda_j h_j^2(X) \delta_j(X),$$

где

$$\lambda = (\lambda_0, \dots, \lambda_{N_1})$$

и

$$\delta_j(X) = \begin{cases} 0, & \text{если } h_j(X) \leq 0, \\ 1, & \text{если } h_j(X) > 0. \end{cases}$$

Очевидно, $\varphi_\lambda(X) = \varphi(X)$, если $X \in \Omega$. Учитывая, что

$$h_j(X) \delta_j(X) = \max\{h_j(X), 0\},$$

заключаем, что функция $\varphi_\lambda(X)$ при каждом фиксированном λ непрерывна по X на E_n .

Пусть $\bar{m}(\lambda) = \min_{j \in [0: N_1]} \lambda_j$. Будем писать $\lambda > 0$, если $m(\lambda) > 0$. Через X_λ обозначим точку, для которой

$$\varphi_\lambda(X_\lambda) = \min_{X \in E_n} \varphi_\lambda(X).$$

При любом $\lambda > 0$ такая точка существует. Действительно, положим

$$M_\lambda(\bar{X}) = \{X | \varphi_\lambda(X) \leq \varphi_\lambda(\bar{X})\}.$$

Тогда $M_\lambda(\bar{X}) \subset M_0(\bar{X})$, ибо для каждого $X \in M_\lambda(\bar{X})$ будет

$$\varphi(X) \leq \varphi_\lambda(X) \leq \varphi_\lambda(\bar{X}) = \varphi(\bar{X}).$$

Значит, множество $M_\lambda(\bar{X})$ ограничено. В силу непрерывности $\varphi_\lambda(X)$ по X оно и замкнуто. Отсюда следует, что существует точка X_λ , для которой

$$\varphi_\lambda(X_\lambda) = \min_{X \in M_\lambda(\bar{X})} \varphi_\lambda(X) = \min_{X \in E_n} \varphi_\lambda(X),$$

при этом $X_\lambda \in M_0(\bar{X})$ при всех $\lambda > 0$.

Заметим также, что если при некотором $\lambda > 0$ оказалось $X_\lambda \in \Omega$, то X_λ является точкой минимума функции $\varphi(X)$ на Ω .

Лемма 10.1. Если функции $f_i(X)$ и $h_j(X)$ непрерывны на E_n и множество $M_0(\bar{X})$ ограничено, то

$$\lim_{m(\lambda) \rightarrow \infty} \varphi(X_\lambda) = \varphi(\bar{X}).$$

Доказательство. Допустим противное. Тогда найдется последовательность векторов $\lambda_k > 0$, $m(\lambda_k) \xrightarrow{k \rightarrow \infty} \infty$, и число $\varepsilon > 0$ такие, что

$$\varphi(X_{\lambda_k}) \leq \varphi(\bar{X}) - \varepsilon. \quad (10.1)$$

Не ограничивая общности, будем считать, что

$$X_{\lambda_k} \xrightarrow{k \rightarrow \infty} X^* \in M_0(\bar{X}).$$

Из (10.1) следует

$$\varphi(X^*) \leq \varphi(\bar{X}) - \varepsilon.$$

В силу определения $\bar{X}$ имеем $X^* \notin \Omega$. Значит,

$$\sum_{j=0}^{N_1} \lambda_j^{(k)} h_j^2(X_{\lambda_k}) \delta_j(X_{\lambda_k}) \xrightarrow{m(\lambda_k) \rightarrow \infty} \infty.$$

Но последовательность $\{\varphi(X_{\lambda_k})\}$ ограничена, поэтому

$$\varphi_{\lambda_k}(X_{\lambda_k}) \xrightarrow{k \rightarrow \infty} \infty. \quad (10.2)$$

С другой стороны, вспоминая определение X_{λ_k} , запишем

$$\varphi_{\lambda_k}(X_{\lambda_k}) \leq \varphi_{\lambda_k}(\bar{X}) = \varphi(\bar{X}),$$

что противоречит (10.2).

Лемма доказана.

Следствие. Поскольку при $\lambda > 0$

$$\varphi(X_\lambda) \leq \varphi_\lambda(X_\lambda) \leq \varphi(\bar{X}),$$

то

$$\varphi_\lambda(X_\lambda) \xrightarrow{m(\lambda) \rightarrow \infty} \varphi(\bar{X}).$$

Теорема 10.1. Пусть выполнены условия леммы 10.1 и $\{\lambda_k\}$ — произвольная последовательность векторов такая, что $m(\lambda_k) \xrightarrow{k \rightarrow \infty} \infty$. Тогда любая предельная точка X^* последовательности $\{X_{\lambda_k}\}$, где X_{λ_k} — точка минимума функции $\varphi_{\lambda_k}(X)$ на E_n , является точкой минимума функции $\varphi(X)$ на Ω .

Доказательство. В силу леммы 10.1.

$$\varphi(X^*) = \min_{X \in \Omega} \varphi(X).$$

Осталось показать, что $X^* \in \Omega$. Допустив противное, придем к противоречию так же, как в лемме 10.1.

Теорема доказана.

2. Положим

$$f_{\lambda i}(X) = f_i(X) + \sum_{j=0}^{N_1} \lambda_j h_j^2(X) \delta_j(X).$$

Тогда будем иметь

$$\Phi_{\lambda}(X) = \max_{i \in [0 : N]} f_{\lambda i}(X).$$

В теореме 10.1 задача минимизации функции максимума $\varphi(X)$ на Ω сведена к задаче минимизации функций максимума $\Phi_{\lambda_k}(X)$ на E_n при различных λ_k таких, что $m(\lambda_k) \xrightarrow{k \rightarrow \infty} \infty$.

Функция $\sum_{j=0}^{N_1} \lambda_j h_j^2(X) \delta_j(X)$ называется *функцией штрафа*, а методы минимизации $\varphi(X)$ на Ω , использующие $\Phi_{\lambda}(X)$, называются *методами штрафных функций*.

Сделаем еще одно полезное замечание. Если функции $f_i(X)$ и $h_j(X)$ непрерывно дифференцируемы на E_n , то и $f_{\lambda i}(X)$, $i \in [0 : N]$, при любом фиксированном λ являются непрерывно дифференцируемыми на E_n , причем

$$\frac{\partial f_{\lambda i}(X)}{\partial X} = \frac{\partial f_i(X)}{\partial X} + \sum_{j=0}^{N_1} \lambda_j (|h_j(X)| + h_j(X)) \frac{\partial h_j(X)}{\partial X}. \quad (10.3)$$

Достаточно доказать непрерывную дифференцируемость функций

$$\Phi_j(X) = h_j^2(X) \delta_j(X), \quad j \in [0 : N_1].$$

Заметим, что

$$\Phi_j(X) = \frac{1}{2} h_j(X) (|h_j(X)| + h_j(X)) = \frac{1}{2} (\xi(h_j(X)) + h_j^2(X)),$$

где $\xi(t) = t|t|$, $-\infty < t < \infty$. Функция $\xi(t)$ непрерывно дифференцируема на всей числовой прямой, причем $\xi'(t) = 2|t|$. Поэтому $\xi(h_j(X))$ непрерывно дифференцируема на E_n и

$$\frac{\partial \xi(h_j(X))}{\partial X} = \xi'(h_j(X)) \frac{\partial h_j(X)}{\partial X} = 2|h_j(X)| \frac{\partial h_j(X)}{\partial X}.$$

Значит, непрерывно дифференцируемыми на E_n будут и $\Phi_j(X)$; при этом

$$\frac{\partial \Phi_j(X)}{\partial X} = (|h_j(X)| + h_j(X)) \frac{\partial h_j(X)}{\partial X}. \quad (10.4)$$

Формула (10.3) следует теперь из определения $f_{\lambda i}(X)$ и (10.4).

§ 11. Заключительные замечания

1. Сделаем некоторые замечания о методах последовательных приближений для нахождения стационарных точек функции $\Phi(X)$ на Ω . Если множество Ω задается линейными неравенствами, то весьма эффективным является метод, основанный на нахождении (ε, μ) -стационарных точек. Идея этого метода изложена в § 8. Однако следует иметь в виду, что в этом случае поиск направления спуска сводится к задаче квадратичного программирования:

$$\left\| \sum_i \alpha_i \frac{\partial f_i(X_0)}{\partial X} + \sum_j \beta_j \frac{\partial h_j(X_0)}{\partial X} \right\|^2 \rightarrow \min,$$

где $\sum_i \alpha_i = 1$, $\alpha_i \geq 0$, $\beta_j \geq 0$.

Методы §§ 6 и 9 имеют в основном теоретическое значение. С помощью метода штрафных функций (§ 10) минимаксная задача с ограничениями сводится к минимаксной задаче без ограничений. В этом случае могут быть использованы методы главы III. При пользовании методом, изложенным в § 7, приходится решать те же вспомогательные задачи, что и в случае отсутствия ограничений на параметры (см. § 9 гл. III). Однако получающееся направление спуска ни в каком разумном смысле не является скорейшим.

2. Во всех методах в качестве начального приближения берется произвольная точка X_0 из Ω . Заметим, что не всегда легко определить, принадлежит ли данная точка множеству Ω или нет. В этом случае можно рекомендовать одним из методов главы III решать следующую минимаксную задачу:

$$\max_{j \in [0 : N_1]} h_j(X) \rightarrow \min_{X \in E_n}.$$

Поскольку функции $h_j(X)$ выпуклы на E_n , то будет построена последовательность $\{X'_k\}$, любая предельная точка которой является точкой минимума функции $\varphi_1(X)$ на E_n . В качестве X_0 можно взять любую точку последовательности $\{X'_k\}$, для которой

$$\max_{j \in [0 : N_1]} h_j(X'_k) \leq 0.$$

Заметим, что таким же образом можно проверить выполнимость условия Слейтера.

3. Исходная задача

$$\max_{i \in [0 : M]} f_i(X) \rightarrow \min_{X \in \Omega},$$

где $\Omega = \{X \mid h_j(X) \leq 0, j \in [0 : N_1]\}$, равносильна следующей задаче нелинейного программирования в пространстве векторов $W = (x_1, \dots, x_n, u)$: найти $\min u$ при дополнительных условиях

$$\begin{aligned} f_i(x_1, \dots, x_n) - u &\leq 0, \quad i \in [0 : M]; \\ h_j(x_1, \dots, x_n) &\leq 0, \quad j \in [0 : N_1]. \end{aligned}$$

Отсюда следует, что исходная задача сводится к задаче линейного программирования, если функции $f_i(X)$ и $h_j(X)$ линейны.

4. В § 6 главы IV доказано следующее утверждение: если для некоторой точки $X_0 \in \Omega$ и некоторого индексного множества $Q \subset [0 : N]$ будет

$$\min_{Z \in \Omega} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) = -a \leq 0,$$

причем функции $f_i(X)$ выпуклы на Ω при $i \in Q$, то

$$\min_{i \in Q} f_i(X_0) - a \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0). \quad (11.1)$$

Предположим теперь, что множество Ω задано в виде $\Omega = \{X \mid h_j(X) \leq 0, j \in [0 : N_1]\}$ и ограничено. Последнее означает, что найдутся константы $K_l, l \in [1 : n]$, такие, что $|x_l| \leq K_l$ для всех $X = (x_1, \dots, x_n) \in \Omega$. Положим

$$h_{N_1+l}(X) = x_l - K_l, \quad l \in [1 : n];$$

$$h_{N_1+n+l}(X) = -x_l - K_l, \quad l \in [1 : n].$$

Тогда

$$\Omega = \{X \mid h_j(X) \leq 0, j \in [0 : N_1 + 2n]\}.$$

Введем множество

$$\Omega(X_0) = \left\{ Z \mid h_j(X_0) + \left(\frac{\partial h_j(X_0)}{\partial X}, Z - X_0 \right) \leq 0, j \in [0 : N_1 + 2n] \right\}.$$

Очевидно, $\Omega(X_0)$ ограничено и $\Omega \subset \Omega(X_0)$.

Предположим теперь, что функции $f_i(X)$ непрерывно дифференцируемы на открытом множестве Ω'' , содержащем $\Omega(X_0)$.

Поскольку

$$\min_{Z \in \Omega(X_0)} \max_{i \in Q} \left(\frac{\partial f_i(X_0)}{\partial X}, Z - X_0 \right) \stackrel{df}{=} -a' \leq -a,$$

то в силу (11.1)

$$\min_{i \in Q} f_i(X_0) - a' \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0). \quad (11.2)$$

Левое неравенство в (11.2) более грубое, чем соответствующее неравенство в (11.1), однако вычисление a' сводится к задаче линейного программирования, в то время как вычисление a может представлять более сложную задачу.

НЕПРЕРЫВНАЯ МИНИМАКСНАЯ ЗАДАЧА

§ 1. Постановка задачи

Пусть Ω' — открытое множество в E_n , G — ограниченное замкнутое множество в E_m , $F(X, Y)$ — функция, непрерывная и непрерывно дифференцируемая по X на $\Omega' \times G$. Рассматривается задача минимизации функции

$$\varphi(X) = \max_{Y \in G} F(X, Y)$$

на выпуклом замкнутом множестве $\Omega \subset \Omega'$.

В §§ 2 и 3 обобщаются на непрерывный случай результаты, полученные в главах III—V (дифференцируемость по направлению функции $\varphi(X)$, необходимые условия минимакса, их геометрическая интерпретация, некоторые оценки).

§ 4 посвящен сеточным методам нахождения стационарных точек функции $\varphi(X)$ на Ω ; отдельно рассматривается случай $\Omega = E_n$.

В § 5 приводится один частный случай теоремы о минимаксе, т. е. выясняются условия, при которых

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) = \max_{Y \in G} \min_{X \in \Omega} F(X, Y).$$

В § 6 рассмотрен вопрос о нахождении седловых точек на многогранниках.

В последних двух параграфах как следствие общих теорем получены некоторые классические результаты теории чебышевских приближений.

§ 2. Основные теоремы

1. Пусть задана функция $F(X, Y)$, $X \in \Omega'$, $Y \in G$, где Ω' — открытое множество в E_n , а G — ограниченное замкнутое множество в E_m . Будем предполагать, что функция $F(X, Y)$ непрерывна вместе с $\partial F(X, Y)/\partial X$ по совокупности переменных в $\Omega' \times G$.

Положим

$$\varphi(X) = \max_{Y \in G} F(X, Y).$$

Отметим простейшие свойства функции $\varphi(X)$.

I. Функция $\varphi(X)$ непрерывна на множестве Ω' .

II. Если при каждом $Y \in G$ функция $F(X, Y)$ выпукла по X на выпуклом множестве $\Omega \subset \Omega'$, то $\varphi(X)$ выпукла на Ω .

III. Если $\Omega \subset \Omega'$ — замкнутое множество и при некотором $X_0 \in \Omega$ множество

$$M(X_0) = \{X \in \Omega \mid \varphi(X) \leq \varphi(X_0)\}$$

ограничено, то $\varphi(X)$ достигает своего инфимума на Ω , т. е. существует точка $X^* \in \Omega$, для которой

$$\varphi(X^*) = \inf_{X \in \Omega} \varphi(X),$$

или, что то же самое,

$$\max_{Y \in G} F(X^*, Y) = \inf_{X \in \Omega} \max_{Y \in G} F(X, Y).$$

Доказательство этих простых утверждений предоставляется читателю (ср. § 2 гл. III).

При фиксированном $X \in \Omega'$ введем множество

$$R(X) = \{Y \in G \mid F(X, Y) = \varphi(X)\}.$$

Очевидно, $R(X) \subset G$ является ограниченным замкнутым множеством в E_m .

Теорема 2.1. *Функция $\varphi(X)$ имеет в каждой точке $X \in \Omega'$ производную по любому направлению $g \in E_n$, $\|g\| = 1$, причем*

$$\frac{\partial \varphi(X)}{\partial g} = \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right). \quad (2.1)$$

Для доказательства этой теоремы понадобятся две леммы.

Лемма 2.1. *Справедливо предельное соотношение*

$$\lim_{\alpha \rightarrow +0} \max_{Y \in R(X+\alpha g)} \min_{V \in R(X)} \|Y - V\| = 0.$$

Доказательство. Допустим противное. Тогда найдется последовательность положительных чисел $\{\alpha_k\}$,

стремящаяся к нулю, и число $a > 0$ такие, что

$$\max_{Y \in R(X + \alpha_k g)} \min_{V \in R(X)} \|Y - V\| \geq a.$$

Выделим точки $Y_k \in R(X + \alpha_k g)$, для которых

$$\min_{V \in R(X)} \|Y_k - V\| \geq a. \quad (2.2)$$

Поскольку все Y_k принадлежат G , то, не ограничивая общности, можно считать, что последовательность $\{Y_k\}$ сходится. Переходя к пределу в равенстве

$$F(X + \alpha_k g, Y_k) = \varphi(X + \alpha_k g),$$

получим

$$F(X, Y^*) = \varphi(X).$$

Таким образом, $Y^* \in R(X)$.

Запишем неравенства

$$0 \leq \min_{V \in R(X)} \|Y_k - V\| \leq \|Y_k - Y^*\|.$$

Отсюда следует, что $\min_{V \in R(X)} \|Y_k - V\| \xrightarrow{k \rightarrow \infty} 0$, что противоречит (2.2).

Лемма доказана.

З а м е ч а н и е. Смысл этой леммы заключается в следующем: *по любому $\varepsilon > 0$ найдется такое $\alpha_0 > 0$, что при каждом $\alpha \in (0, \alpha_0)$ для всякого $Y \in R(X + \alpha g)$ можно указать такое $V \in R(X)$, что*

$$\|Y - V\| < \varepsilon.$$

Введем обозначение: $Q(X, \alpha) = R(X + \alpha g) \cup R(X)$.

Л е м м а 2.2. *Равномерно по $g \in E_n$, $\|g\| = 1$, справедливо предельное соотношение*

$$\lim_{\alpha \rightarrow +0} \left\{ \max_{Y \in Q(X, \alpha)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) - \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \right\} = 0. \quad (2.3)$$

Д о к а з а т е л ь с т в о. Поскольку при всех $\alpha > 0$

$$Q(X, \alpha) \supset R(X),$$

то выражение, стоящее в фигурных скобках в формуле (2.3), неотрицательно. Возьмем $\varepsilon > 0$. Тогда най-

дется $\delta > 0$ такое, то из неравенства $\|Y' - Y\| < \delta$ ($Y, Y' \in G$) будет следовать

$$\left| \left(\frac{\partial F(X, Y')}{\partial X}, g \right) - \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \right| < \varepsilon$$

равномерно по g , $\|g\| = 1$. В силу леммы 2.1 по уже выбранному $\delta > 0$ можно указать такое $\alpha_0 > 0$, что для $\alpha \in (0, \alpha_0)$ и любого $Y \in Q(X, \alpha)$ найдется $V \in R(X)$, для которого

$$\|Y - V\| < \delta.$$

Зафиксируем $\alpha \in (0, \alpha_0)$. Пусть $Y_\alpha \in Q(X, \alpha)$ — такая точка, что

$$\max_{Y \in Q(X, \alpha)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) = \left(\frac{\partial F(X, Y_\alpha)}{\partial X}, g \right).$$

Обозначим через $V_\alpha \in R(X)$ точку, для которой

$$\|Y_\alpha - V_\alpha\| < \delta.$$

Имеем равномерно по g , $\|g\| = 1$,

$$\begin{aligned} 0 \leq \max_{Y \in Q(X, \alpha)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) - \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) &\leq \\ &\leq \left(\frac{\partial F(X, Y_\alpha)}{\partial X}, g \right) - \left(\frac{\partial F(X, V_\alpha)}{\partial X}, g \right) < \varepsilon. \end{aligned}$$

Лемма доказана.

Доказательство теоремы. Заметим, во-первых, что

$$F(X + \alpha g, Y) = F(X, Y) + \alpha \left(\frac{\partial F(X, Y)}{\partial X}, g \right) + o(Y, g; \alpha),$$

где

$$\frac{o(Y, g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$$

равномерно по $Y \in G$ и g , $\|g\| = 1$.

Отсюда и из определения $Q(X, \alpha)$ следует

$$\begin{aligned} \Phi(X + \alpha g) &\stackrel{df}{=} \max_{Y \in G} F(X + \alpha g, Y) = \max_{Y \in R(X + \alpha g)} F(X + \alpha g, Y) \leq \\ &\leq \max_{Y \in Q(X, \alpha)} F(X + \alpha g, Y) \leq \max_{Y \in Q(X, \alpha)} F(X, Y) + \\ &+ \alpha \max_{Y \in Q(X, \alpha)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) + \max_{Y \in G} o(Y, g; \alpha). \quad (2.4) \end{aligned}$$

Будем использовать обозначение

$$\max_{\|g\|=1} \max_{Y \in G} |o(Y, g; \alpha)| = o(\alpha).$$

Ясно, что $\frac{o(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$. Поскольку $\max_{Y \in Q(X, \alpha)} F(X, Y) = \varphi(X)$, то, учитывая (2.4), получим

$$\varphi(X + \alpha g) - \varphi(X) \leq \alpha \max_{Y \in Q(X, \alpha)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) + o(\alpha).$$

Воспользуемся результатом леммы 2.2. Это дает

$$\varphi(X + \alpha g) - \varphi(X) \leq \alpha \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) + \bar{o}(\alpha), \quad (2.5)$$

где по-прежнему

$$\frac{\bar{o}(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0.$$

С другой стороны,

$$\varphi(X + \alpha g) \geq$$

$$\begin{aligned} &\geq \max_{Y \in G} \left[F(X, Y) + \alpha \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \right] - \max_{Y \in G} |o(Y, g; \alpha)| \geq \\ &\geq \max_{Y \in R(X)} \left[F(X, Y) + \alpha \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \right] - o(\alpha) = \\ &= \varphi(X) + \alpha \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) - o(\alpha). \end{aligned}$$

Отсюда следует

$$\varphi(X + \alpha g) - \varphi(X) \geq \alpha \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) - o(\alpha). \quad (2.6)$$

Объединяя (2.5) и (2.6), получим

$$\begin{aligned} \alpha \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) - o(\alpha) &\leq \varphi(X + \alpha g) - \varphi(X) \leq \\ &\leq \alpha \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) + \bar{o}(\alpha). \end{aligned} \quad (2.7)$$

Таким образом,

$$\frac{\partial \varphi(X)}{\partial g} \stackrel{df}{=} \lim_{\alpha \rightarrow +0} \frac{1}{\alpha} [\varphi(X + \alpha g) - \varphi(X)] = \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right).$$

Теорема доказана.

Более того, в силу (2.7) можно считать доказанной следующую формулу:

$$\varphi(X + \alpha g) = \varphi(X) + \alpha \frac{\partial \varphi(X)}{\partial g} + o(g; \alpha), \quad (2.8)$$

где $\frac{o(g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow +0} 0$ равномерно по g , $\|g\| = 1$.

2. В дальнейшем будем рассматривать задачу минимизации функции $\varphi(X)$ на замкнутом выпуклом множестве $\Omega \subset \Omega'$.

Теорема 2.2. Для того чтобы функция $\varphi(X)$ достигала минимального на Ω значения в точке $X^ \in \Omega$, необходимо, а в случае выпуклости $\varphi(X)$ на Ω — и достаточно, чтобы*

$$\inf_{Z \in \Omega} \max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, Z - X^* \right) = 0. \quad (2.9)$$

Если воспользоваться понятием производной по направлению и формулой (2.8), то одно доказательство этой теоремы может быть получено дословно так же, как доказательство теоремы 2.1 главы IV.

Приведем другое независимое доказательство сформулированной теоремы.

Необходимость. Пусть в точке $X^* \in \Omega$ функция $\varphi(X)$ достигает минимального на Ω значения

$$\varphi(X^*) = \min_{X \in \Omega} \varphi(X), \quad (2.10)$$

однако равенство (2.9) не выполняется. Тогда найдется точка $Z_1 \in \Omega$ такая, что

$$\max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \stackrel{df}{=} -\varepsilon < 0. \quad (2.11)$$

Обозначим через $\tilde{R}$ множество тех $Y \in G$, для которых

$$\left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \geq -\frac{\varepsilon}{2}. \quad (2.12)$$

Очевидно, $\tilde{R}$ либо пустое множество, либо компакт (ограниченное замкнутое множество в E_m).

Предположим вначале, что $\tilde{R}$ — компакт. В силу (2.11) и (2.12) $\tilde{R}$ не имеет общих точек с $R(X^*)$, поэтому

$$\rho \stackrel{df}{=} \max_{Y \in \tilde{R}} F(X^*, Y) < \varphi(X^*). \quad (2.13)$$

Положим

$$X(\lambda) = X^* + \lambda(Z_1 - X^*).$$

В силу выпуклости Ω при $0 \leq \lambda \leq 1$

$$X(\lambda) \in \Omega.$$

Обозначим

$$K = \max_{0 \leq \lambda \leq 1} \max_{Y \in G} \left| \left(\frac{\partial F(X(\lambda), Y)}{\partial X}, Z_1 - X^* \right) \right|.$$

Очевидно, K — конечное число. По уже имеющемуся $\varepsilon > 0$ найдем $\delta > 0$ такое, что из неравенства $\|X(\lambda) - X^*\| \leq \delta$ будет следовать

$$\left| \left(\frac{\partial F(X(\lambda), Y)}{\partial X}, Z_1 - X^* \right) - \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \right| \leq \frac{\varepsilon}{4}$$

равномерно по $Y \in G$.

Положим, наконец, $h = \varphi(X^*) - \rho$ и

$$\lambda_0 = \min \left\{ \frac{\delta}{\|Z_1 - X^*\|}, \frac{h}{2K}, 1 \right\}, \quad 0 < \lambda_0 \leq 1.$$

Покажем, что $\varphi(X(\lambda_0)) < \varphi(X^*)$. Поскольку это неравенство противоречит (2.10), то тем самым будет доказана необходимость условия (2.9) в случае, если $\tilde{R}$ — непустое множество.

Запишем для произвольного $Y \in G$ формулу Ньютона — Лейбница

$$F(X(\lambda_0), Y) = F(X^*, Y) + \int_0^{\lambda_0} \left(\frac{\partial F(X(t), Y)}{\partial X}, Z_1 - X^* \right) dt. \quad (2.14)$$

Пусть $Y \in \tilde{R}$. Тогда в силу (2.13), (2.14) и определения λ_0 получим

$$F(X(\lambda_0), Y) \leq \varphi(X^*) - h + K\lambda_0 \leq \varphi(X^*) - \frac{h}{2}.$$

Пусть $Y \in G \setminus \tilde{R}$. В этом случае

$$\begin{aligned} F(X(\lambda_0), Y) &= F(X^*, Y) + \lambda_0 \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) + \\ &+ \int_0^{\lambda_0} \left[\left(\frac{\partial F(X(t), Y)}{\partial X}, Z_1 - X^* \right) - \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \right] dt. \end{aligned}$$

Согласно (2.12) и правилу выбора λ_0

$$F(X(\lambda_0), Y) \leq \varphi(X^*) - \lambda_0 \frac{\varepsilon}{2} + \lambda_0 \frac{\varepsilon}{4} = \varphi(X^*) - \frac{\lambda_0 \varepsilon}{4}. \quad (2.15)$$

Итак,

$$\max_{Y \in G} F(X(\lambda_0), Y) < \varphi(X^*), \quad (2.16)$$

что противоречит (2.10).

Если $\tilde{R}$ — пустое множество, то для всех $Y \in G$

$$\left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) < -\frac{\varepsilon}{2}.$$

В этом случае следует положить

$$\lambda_0 = \min \left\{ \frac{\delta}{\|Z_1 - X^*\|}, 1 \right\}, \quad 0 < \lambda_0 \leq 1.$$

Тогда неравенство (2.15) будет выполняться для всех $Y \in G$, откуда следует (2.16).

Необходимость доказана полностью.

Достаточность. Предположим, что для $X^* \in \Omega$ имеет место (2.9) и что функция $\varphi(X)$ выпукла на Ω , т. е. для $X_1, X_2 \in \Omega$ при $0 \leq \lambda \leq 1$

$$\varphi(X_1 + \lambda(X_2 - X_1)) \leq \varphi(X_1) + \lambda[\varphi(X_2) - \varphi(X_1)]. \quad (2.17)$$

Докажем, что

$$\varphi(X^*) = \min_{X \in \Omega} \varphi(X).$$

Допустим противное. Тогда найдется вектор $Z_1 \in \Omega$ такой, что $\varphi(Z_1) < \varphi(X^*)$.

Положим $X(\lambda) = X^* + \lambda(Z_1 - X^*)$, $0 \leq \lambda \leq 1$, и

$$\varepsilon = \frac{1}{2} [\varphi(X^*) - \varphi(Z_1)] > 0.$$

По $\varepsilon > 0$ найдется такое $\delta > 0$, что из неравенства $\|X(\lambda) - X^*\| \leq \delta$ будет следовать

$$\left| \left(\frac{\partial F(X(\lambda), Y)}{\partial X}, Z_1 - X^* \right) - \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \right| \leq \varepsilon$$

равномерно по $Y \in G$.

Введем число λ_0 ,

$$\lambda_0 = \min \left\{ \frac{\delta}{\|Z_1 - X^*\|}, 1 \right\}, \quad 0 < \lambda_0 \leq 1,$$

и перепишем равенство (2.14) в следующем виде:

$$\lambda_0 \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) = F(X(\lambda_0), Y) - F(X^*, Y) - \\ - \int_0^{\lambda_0} \left[\left(\frac{\partial F(X(t), Y)}{\partial X}, Z_1 - X^* \right) - \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \right] dt.$$

Для $Y \in R(X^*)$ в силу (2.17) и выбора ε получим

$$\lambda_0 \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \leq \varphi(X(\lambda_0)) - \varphi(X^*) + \varepsilon \lambda_0 \leq \\ \leq \lambda_0 [\varphi(Z_1) - \varphi(X^*)] + \varepsilon \lambda_0 = -\varepsilon \lambda_0.$$

Значит,

$$\max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, Z_1 - X^* \right) \leq -\varepsilon < 0,$$

что противоречит (2.9).

Теорема доказана.

З а м е ч а н и е. Если $\Omega = E_n$, то соотношение (2.9) эквивалентно следующему неравенству:

$$\min_{\|g\|=1} \max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, g \right) \geq 0.$$

3. Теорема 2.3. Если для некоторой точки $X_0 \in \Omega$ и некоторого замкнутого множества $Q \subset G$ выполняется неравенство

$$\inf_{Z \in \Omega} \max_{Y \in Q} \left(\frac{\partial F(X_0, Y)}{\partial X}, Z - X_0 \right) \stackrel{df}{=} -a \leq 0, \quad (2.18)$$

а функция $F(X, Y)$ выпукла по X на Ω для всех $Y \in Q$, то

$$\min_{Y \in Q} F(X_0, Y) - a \leq \inf_{X \in \Omega} \varphi(X) \leq \varphi(X_0).$$

Доказательство этой теоремы вполне аналогично доказательству теоремы 5.1 главы IV.

С л е д с т в и е 1. Если в условиях теоремы $Q \subset R(X_0)$ и $a = 0$, то

$$\inf_{X \in \Omega} \varphi(X) = \varphi(X_0),$$

т. е. X_0 является точкой минимума функции $\varphi(X)$ на замкнутом выпуклом множестве Ω ,

Следствие 2. Если $\Omega = E_n$ и $a = 0$, то (2.18) эквивалентно следующему условию:

$$\min_{\|g\|=1} \max_{Y \in Q} \left(\frac{\partial F(X_0, Y)}{\partial X}, g \right) \geq 0.$$

При выполнении этого условия, а также в случае выпуклости функции $F(X, Y)$ по X для всех $Y \in Q$, имеют место неравенства

$$\min_{Y \in Q} F(X_0, Y) \leq \inf_{X \in E_n} \varphi(X) \leq \varphi(X_0).$$

§ 3. Геометрическая интерпретация необходимого условия минимакса. Некоторые следствия

1. Зафиксируем $X \in \Omega$. Пусть, как и раньше,

$$\Gamma(X) = \{V = \lambda(Z - X) \mid \lambda > 0, Z \in \Omega\}.$$

Через $\bar{\Gamma}(X)$ обозначим конус возможных направлений множества Ω в точке X (замыкание конуса $\Gamma(X)$), а через $\Gamma^+(X)$ — сопряженный к $\bar{\Gamma}(X)$ конус.

Положим далее

$$H(X) = \left\{ Z = \frac{\partial F(X, Y)}{\partial X} \mid Y \in R(X) \right\}.$$

Нетрудно понять, что $H(X)$ — ограниченное замкнутое множество в E_n . Обозначим через $L(X)$ выпуклую оболочку, натянутую на точки множества $H(X)$:

$$L(X) = \left\{ Z = \sum_{k=1}^r \alpha_k Z_k \mid Z_k \in H(X), \right. \\ \left. \alpha_k \geq 0, \sum_{k=1}^r \alpha_k = 1, r - \text{произвольное натуральное число} \right\}.$$

В силу леммы 1.2 Приложения II $L(X)$ — ограниченное замкнутое выпуклое множество.

Теорема 3.1. Равенство (2.9) равносильно следующему соотношению:

$$L(X^*) \cap \Gamma^+(X^*) \neq \emptyset. \quad (3.1)$$

Доказательство. Покажем, что из (2.9) следует (3.1). Допустим противное. Пусть равенство (2.9) выполняется, однако

$$L(X^*) \cap \Gamma^+(X^*) = \emptyset.$$

Тогда в силу теоремы отделимости (теорема 2.1 Приложения II) найдется вектор $V_0 \in \Gamma^{++}(X^*) = \bar{\Gamma}(X^*)$ такой, что

$$\max_{V \in L(X^*)} (V_0, Z) \stackrel{df}{=} -a < 0.$$

Функция (V, Z) непрерывна по V равномерно относительно $Z \in L(X^*)$. Поэтому для $V_0 \in \bar{\Gamma}(X^*)$ можно указать вектор $V_1 = \lambda_1(X_1 - X^*) \in \Gamma(X^*)$, для которого

$$\max_{Z \in L(X^*)} (V_1, Z) \leq -\frac{a}{2}.$$

Отсюда следует, что

$$\max_{Z \in L(X^*)} (Z, X_1 - X^*) \leq -\frac{a}{2\lambda_1}.$$

В частности,

$$\max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, X_1 - X^* \right) \leq -\frac{a}{2\lambda_1},$$

что противоречит (2.9).

Покажем, что из (3.1) следует (2.9). Допустим противное. Тогда найдется вектор $V_0 \in \Gamma(X^*) \subset \bar{\Gamma}(X^*)$, для которого

$$\max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, V_0 \right) < 0.$$

Последнее неравенство можно переписать в виде

$$\max_{Z \in H(X^*)} (V_0, Z) < 0.$$

Нетрудно проверить (см. доказательство леммы 3.1 главы III), что

$$\max_{Z \in L(X^*)} (V_0, Z) = \max_{Z \in H(X^*)} (V_0, Z) < 0. \quad (3.2)$$

В силу того, что $V_0 \in \bar{\Gamma}(X^*) = \Gamma^{++}(X^*)$, выполняется неравенство

$$(V_0, Z) \geq 0 \quad \text{для всех } Z \in \Gamma^+(X^*). \quad (3.3)$$

Сопоставляя (3.2) и (3.3), заключаем, что

$$L(X^*) \cap \Gamma^+(X^*) = \emptyset.$$

Но это противоречит (3.1).

Теорема доказана.

Следствие. Если $\Gamma(X^*) = E_n$, то соотношение (3.1) эквивалентно следующему включению:

$$0 \in L(X^*). \quad (3.4)$$

Действительно, в этом случае $\Gamma^+(X^*) = \{0\}$.

2. Теорема 3.2. Для того чтобы функция $\varphi(X)$ достигала минимального на Ω значения в точке $X^* \in \Omega$, необходимо, а в случае выпуклости $\varphi(X)$ на Ω и достаточно, чтобы нашлись r точек $Y_1, \dots, Y_r$ из $R(X^*)$, $1 \leq r \leq n+1$, и r неотрицательных чисел $\alpha_k \geq 0$,

$\sum_{k=1}^r \alpha_k = 1$, такие, что

$$\inf_{Z \in \Omega} \sum_{k=1}^r \alpha_k \left(\frac{\partial F(X^*, Y_k)}{\partial X}, Z - X^* \right) = 0. \quad (3.5)$$

Доказательство. Необходимость. Пусть $X^* \in \Omega$ — точка минимума функции $\varphi(X)$ на Ω . Тогда выполняется соотношение (3.1). Это значит, что существует точка Z_0 из $L(X^*)$, принадлежащая $\Gamma^+(X^*)$. Поскольку любая точка из выпуклой оболочки может быть представлена в виде выпуклой комбинации не более чем $(n+1)$ -й точки исходного множества (см. лемму 1.1 Приложения II), то найдутся точки $Y_1, \dots, Y_r$ из $R(X^*)$

и числа $\alpha_k \geq 0$, $\sum_{k=1}^r \alpha_k = 1$, $1 \leq r \leq n+1$, такие, что

$$Z_0 = \sum_{k=1}^r \alpha_k \frac{\partial F(X^*, Y_k)}{\partial X}.$$

Напомним, что $Z_0^* \in \Gamma^+(X^*)$. Это значит, что для любого $V \in \bar{\Gamma}(X^*)$ будет выполняться неравенство

$$(Z_0, V) \geq 0.$$

В частности, для всех $Z \in \Omega$

$$\sum_{k=1}^r \alpha_k \left(\frac{\partial F(X^*, Y_k)}{\partial X}, Z - X^* \right) \geq 0,$$

откуда следует (3.5).

Достаточность. Пусть выполняется равенство (3.5). Тогда для всех $V \in \bar{\Gamma}(X^*)$

$$\left(\sum_{k=1}^r \alpha_k \frac{\partial F(X^*, Y_k)}{\partial X}, V \right) \geq 0.$$

В силу непрерывности скалярного произведения последнее неравенство имеет место для всех $V \in \bar{\Gamma}(X^*)$. Значит,

$$\sum_{k=1}^r \alpha_k \frac{\partial F(X^*, Y_k)}{\partial X} \in \Gamma^+(X^*).$$

С другой стороны, по определению $L(X^*)$

$$\sum_{k=1}^r \alpha_k \frac{\partial F(X^*, Y_k)}{\partial X} \in L(X^*)$$

(напомним, что $Y_k \in R(X^*)$, $k \in [1 : r]$). Следовательно, выполняется соотношение (3.1). Остается сослаться на теоремы 3.1 и 2.2.

Теорема доказана.

З а м е ч а н и е. Если $\Gamma(X^*) = E_n$, то (3.5) эквивалентно следующему равенству:

$$\sum_{k=1}^r \alpha_k \frac{\partial F(X^*, Y_k)}{\partial X} = 0. \quad (3.6)$$

Действительно, в этом случае (3.1) равносильно (3.4), а (3.4) в силу леммы 1.1 Приложения II — равенству (3.6).

Теорема 3.3. Если функция $F(X, Y)$ выпукла по X на Ω при каждом фиксированном $Y \in G$ и существует точка $X^* \in \Omega$, для которой

$$\inf_{X \in \Omega} \max_{Y \in G} F(X, Y) = \max_{Y \in G} F(X^*, Y), \quad (3.7)$$

то найдется конечное множество $G_r = \{Y_k \in G \mid k \in [1:r]\}$, $1 \leq r \leq n+1$, такое, что

$$\inf_{X \in \Omega} \max_{Y \in G} F(X, Y) = \inf_{X \in \Omega} \max_{Y \in G_r} F(X, Y),$$

причем

$$\inf_{X \in \Omega} \max_{Y \in G_r} F(X, Y) = \max_{Y \in G_r} F(X^*, Y).$$

Доказательство. В силу (3.7) и теоремы 3.2 найдутся точки $Y_1, \dots, Y_r$ из $R(X^*)$ такие, что будет выполняться равенство (3.5). Докажем, что $G_r \stackrel{df}{=} \{Y_k \mid k \in [1:r]\}$ может быть взято в качестве искомого множества.

Действительно, в силу выпуклости на Ω функции

$$\max_{Y \in G_r} F(X, Y)$$

и справедливости (3.5) получим на основании теоремы 3.2 (в части достаточности)

$$\inf_{X \in \Omega} \max_{Y \in G_r} F(X, Y) = \max_{Y \in G_r} F(X^*, Y). \quad (3.8)$$

Остается заметить, что точки множества G_r взяты из $R(X^*)$, поэтому

$$\max_{Y \in G} F(X^*, Y) = \max_{Y \in G_r} F(X^*, Y). \quad (3.9)$$

Объединяя (3.7), (3.8) и (3.9), получим требуемое. Теорема доказана.

3*. Рассмотрим случай, когда множество Ω задано специальным образом.

Пусть

$$\Omega = \{X \in E_n \mid h(X, \beta) \leq 0 \text{ для всех } \beta \in \omega\},$$

где $\omega \subset E_s$ — ограниченное замкнутое множество, а функция $h(X, \beta)$ непрерывна вместе с $\partial h(X, \beta)/\partial X$ по совокупности переменных на $E_n \times \omega$. Предположим, кроме того, что функция $h(X, \beta)$ выпукла по X при каждом фиксированном $\beta \in \omega$ и что существует точка $\bar{X} \in E_n$, для которой

$$\max_{\beta \in \omega} h(\bar{X}, \beta) < 0. \quad (3.10)$$

Условие (3.10) будем называть условием Слейтера. В силу сделанных предположений множество Ω является выпуклым и замкнутым множеством. По (3.10) оно имеет внутренние точки.

Пусть $X_0 \in \Omega$. Рассмотрим множество

$$\Gamma(X_0) = \{V \in E_n \mid \max_{\beta \in \omega} h(X_0 + \alpha V, \beta) \leq 0\}$$

для всех $\alpha \in [0, \alpha_0(V)]$, $\alpha_0(V) > 0$.

Через $\bar{\Gamma}(X_0)$ обозначим замыкание множества $\Gamma(X_0)$.

Очевидно, $\bar{\Gamma}(X_0)$ является конусом возможных направлений множества Ω в точке X_0 (см. п. 1).

Введем обозначение

$$B(X_0) = \left\{ V \in E_n \mid \left(V, \frac{\partial h(X_0, \beta)}{\partial X} \right) \leq 0 \text{ для всех } \beta \in Q(X_0) \right\},$$

где $Q(X_0) = \{\beta \in \omega \mid h(X_0, \beta) = 0\}$. Если $Q(X_0) = \emptyset$, то полагаем $B(X_0) = E_n$.

Лемма 3.1. Справедливо равенство

$$\bar{\Gamma}(X_0) = B(X_0).$$

Доказательство. Если $Q(X_0) = \emptyset$, то утверждение леммы тривиально. Действительно, в этом случае X_0 — внутренняя точка множества Ω , поэтому

$$\Gamma(X_0) = \bar{\Gamma}(X_0) = E_n.$$

С другой стороны, по определению $B(X_0) = E_n$. Таким образом, $\bar{\Gamma}(X_0) = B(X_0)$.

В дальнейшем будем считать, что $Q(X_0) \neq \emptyset$. Заметим, что в этом случае

$$\max_{\beta \in \omega} h(X_0, \beta) = 0.$$

Докажем вначале включение

$$\bar{\Gamma}(X_0) \subset B(X_0).$$

Пусть $V_0 \in \bar{\Gamma}(X_0)$. Это означает, что существует такая последовательность векторов $\{V_i\}$, что

$$V_0 = \lim_{i \rightarrow \infty} V_i; \quad V_i \in \Gamma(X_0), \quad i = 1, 2, \dots$$

Зафиксируем любое i . Для всех $\beta \in \omega$ и всех $\alpha \in (0, \alpha_0(V_i))$ по определению $\Gamma(X_0)$ будет

$$h(X_0 + \alpha V_i, \beta) \leq 0.$$

Отсюда имеем

$$h(X_0 + \alpha V_i, \beta) = h(X_0, \beta) + \alpha \left(\frac{\partial h(X_0 + \tau_i V_i, \beta)}{\partial X}, V_i \right) \leq 0,$$

где $\tau_i \stackrel{\text{df}}{=} \tau_i(\alpha) \in (0, \alpha)$.

Если $\beta \in Q(X_0)$, то $h(X_0, \beta) = 0$, и потому

$$\left(\frac{\partial h(X_0 + \tau_i(\alpha) V_i, \beta)}{\partial X}, V_i \right) \leq 0. \quad (3.11)$$

Неравенство (3.11) справедливо для всех $\alpha \in (0, \alpha_0(V_i))$. Устремляя α к нулю, получим

$$\left(\frac{\partial h(X_0, \beta)}{\partial X}, V_i \right) \leq 0. \quad (3.12)$$

В неравенстве (3.12) перейдем к пределу при $i \rightarrow \infty$. Это дает для всех $\beta \in Q(X_0)$

$$\left(\frac{\partial h(X_0, \beta)}{\partial X}, V_0 \right) \leq 0.$$

Таким образом, $V_0 \in B(X_0)$, и включение $\bar{\Gamma}(X_0) \subset B(X_0)$ доказано.

Докажем обратное включение

$$B(X_0) \subset \bar{\Gamma}(X_0).$$

Пусть $V_0 \in B(X_0)$. Это означает, что

$$\left(V_0, \frac{\partial h(X_0, \beta)}{\partial X} \right) \leq 0 \quad \text{для всех } \beta \in Q(X_0). \quad (3.13)$$

Прежде всего установим, что по любому $\varepsilon > 0$ найдется такое $\bar{\varepsilon} > 0$, что для всех $\varepsilon \in (0, \bar{\varepsilon}]$ будет

$$\max_{\beta \in Q_\varepsilon(X_0)} \left(V_0, \frac{\partial h(X_0, \beta)}{\partial X} \right) \leq \varepsilon, \quad (3.14)$$

где $Q_\varepsilon(X_0) = \{\beta \in \omega \mid -\varepsilon \leq h(X_0, \beta) \leq 0\}$.

Действительно, допустим противное. Тогда для некоторого $\bar{\alpha} > 0$ и некоторой последовательности

положительных чисел $\{\varepsilon_v\}$, стремящейся к нулю, будет выполняться неравенство

$$\max_{\beta \in Q_{\varepsilon_v}(X_0)} \left(V_0, \frac{\partial h(X_0, \beta)}{\partial X} \right) \geq \bar{a} > 0, \quad v = 1, 2, \dots \quad (3.15)$$

Очевидно, максимум в (3.15) достигается. Значит, при каждом v найдется точка $\beta_v \in Q_{\varepsilon_v}(X_0)$, для которой

$$\left(V_0, \frac{\partial h(X_0, \beta_v)}{\partial X} \right) \geq \bar{a}. \quad (3.16)$$

При этом

$$-\varepsilon_v \leq h(X_0, \beta_v) \leq 0. \quad (3.17)$$

Поскольку все точки β_v , $v = 1, 2, \dots$, принадлежат ω , то из последовательности $\{\beta_v\}$ можно выделить сходящуюся подпоследовательность. Не ограничивая общности, будем считать, что вся последовательность $\{\beta_v\}$ сходится к точке $\beta^* \in \omega$. Переходя в (3.16) и (3.17) к пределу при $v \rightarrow \infty$, получим

$$\left(V_0, \frac{\partial h(X_0, \beta^*)}{\partial X} \right) \geq \bar{a},$$

причем $h(X_0, \beta^*) = 0$, т. е. $\beta^* \in Q(X_0)$.

Тогда тем более

$$\max_{\beta \in Q(X_0)} \left(V_0, \frac{\partial h(X_0, \beta)}{\partial X} \right) \geq \bar{a},$$

что противоречит (3.13). Итак, доказано, что по любому $a > 0$ найдется такое $\bar{\varepsilon} > 0$, что для всех $\varepsilon \in (0, \bar{\varepsilon}]$ будет выполняться неравенство (3.14).

Положим $W_0 = \bar{X} - X_0$, где $\bar{X}$ — точка, фигурирующая в (3.10), и покажем, что при любом $\gamma > 0$ точка $V_0(\gamma) \stackrel{df}{=} V_0 + \gamma W_0$ принадлежат $\Gamma(X_0)$, т. е. что для всех $\alpha \in [0, \alpha_0(\gamma)]$, $\alpha_0(\gamma) > 0$,

$$\max_{\beta \in \omega} h(X_0 + \alpha V_0(\gamma), \beta) \leq 0.$$

Зафиксируем $\gamma > 0$. Введем обозначения

$$\varphi_1(X) = \max_{\beta \in \omega} h(X, \beta),$$

$$a = -\frac{1}{\gamma} \varphi_1(\bar{X}) > 0.$$

Как уже отмечалось, по $a > 0$ найдется такое $\bar{\varepsilon} > 0$, что для $\varepsilon \in (0, \bar{\varepsilon}]$ будет

$$\max_{\beta \in Q_\varepsilon(X_0)} \left(\frac{\partial h(X_0, \beta)}{\partial X}, V_0 \right) \leq a.$$

Далее заметим (см. Приложение II, § 3), что в силу выпуклости $h(X, \beta)$ по X

$$\left(\frac{\partial h(X_0, \beta)}{\partial X}, \bar{X} - X_0 \right) \leq h(\bar{X}, \beta) - h(X_0, \beta). \quad (3.18)$$

Если $\varepsilon \in (0, -\frac{1}{2} \varphi_1(\bar{X})]$, то для $\beta \in Q_\varepsilon(X_0)$ получим на основании (3.18)

$$\left(\frac{\partial h(X_0, \beta)}{\partial X}, W_0 \right) \leq \varphi_1(\bar{X}) + \varepsilon \leq \frac{1}{2} \varphi_1(\bar{X}) = -\frac{4a}{\gamma}.$$

Зафиксируем некоторое ε , $0 < \varepsilon < \min \left\{ \bar{\varepsilon}, -\frac{1}{2} \varphi_1(\bar{X}) \right\}$. Тогда будем иметь

$$\begin{aligned} \max_{\beta \in Q_\varepsilon(X_0)} \left(\frac{\partial h(X_0, \beta)}{\partial X}, V_0 \right) &\leq a, \\ \gamma \max_{\beta \in Q_\varepsilon(X_0)} \left(\frac{\partial h(X_0, \beta)}{\partial X}, W_0 \right) &\leq -4a. \end{aligned}$$

откуда следует, что

$$\max_{\beta \in Q_\varepsilon(X_0)} \left(\frac{\partial h(X_0, \beta)}{\partial X}, V_0(\gamma) \right) \leq -3a. \quad (3.19)$$

Далее

$$\begin{aligned} h(X_0 + \alpha V_0(\gamma), \beta) &= \\ &= h(X_0, \beta) + \alpha \left(\frac{\partial h(X_0 + \tau V_0(\gamma), \beta)}{\partial X}, V_0(\gamma) \right) \leq \\ &\leq \alpha \left(\frac{\partial h(X_0 + \tau V_0(\gamma), \beta)}{\partial X}, V_0(\gamma) \right), \end{aligned} \quad (3.20)$$

где $\tau \in (0, \alpha)$. В силу (3.19), (3.20) и непрерывности $\partial h(X, \beta)/\partial X$ по совокупности переменных на $E_n \times \omega$, получим, что для достаточно малых $\alpha \in [0, \alpha_0]$, $\alpha_0 > 0$, и всех $\beta \in Q_\varepsilon(X_0)$

$$h(X_0 + \alpha V_0(\gamma), \beta) \leq -2\alpha a \leq 0. \quad (3.21)$$

Пусть теперь $\beta \notin Q_\varepsilon(X_0)$. Тогда $h(X_0, \beta) < -\varepsilon$. Нетрудно понять, что найдется такое $\alpha_0(\gamma)$, $0 < \alpha_0(\gamma) \leq \alpha_0$, что для $\alpha \in [0, \alpha_0(\gamma)]$ и $\beta \notin Q_\varepsilon(X_0)$ будет

$$h(X_0 + \alpha V_0(\gamma), \beta) \leq -\frac{1}{2} \varepsilon. \quad (3.22)$$

Объединяя (3.21) и (3.22), заключаем, что

$$\max_{\beta \in \omega} h(X_0 + \alpha V_0(\gamma), \beta) \leq 0$$

при всех $\alpha \in [0, \alpha_0(\gamma)]$. Но это и означает, что $V_0(\gamma) \in \Gamma(X_0)$.

Поскольку последнее включение имеет место для любого $\gamma > 0$, то устремляя γ к нулю, получим: $V_0 \in \bar{\Gamma}(X_0)$.

Итак, установлено, что

$$B(X_0) \subset \bar{\Gamma}(X_0).$$

Лемма доказана полностью.

З а м е ч а н и е. Условие Слейтера (3.10) существенно для справедливости леммы. Это видно из следующего примера.

Пусть (рис. 33)

$$\Omega = \{X = (x_1, x_2) \mid (x_1 - 1)^2 + x_2^2 - 1 \leq 0,$$

$$(x_1 - 4)^2 + x_2^2 - 4 \leq 0\}.$$

В этом случае $\omega = \{1, 2\}$,

$$h(X, 1) = (x_1 - 1)^2 + x_2^2 - 1,$$

$$h(X, 2) = (x_1 - 4)^2 + x_2^2 - 4.$$

Множество Ω состоит из единственной точки $X_0 = (2, 0)$. Поэтому $\Gamma(X_0) = \bar{\Gamma}(X_0) = \{0\}$. Найдем множество $B(X_0)$. Поскольку

$$Q(X_0) = \{1, 2\} \text{ и } \frac{\partial h(X_0, 1)}{\partial X} = (2, 0), \quad \frac{\partial h(X_0, 2)}{\partial X} = (-4, 0),$$

то

$$B(X_0) = \{V = (v_1, v_2) \mid v_1 = 0\}.$$

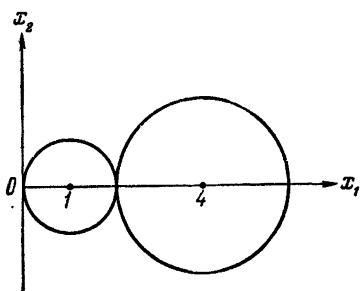


Рис. 33.

Таким образом, $\bar{\Gamma}(X_0) \neq B(X_0)$. Причина здесь кроется в том, что множество Ω не имеет внутренних точек.

4. В предыдущем пункте было установлено, что при выполнении условия Слейтера

$$\bar{\Gamma}(X_0) = \begin{cases} \left\{ V \mid \left(V, \frac{\partial h(X_0, \beta)}{\partial X} \right) \leq 0 \text{ для всех } \beta \in Q(X_0) \right\}, & \text{если } Q(X_0) \neq \emptyset, \\ E_n, & \text{если } Q(X_0) = \emptyset. \end{cases}$$

Обозначим через $K(X_0)$ выпуклую каноническую оболочку множества $H'(X_0) = \left\{ -\frac{\partial h(X_0, \beta)}{\partial X} \mid \beta \in Q(X_0) \right\}$:

$$K(X_0) = \left\{ V = -\sum_{i=1}^r \lambda_i \frac{\partial h(X_0, \beta_i)}{\partial X} \mid \lambda_i \geq 0, \beta_i \in Q(X_0), \right. \\ \left. r - \text{произвольное натуральное число} \right\}.$$

Если $Q(X_0) = \emptyset$, то по определению полагаем $K(X_0) = \{0\}$. Нетрудно проверить (см. лемму 2.5 Приложения II), что $K(X_0)$ — замкнутый выпуклый конус.

Лемма 3.2. При выполнении условия Слейтера справедливо равенство

$$\Gamma^+(X_0) = K(X_0).$$

Доказательство. Если $Q(X_0) = \emptyset$, то утверждение леммы тривиально. Поэтому будем считать, что $Q(X_0) \neq \emptyset$. Докажем равенство

$$B(X_0) = K^+(X_0). \quad (3.23)$$

Пусть $V_0 \in B(X_0)$. В этом случае

$$\left(-\frac{\partial h(X_0, \beta)}{\partial X}, V_0 \right) \geq 0 \text{ для всех } \beta \in Q(X_0).$$

Отсюда очевидным образом следует, что для всех $Z \in K(X_0)$

$$(Z, V_0) \geq 0,$$

т. е. $V_0 \in K^+(X_0)$. Значит,

$$B(X_0) \subset K^+(X_0). \quad (3.24)$$

Обратно, пусть $V_0 \in K^+(X_0)$. Тогда для всех $Z \in K(X_0)$

$$(Z, V_0) \geq 0.$$

В частности, для всех $\beta \in Q(X_0)$

$$\left(-\frac{\partial h(X_0, \beta)}{\partial X}, V_0\right) \geq 0,$$

или, что то же самое,

$$\left(\frac{\partial h(X_0, \beta)}{\partial X}, V_0\right) \leq 0.$$

Но это означает, что $V_0 \in B(X_0)$. Таким образом,

$$K^+(X_0) \subset B(X_0). \quad (3.25)$$

Объединяя (3.24) и (3.25), получим (3.23). В равенстве (3.23) перейдем к сопряженным конусам. Это дает

$$\Gamma^+(X_0) = K(X_0),$$

ибо $B(X_0) = \bar{\Gamma}(X_0)$ и $K^{++}(X_0) = K(X_0)$.

Лемма доказана.

Теорема 3.4. Пусть

$$\Omega = \{X \in E_n \mid h(X, \beta) \leq 0 \text{ для всех } \beta \in \omega\},$$

и выполнено условие Слейтера (3.10). Для того чтобы X^* была точкой минимума функции $\varphi(X)$ на Ω , необходимо, (а в случае выпуклости $\varphi(X)$ на Ω и достаточно), чтобы нашлись точки $Y_1, \dots, Y_r$ из $R(X^*)$ и $\beta_1, \dots, \beta_{r_1}$ из $Q(X^*)$ и неотрицательные числа

$$\lambda_{11}, \dots, \lambda_{1r}; \lambda_{21}, \dots, \lambda_{2r_1}; \sum_{i=1}^r \lambda_{1i} = 1,$$

где $1 \leq r \leq n+1$, $1 \leq r_1 \leq n$, такие, что

$$\sum_{i=1}^r \lambda_{1i} \frac{\partial F(X^*, Y_i)}{\partial X} + \sum_{i=1}^{r_1} \lambda_{2i} \frac{\partial h(X^*, \beta_i)}{\partial X} = 0. \quad (3.26)$$

Доказательство. В силу теорем 2.2 и 3.1 достаточно доказать, что (3.26) равносильно соотношению (3.1). Покажем вначале, что (3.26) следует из (3.1). Пусть $Z^* \in L(X^*) \cap \Gamma^+(X^*)$. Равенство (3.26) получится, если воспользоваться леммами 1.1 и 2.4 Приложения II и леммой 3.2. То, что (3.1) следует из (3.26), очевидно.

Теорема доказана.

§ 4. О сходимости сеточного метода

1. Определение. Точка $X^* \in \Omega$, для которой выполняется равенство

$$\inf_{Z \in \Omega} \max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, Z - X^* \right) = 0,$$

называется *стационарной точкой функции $\varphi(X)$ на множестве Ω* .

Для нахождения стационарных точек может быть использован сеточный метод, сущность которого заключается в следующем.

На множестве G при натуральных N вводятся *конечные сетки* $G_N = \{Y_{Ni} \in G | i \in [0: N]\}$. Предполагается выполненным следующее условие на поведение сеток G_N при $N \rightarrow \infty$.

Для любого $\varepsilon > 0$ можно указать такой номер N_0 , что для $N > N_0$ расстояние между произвольной точкой $Y \in G$ и ближайшей к Y точкой сетки G_N будет меньше ε .

Последовательность таких сеток $\{G_N\}$ будем называть *всюду плотной в G* .

Пусть

$$f_{Ni}(X) = F(X, Y_{Ni}), \quad i \in [0: N];$$

$$\varphi_N(X) \stackrel{\text{df}}{=} \max_{i \in [0: N]} f_{Ni}(X) = \max_{Y \in G_N} F(X, Y).$$

Будем считать, что при достаточно больших N можно найти хотя бы одну стационарную точку $X_N \in \Omega$ функции $\varphi_N(X)$ на множестве Ω .

Таким образом, при достаточно больших N выполняется равенство

$$\inf_{Z \in \Omega} \max_{i \in \tilde{R}(X_N)} \left(\frac{\partial f_{Ni}(X_N)}{\partial X}, Z - X_N \right) = 0, \quad (4.1)$$

где

$$\tilde{R}(X) = \{i \in [0: N] | f_{Ni}(X) = \varphi_N(X)\}.$$

При сделанных предположениях справедлива

Теорема 4.1. *Всякая предельная точка последовательности $\{X_N\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω .*

Доказательство. Пусть для определенности вся последовательность $\{X_n\}$ сходится к точке $X^* \in \Omega$, однако X^* не является стационарной точкой функции $\varphi(X)$ на множестве Ω . Тогда существует такая точка $\bar{Z} \in \Omega$, что

$$\max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, \bar{Z} - X^* \right) < 0. \quad (4.2)$$

Докажем, что найдется число $\bar{\varepsilon} > 0$, при котором

$$\max_{Y \in R_{\bar{\varepsilon}}(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, \bar{Z} - X^* \right) \stackrel{df}{=} -r_{\bar{\varepsilon}} < 0, \quad (4.3)$$

где

$$R_{\varepsilon}(X) = \{Y \in G \mid \varphi(X) - F(X, Y) \leq \varepsilon\}.$$

Если бы это было не так, то для любой последовательности $\{\varepsilon_v\}$, $\varepsilon_v > 0$, монотонно стремящейся к нулю, выполнялись бы неравенства

$$\max_{Y \in R_{\varepsilon_v}(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, \bar{Z} - X^* \right) \geq 0. \quad (4.4)$$

Учитывая, что $R_{\varepsilon_v}(X^*)$ — ограниченные замкнутые множества в E_m и что, значит, максимумы в (4.4) достигаются, укажем $Y_v \in R_{\varepsilon_v}(X^*)$ такие, что

$$\left(\frac{\partial F(X^*, Y_v)}{\partial X}, \bar{Z} - X^* \right) \geq 0. \quad (4.5)$$

При этом

$$\varphi(X^*) - F(X^*, Y_v) \leq \varepsilon_v. \quad (4.6)$$

Далее $Y_v \in G$, и поэтому из последовательности $\{Y_v\}$ можно выделить сходящуюся подпоследовательность. Не ограничивая общности, будем считать, что вся последовательность $\{Y_v\}$ сходится к $Y^* \in G$. Тогда, с одной стороны, в силу (4.6) $Y^* \in R(X^*)$, а с другой — в силу (4.5)

$$\left(\frac{\partial F(X^*, Y^*)}{\partial X}, \bar{Z} - X^* \right) \geq 0.$$

Значит, и подаловно

$$\max_{Y \in R(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, \bar{Z} - X^* \right) \geq 0,$$

что противоречит (4.2).

Итак, доказано существование $\bar{\varepsilon} > 0$, для которого выполняется неравенство (4.3).

По $\bar{\varepsilon} > 0$ выберем такое $\delta > 0$, что из неравенства $\|Y' - Y''\| < \delta$ будет следовать

$$|F(X^*, Y') - F(X^*, Y'')| < \frac{\bar{\varepsilon}}{3}.$$

Далее, N_0 возьмем столь большим, чтобы при $N > N_0$ расстояние между произвольным $Y \in G$ и ближайшим к Y узлом сетки G_N было меньше δ и чтобы

$$\max_{Y \in G} |F(X_N, Y) - F(X^*, Y)| < \frac{\bar{\varepsilon}}{3}.$$

Докажем, что в этом случае при всех $i \in \tilde{R}(X_N)$

$$Y_{Ni} \in R_{\bar{\varepsilon}}(X^*). \quad (4.7)$$

Действительно, пусть $i \in \tilde{R}(X_N)$. Тогда

$$\begin{aligned} F(X_N, Y_{Ni}) &\stackrel{df}{=} f_{Ni}(X_N) = \max_{k \in [0:N]} f_{Nk}(X_N) = \\ &= \max_{k \in [0:N]} F(X_N, Y_{Nk}) = \max_{Y \in G_N} F(X_N, Y). \end{aligned}$$

Далее,

$$\begin{aligned} \max_{Y \in G} F(X^*, Y) &= F(X^*, \bar{Y}) \leq \\ &\leq \max_{Y \in G_N} F(X^*, Y) + \frac{\bar{\varepsilon}}{3} \leq \max_{Y \in G_N} F(X_N, Y) + \frac{2\bar{\varepsilon}}{3} = \\ &= F(X_N, Y_{Ni}) + \frac{2\bar{\varepsilon}}{3} \leq F(X^*, Y_{Ni}) + \bar{\varepsilon}. \end{aligned}$$

Таким образом, $Y_{Ni} \in R_{\bar{\varepsilon}}(X^*)$, и включение (4.7) для $N > N_0$ доказано.

Теперь выберем столь большое $\bar{N} > N_0$, чтобы выполнялись неравенства

$$\begin{aligned} \max_{Y \in G} \left| \left(\frac{\partial F(X^*, Y)}{\partial X}, X_{\bar{N}} - X^* \right) \right| &< \frac{r_{\bar{\varepsilon}}}{3}, \\ \max_{Y \in G} \left| \left(\frac{\partial F(X^*, Y)}{\partial X} - \frac{\partial F(X_{\bar{N}}, Y)}{\partial X}, \bar{Z} - X_{\bar{N}} \right) \right| &< \frac{r_{\bar{\varepsilon}}}{3}. \end{aligned}$$

Получим в силу (4.7) и (4.3)

$$\begin{aligned}
 \max_{i \in \tilde{R}(X_N)} \left(\frac{\partial f_{\bar{N}i}(X_N)}{\partial X}, \bar{Z} - X_N \right) &\leq \\
 &\leq \max_{Y \in R_{\bar{\varepsilon}}(X^*)} \left(\frac{\partial F(X_N, Y)}{\partial X}, \bar{Z} - X_N \right) \leq \\
 &\leq \max_{Y \in R_{\bar{\varepsilon}}(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, \bar{Z} - X_N \right) + \frac{r_{\bar{\varepsilon}}}{3} \leq \\
 &\leq \max_{Y \in R_{\bar{\varepsilon}}(X^*)} \left(\frac{\partial F(X^*, Y)}{\partial X}, \bar{Z} - X^* \right) + \frac{2r_{\bar{\varepsilon}}}{3} = -\frac{r_{\bar{\varepsilon}}}{3} < 0,
 \end{aligned}$$

что противоречит (4.1)

Теорема доказана.

З а м е ч а н и е. Точки X_N являются стационарными точками в дискретных минимаксных задачах. Для их нахождения могут быть использованы методы, развитые в главах III—V.

2*. Рассмотрим отдельно случай $\Omega = E_n$. Пусть по-прежнему $\{G_N\}$ — последовательность сеток, всюду плотная в G при $N \rightarrow \infty$ (см. п. 1). Введем обозначения:

$$\begin{aligned}
 R_{\varepsilon}(X) &= \{Y \in G \mid \varphi(X) - F(X, Y) \leq \varepsilon\}, \\
 R_{N\varepsilon}(X) &= \{Y \in G_N \mid \varphi_N(X) - F(X, Y) \leq \varepsilon\}, \\
 \chi_{\varepsilon}(X, g) &= \max_{Y \in R_{\varepsilon}(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right), \\
 \chi_{N\varepsilon}(X, g) &= \max_{Y \in R_{N\varepsilon}(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right), \\
 \psi_{\varepsilon}(X) &= \min_{\|g\|=1} \chi_{\varepsilon}(X, g), \\
 \psi_{N\varepsilon}(X) &= \min_{\|g\|=1} \chi_{N\varepsilon}(X, g), \\
 D(X) &= \inf_{\varepsilon \in [0, \bar{\varepsilon}]} \varepsilon \psi_{\varepsilon}(X), \\
 D_N(X) &= \inf_{\varepsilon \in [0, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X),
 \end{aligned}$$

где $\bar{\varepsilon} > 0$ — любое фиксированное число.

Нашей ближайшей целью является доказательство следующей теоремы.

Теорема 4.2. При $N \rightarrow \infty$ имеет место предельное соотношение

$$D_N(X) \rightarrow D(X),$$

причем сходимость равномерная по $X \in \bar{S}$, где $\bar{S} \subset E_n$ — произвольное ограниченное замкнутое множество.

Предварительно докажем одну лемму.

Лемма 4.1. Зафиксируем $\varepsilon_0 > 0$. Для любых $\rho > 0$, $\delta > 0$ ($\delta \leq \varepsilon_0$) и $\varepsilon \geq \varepsilon_0$ найдется такое N_0 , что при $N > N_0$ для всех $X \in \bar{S}$ будет

$$\psi_{N, \varepsilon - \delta}(X) \leq \psi_\varepsilon(X) \leq \psi_{N, \varepsilon + \delta}(X) + \rho. \quad (4.8)$$

Доказательство. Заметим, во-первых, что если $\delta \leq \varepsilon_0$ и $\varepsilon \geq \varepsilon_0$, то $\varepsilon - \delta \geq 0$.

Нетрудно показать, что

$$\varphi_N(X) \xrightarrow{N \rightarrow \infty} \varphi(X)$$

равномерно по $X \in \bar{S}$. Поэтому найдется столь большое N_1 , что при $N > N_1$ для всех $X \in \bar{S}$ будет выполняться неравенство

$$\varphi(X) - \varphi_N(X) \leq \delta. \quad (4.9)$$

Зафиксируем теперь $X \in \bar{S}$ и $g \in E_n$, $\|g\| = 1$, и пусть

$$\chi_{N, \varepsilon - \delta}(X, g) = \left(\frac{\partial F(X, \bar{Y}_N)}{\partial X}, g \right),$$

где $\bar{Y}_N \stackrel{\text{def}}{=} \bar{Y}_N(X, g) \in R_{N, \varepsilon - \delta}(X)$. Тогда

$$\varphi_N(X) - F(X, \bar{Y}_N) \leq \varepsilon - \delta. \quad (4.10)$$

Учитывая (4.9) и (4.10), получим для $N > N_1$

$$\begin{aligned} \varphi(X) - F(X, \bar{Y}_N) &= (\varphi(X) - \varphi_N(X)) + \\ &+ (\varphi_N(X) - F(X, \bar{Y}_N)) \leq \varepsilon. \end{aligned}$$

Значит, $\bar{Y}_N \in R_\varepsilon(X)$. Теперь имеем

$$\begin{aligned} \chi_{N, \varepsilon - \delta}(X, g) &= \\ &= \left(\frac{\partial F(X, \bar{Y}_N)}{\partial X}, g \right) \leq \max_{Y \in R_\varepsilon(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) = \chi_\varepsilon(X, g). \end{aligned}$$

Поскольку последнее неравенство справедливо для всех g , $\|g\| = 1$, то

$$\psi_{N, \varepsilon - \delta}(X) \leq \psi_\varepsilon(X).$$

Получили, что для $N > N_1$ и $X \in \bar{S}$ выполняется левое неравенство в (4.8).

Переходим к доказательству правого неравенства. В силу замкнутости и ограниченности множества G по $\rho > 0$ и $\delta > 0$ можно указать такое $\gamma > 0$, что при $\|Y' - Y''\| \leq \gamma$, $Y', Y'' \in G$, равномерно по $X \in \bar{S}$ будут выполняться неравенства

$$|F(X, Y') - F(X, Y'')| \leq \delta, \quad (4.11)$$

$$\left\| \frac{\partial F(X, Y')}{\partial X} - \frac{\partial F(X, Y'')}{\partial X} \right\| \leq \rho. \quad (4.12)$$

Выберем теперь столь большое N_2 , чтобы при $N > N_2$ расстояние между произвольным $Y \in G$ и ближайшим к Y узлом сетки G_N было меньше γ . Зафиксируем $X \in \bar{S}$ и $g \in E_n$, $\|g\| = 1$, и пусть

$$\chi_e(X, g) = \left(\frac{\partial F(X, \tilde{Y})}{\partial X}, g \right),$$

где $\tilde{Y} \stackrel{\text{def}}{=} \tilde{Y}(X, g) \in R_e(X)$. Тогда

$$\varphi(X) - F(X, \tilde{Y}) \leq \varepsilon. \quad (4.13)$$

Обозначим через $Y'_N \in G_N$ ближайший к $\tilde{Y}$ узел из G_N . При $N > N_2$ получим в силу (4.11) и (4.13)

$$\begin{aligned} \varphi_N(X) - F(X, Y'_N) &\leq \varphi(X) - F(X, Y'_N) = \\ &= (\varphi(X) - F(X, \tilde{Y})) + (F(X, \tilde{Y}) - F(X, Y'_N)) \leq \varepsilon + \delta. \end{aligned}$$

Значит, $Y'_N \in R_{N, \varepsilon + \delta}(X)$. Отсюда и из (4.12) следует, что

$$\begin{aligned} \chi_e(X, g) &= \left(\frac{\partial F(X, \tilde{Y})}{\partial X}, g \right) = \\ &= \left(\frac{\partial F(X, Y'_N)}{\partial X}, g \right) + \left(\frac{\partial F(X, \tilde{Y})}{\partial X} - \frac{\partial F(X, Y'_N)}{\partial X}, g \right) \leq \\ &\leq \max_{Y \in R_{N, \varepsilon + \delta}(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) + \rho = \chi_{N, \varepsilon + \delta}(X, g) + \rho. \end{aligned}$$

Последнее неравенство справедливо для всех g , $\|g\| = 1$, поэтому

$$\psi_e(X) \leq \psi_{N, \varepsilon + \delta}(X) + \rho.$$

Получили, что при $N > N_2$ равномерно по $X \in \bar{S}$ выполняется правое неравенство в (4.8). Оба неравенства в (4.8) будут выполняться при $N > N_0 \stackrel{\text{def}}{=} \max\{N_1, N_2\}$.

Лемма доказана.

Доказательство теоремы. Положим

$$c = \max_{X \in \bar{S}} \max_{Y \in G} \left\| \frac{\partial F(X, Y)}{\partial X} \right\|.$$

Ясно, что

$$|\psi_{N\varepsilon}(X)| \leq c, \quad |\psi_\varepsilon(X)| \leq c. \quad (4.14)$$

Возьмем произвольное $\alpha_0 > 0$. Числа $\varepsilon_0 > 0$, $\rho > 0$ и $\delta > 0$ выберем так, чтобы

$$4\varepsilon_0 < \bar{\varepsilon}, \quad 4\varepsilon_0 c \leq \alpha_0, \quad \delta < \varepsilon_0, \quad \delta c + \bar{\varepsilon}\rho < \varepsilon_0 c. \quad (4.15)$$

По лемме 4.1 найдется столь большое N_0 , что для всех $N > N_0$ и $\varepsilon \geq \varepsilon_0$ равномерно по $X \in \bar{S}$ будет

$$\psi_{N, \varepsilon - \delta}(X) \leq \psi_\varepsilon(X) \leq \psi_{N, \varepsilon + \delta}(X) + \rho.$$

Отсюда при $\varepsilon \geq \varepsilon_0$

$$\varepsilon \psi_{N, \varepsilon - \delta}(X) \leq \varepsilon \psi_\varepsilon(X) \leq \varepsilon \psi_{N, \varepsilon + \delta}(X) + \varepsilon \rho.$$

Преобразуем эти неравенства

$$\begin{aligned} (\varepsilon - \delta) \psi_{N, \varepsilon - \delta}(X) + \delta \psi_{N, \varepsilon - \delta}(X) &\leq \varepsilon \psi_\varepsilon(X) \leq \\ &\leq (\varepsilon + \delta) \psi_{N, \varepsilon + \delta}(X) - \delta \psi_{N, \varepsilon + \delta}(X) + \varepsilon \rho. \end{aligned}$$

Учитывая (4.14), получим для $\varepsilon_0 \leq \varepsilon \leq \bar{\varepsilon}$

$$\begin{aligned} (\varepsilon - \delta) \psi_{N, \varepsilon - \delta}(X) - \delta c &\leq \varepsilon \psi_\varepsilon(X) \leq \\ &\leq (\varepsilon + \delta) \psi_{N, \varepsilon + \delta}(X) + \delta c + \bar{\varepsilon}\rho. \end{aligned}$$

Значит,

$$\begin{aligned} \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} (\varepsilon - \delta) \psi_{N, \varepsilon - \delta}(X) - \delta c &\leq \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} \varepsilon \psi_\varepsilon(X) \leq \\ &\leq \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} (\varepsilon + \delta) \psi_{N, \varepsilon + \delta}(X) + \delta c + \bar{\varepsilon}\rho. \end{aligned} \quad (4.16)$$

Так как

$$\begin{aligned}
 D_N(X) &\stackrel{\text{df}}{=} \inf_{\varepsilon \in [0, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X) \leq \inf_{\varepsilon \in [\varepsilon_0 - \delta, \bar{\varepsilon} - \delta]} \varepsilon \psi_{N\varepsilon}(X), \\
 \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} (\varepsilon - \delta) \psi_{N, \varepsilon - \delta}(X) &= \inf_{\varepsilon \in [\varepsilon_0 - \delta, \bar{\varepsilon} - \delta]} \varepsilon \psi_{N\varepsilon}(X), \\
 \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} (\varepsilon + \delta) \psi_{N, \varepsilon + \delta}(X) &= \\
 &= \inf_{\varepsilon \in [\varepsilon_0 + \delta, \bar{\varepsilon} + \delta]} \varepsilon \psi_{N\varepsilon}(X) \leq \inf_{\varepsilon \in [\varepsilon_0 + \delta, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X),
 \end{aligned}$$

то из (4.16) и (4.15) получим

$$\begin{aligned}
 D_N(X) - \varepsilon_0 c &\leq \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X) \leq \\
 &\leq \inf_{\varepsilon \in [\varepsilon_0 + \delta, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X) + \varepsilon_0 c. \quad (4.17)
 \end{aligned}$$

Заметим (см. Приложение III, § 2), что

$$D(X) \stackrel{\text{df}}{=} \inf_{\varepsilon \in [0, \bar{\varepsilon}]} \varepsilon \psi_{\varepsilon}(X) = \min \left\{ \inf_{\varepsilon \in [0, \varepsilon_0]} \varepsilon \psi_{\varepsilon}(X), \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} \varepsilon \psi_{\varepsilon}(X) \right\}, \quad (4.18)$$

$$D_N(X) = \min \left\{ \inf_{\varepsilon \in [0, 2\varepsilon_0]} \varepsilon \psi_{N\varepsilon}(X), \inf_{\varepsilon \in [\varepsilon_0 + \delta, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X) \right\}. \quad (4.19)$$

Дальнейшие рассуждения будем проводить отдельно для случаев $D(X) \leq -3\varepsilon_0 c$ и $D(X) > -3\varepsilon_0 c$.

I. $D(X) \leq -3\varepsilon_0 c$.

Учитывая (4.18) и (4.14), запишем

$$D(X) = \inf_{\varepsilon \in [\varepsilon_0, \bar{\varepsilon}]} \varepsilon \psi_{\varepsilon}(X). \quad (4.20)$$

Из (4.17) и I теперь следует

$$D_N(X) \leq D(X) + \varepsilon_0 c \leq -2\varepsilon_0 c,$$

так что в силу (4.19)

$$D_N(X) = \inf_{\varepsilon \in [\varepsilon_0 + \delta, \bar{\varepsilon}]} \varepsilon \psi_{N\varepsilon}(X). \quad (4.21)$$

Объединяя (4.17), (4.20) и (4.21) и учитывая (4.15), получим для всех $N > N_0$

$$|D(X) - D_N(X)| \leq \varepsilon_0 c \leq \alpha_0. \quad (4.22)$$

II. $D(X) > -3\varepsilon_0 c$.

Докажем, что в этом случае при $N > N_0$

$$D_N(X) \geq -4\varepsilon_0 c. \quad (4.23)$$

Допустим противное. Пусть $D_N(X) < -4\varepsilon_0 c$. Тогда найдется такое $\varepsilon' \in [0, \bar{\varepsilon}]$, что

$$\varepsilon' \psi_{N\varepsilon'}(X) < -4\varepsilon_0 c.$$

Отсюда и из (4.14) следует

$$\varepsilon' > 4\varepsilon_0.$$

Поскольку, в частности, $\varepsilon' - \delta \geq \varepsilon_0$, то

$$(\varepsilon' - \delta) \psi_{\varepsilon' - \delta}(X) \leq (\varepsilon' - \delta) \psi_{N\varepsilon'}(X) + (\varepsilon' - \delta) \rho.$$

Усиливая это неравенство, получим

$$(\varepsilon' - \delta) \psi_{\varepsilon' - \delta}(X) \leq \varepsilon' \psi_{N\varepsilon'}(X) + \delta c + \bar{\varepsilon} \rho \leq -3\varepsilon_0 c.$$

Но тогда

$$D(X) \stackrel{df}{=} \inf_{\varepsilon \in [0, \bar{\varepsilon}]} \varepsilon \psi_{\varepsilon}(X) \leq (\varepsilon' - \delta) \psi_{\varepsilon' - \delta}(X) \leq -3\varepsilon_0 c,$$

что противоречит неравенству II. Итак, неравенство (4.23) установлено. Теперь имеем

$$\begin{aligned} -3\varepsilon_0 c &< D(x) \leq 0. \\ -4\varepsilon_0 c &\leq D_N(X) \leq 0. \end{aligned}$$

Отсюда и из (4.15) получаем для всех $N > N_0$

$$|D(X) - D_N(X)| \leq 4\varepsilon_0 c \leq \alpha_0. \quad (4.24)$$

Теорема доказана, ибо неравенства (4.22) и (4.24) справедливы для всех $N > N_0$ и $X \in \bar{S}$.

Следствие. Функция $D(X)$ непрерывна на E_n .

Это утверждение следует из непрерывности функций $D_N(X)$ на E_n (см. § 8 гл. III) и теоремы 4.2.

3*. Теорема 4.3. *Для того чтобы точка X^* была стационарной точкой функции $\varphi(X)$ на E_n , необходимо и достаточно, чтобы*

$$D(X^*) = 0.$$

Предварительно докажем одну лемму.

Лемма 4.2. *При фиксированном $X \in E_n$ функция $\psi_{\varepsilon}(X)$ как функция от ε , $0 \leq \varepsilon < \infty$, непрерывна справа в точке $\varepsilon = 0$.*

Доказательство. Пусть

$$\lambda(\varepsilon) = \psi_{\varepsilon}(X), \quad 0 \leq \varepsilon < \infty.$$

Очевидно, что функция $\lambda(\varepsilon)$ не убывает и ограничена,

$$|\lambda(\varepsilon)| \leq \max_{Y \in G} \left\| \frac{\partial F(X, Y)}{\partial X} \right\|.$$

Предположим, что $\lambda(\varepsilon)$ разрывна в точке $\varepsilon = 0$ справа. Тогда найдется последовательность положительных чисел $\{\varepsilon_\nu\}$, монотонно убывающая и стремящаяся к нулю, для которой

$$\lim_{\nu \rightarrow \infty} \lambda(\varepsilon_\nu) = \bar{\lambda} > \lambda(0). \quad (4.25)$$

В частности, для всех ν

$$\min_{\|g\|=1} \max_{Y \in R_{\varepsilon_\nu}(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \geq \bar{\lambda}. \quad (4.26)$$

Зафиксируем g , $\|g\| = 1$. В силу (4.26)

$$\max_{Y \in R_{\varepsilon_\nu}(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \geq \bar{\lambda}. \quad (4.27)$$

Поскольку $R_{\varepsilon_\nu}(X) \subset G$ — ограниченное замкнутое множество, то максимум в (4.27) достигается. Таким образом, при всех ν можно указать точку $Y_\nu \in R_{\varepsilon_\nu}(X)$, для которой

$$\left(\frac{\partial F(X, Y_\nu)}{\partial X}, g \right) \geq \bar{\lambda}. \quad (4.28)$$

При этом

$$0 \leq \varphi(X) - F(X, Y_\nu) \leq \varepsilon_\nu. \quad (4.29)$$

Не ограничивая общности, будем считать, что последовательность $\{Y_\nu\}$ сходится к точке $Y^* \in G$. Тогда в силу (4.29) $Y^* \in R(X)$. Переходя в неравенстве (4.28) к пределу при $\nu \rightarrow \infty$, получим

$$\left(\frac{\partial F(X, Y^*)}{\partial X}, g \right) \geq \bar{\lambda}.$$

Тем более

$$\max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \geq \bar{\lambda}. \quad (4.30)$$

Неравенство (4.30) справедливо для произвольного g , $\|g\| = 1$, поэтому

$$\min_{\|g\|=1} \max_{Y \in R(X)} \left(\frac{\partial F(X, Y)}{\partial X}, g \right) \geq \bar{\lambda}.$$

Получили, что $\lambda(0) \geq \bar{\lambda}$. Но это противоречит (4.25).

Лемма доказана.

Доказательство теоремы. Необходимость. Пусть X^* — стационарная точка функции $\varphi(X)$ на E_n . Тогда $\psi_0(X^*) \geq 0$. Тем более,

$$\psi_\varepsilon(X^*) \geq 0, \quad 0 \leq \varepsilon \leq \bar{\varepsilon}.$$

Отсюда получаем

$$D(X^*) \geq 0.$$

Обратное неравенство следует из определения $D(X)$. Значит,

$$D(X^*) = 0.$$

Достаточность. Пусть $D(X^*) = 0$. Тогда для всех $\varepsilon \in (0, \bar{\varepsilon}]$ будет

$$\psi_\varepsilon(X^*) \geq 0.$$

Учитывая лемму 4.2, заключаем, что $\psi_0(X^*) \geq 0$. Но это и означает, что X^* — стационарная точка функции $\varphi(X)$ на E_n .

Теорема доказана.

4*. Теперь мы имеем возможность описать еще один вариант сеточного метода для нахождения стационарных точек функции $\varphi(X)$ на E_n . Будем считать, что сетки G_N строятся таким образом, что $G_N \subset G_{N'}$ при $N' > N$.

Возьмем начальное приближение $X_0 \in E_n$. Предположим, что при некотором N_0 множество

$$\bar{M}(X_0) = \{X \in E_n \mid \varphi_{N_0}(X) \leq \varphi(X_0)\}$$

ограничено. Поскольку

$$M(X_0) \stackrel{\text{df}}{=} \{X \mid \varphi(X) \leq \varphi(X_0)\} \subset \bar{M}(X_0)$$

и при $N \geq N_0$

$$M_N(X_0) \stackrel{\text{df}}{=} \{X \mid \varphi_N(X) \leq \varphi_N(X_0)\} \subset \bar{M}(X_0),$$

то можно, в частности, утверждать, что существуют точки, в которых достигается минимум на E_n функций $\varphi(X)$ и $\varphi_N(X)$ при $N \geq N_0$.

Зафиксируем произвольным образом параметр $\varepsilon_0 > 0$ и рассмотрим функцию $\varphi_{N_0}(X)$. Исходя из начального

приближения X_0 и пользуясь D -методом, описанным в § 8 главы III, за конечное число шагов придем к точке X_1 , для которой

$$-D_{N_0}(X_1) \leq \varepsilon_0.$$

Поскольку $\varphi_{N_0}(X_1) \leq \varphi_{N_0}(X_0)$, то $X_1 \in \bar{M}(X_0)$. Положим теперь $N_1 = 2N_0$, $\varepsilon_1 = \varepsilon_0/2$ и рассмотрим функцию $\varphi_{N_1}(X)$. Пусть

$$X'_1 = \begin{cases} X_1, & \text{если } \varphi_{N_1}(X_1) \leq \varphi_{N_1}(X_0), \\ X_0 & - \text{в противном случае.} \end{cases}$$

Очевидно, $X'_1 \in M_{N_1}(X_0)$. Беря X'_1 в качестве начального приближения и пользуясь D -методом, за конечное число шагов придем к точке X_2 , для которой

$$-D_{N_1}(X_2) \leq \varepsilon_1.$$

Поскольку $\varphi_{N_1}(X_2) \leq \varphi_{N_1}(X'_1) \leq \varphi_{N_1}(X_0)$, то

$$X_2 \in M_{N_1}(X_0) \text{ и тем более } X_2 \in \bar{M}(X_0).$$

Пусть уже найдено k -е приближение $X_k \in \bar{M}(X_0)$. Положим $N_k = 2N_{k-1}$, $\varepsilon_k = \varepsilon_{k-1}/2$ и рассмотрим функцию $\varphi_{N_k}(X)$. Пусть

$$X'_k = \begin{cases} X_k, & \text{если } \varphi_{N_k}(X_k) \leq \varphi_{N_k}(X_0), \\ X_0 & - \text{в противном случае.} \end{cases}$$

Очевидно, $X'_k \in M_{N_k}(X_0)$. Беря X'_k в качестве начального приближения и пользуясь D -методом, за конечное число шагов придем к точке X_{k+1} , для которой

$$-D_{N_k}(X_{k+1}) \leq \varepsilon_k.$$

Поскольку $\varphi_{N_k}(X_{k+1}) \leq \varphi_{N_k}(X'_k) \leq \varphi_{N_k}(X_0)$, то $X_{k+1} \in M_{N_k}(X_0)$ и тем более $X_{k+1} \in \bar{M}(X_0)$.

Продолжая аналогично, получим последовательность $\{X_k\}$, причем $X_k \in \bar{M}(X_0)$ и

$$-D_{N_{k-1}}(X_k) \leq \varepsilon_{k-1}, \quad (4.31)$$

где $N_{k-1} = 2^{k-1}N_0$, $\varepsilon_{k-1} = \varepsilon_0/2^{k-1}$, $k = 1, 2, \dots$. Так как $\bar{M}(X_0)$ — ограниченное замкнутое множество, то после-

довательность $\{X_k\}$ имеет хотя бы одну предельную точку.

Теорема 4.4. *Любая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на E_n .*

Доказательство. Пусть

$$X_{k_j} \xrightarrow{k_j \rightarrow \infty} X^*.$$

Имеем

$$|D(X^*)| \leq |D(X^*) - D(X_{k_j})| + \\ + |D(X_{k_j}) - D_{N_{k_j-1}}(X_{k_j})| + |D_{N_{k_j-1}}(X_{k_j})|. \quad (4.32)$$

Первое слагаемое в правой части неравенства (4.32) стремится к нулю при $k_j \rightarrow \infty$ в силу непрерывности функции $D(X)$, второе — в силу ограниченности последовательности $\{X_{k_j}\}$ и теоремы 4.2, третье — по построению (см. (4.31)). Переходя в неравенстве (4.32) к пределу при $k_j \rightarrow \infty$, получим

$$D(X^*) = 0.$$

Остается сослаться на теорему 4.3.

Теорема доказана.

5. Как уже отмечалось, при каждом $X \in E_n$

$$\varphi_N(X) \xrightarrow{N \rightarrow \infty} \varphi(X).$$

В связи с этим интересно отметить, что при некоторых $X \in E_n$ и $g \in E_n$, $\|g\| = 1$, предельное соотношение

$$\frac{\partial \varphi_N(X)}{\partial g} \xrightarrow{N \rightarrow \infty} \frac{\partial \varphi(X)}{\partial g}$$

может не иметь места.

Пример. Пусть

$$F(x, y) = \cos\left(x\left(y + \frac{\pi}{4}\right)\right), \quad G = \left[0, \frac{3\pi}{2}\right]; \\ \varphi(x) = \max_{y \in \left[0, \frac{3\pi}{2}\right]} F(x, y).$$

Возьмем точку $x = 1$. В этом случае (рис. 34)

$$R(x) = \left\{0, \frac{3\pi}{2}\right\}, \quad \varphi(x) = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}.$$

По формуле (2.1)

$$\begin{aligned}\frac{\partial \varphi(x)}{\partial g} &= \max \left\{ -g \frac{\pi}{4} \sin \frac{\pi}{4}, -g \frac{7\pi}{4} \sin \frac{7\pi}{4} \right\} = \\ &= \max \left\{ -g \frac{\pi \sqrt{2}}{8}, g \frac{7\pi \sqrt{2}}{8} \right\}.\end{aligned}$$

Для $g_1 = (+1)$ будет

$$\frac{\partial \varphi(x)}{\partial g_1} = \frac{7\pi \sqrt{2}}{8}, \quad (4.33)$$

а для $g_2 = (-1)$

$$\frac{\partial \varphi(x)}{\partial g_2} = \frac{\pi \sqrt{2}}{8}.$$

Ясно, что $\psi(x) = \min_{\|g\|=1} \frac{\partial \varphi(x)}{\partial g} = \frac{\pi \sqrt{2}}{8} > 0$, т. е. точка $x = 1$ стационарная.

Теперь введем конечные сетки

$$G_N = \left\{ \frac{k}{N} \mid k \in \left[0 : E \left(\frac{3\pi}{2} N \right) \right] \right\}.$$

При любом натуральном N числа k/N рациональны, поэтому $3\pi/2 \notin G_N$ ни при каком N . В то же время $0 \in G_N$.

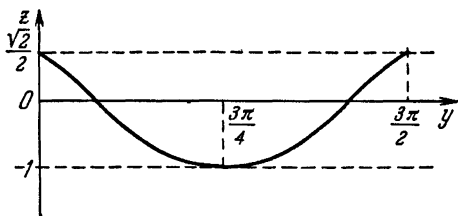


Рис. 34.

Для точки $x = 1$ имеем тогда $R_N(x) = \{0\}$. Поэтому

$$\frac{\partial \varphi_N(x)}{\partial g} = -g \frac{\pi \sqrt{2}}{8}.$$

В частности, при $x = 1$, $g_1 = (+1)$ для всех натуральных N получаем

$$\frac{\partial \varphi_N(x)}{\partial g_1} = -\frac{\pi \sqrt{2}}{8}.$$

Вспоминая (4.33), видим, что

$$\frac{\partial \varphi_N(x)}{\partial g_1} \not\rightarrow \frac{\partial \varphi(x)}{\partial g_1}.$$

Это мы и хотели показать.

§ 5. Частный случай теоремы о минимаксе

1. В этом параграфе будут получены некоторые достаточные условия для того, чтобы выполнялось равенство

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) = \max_{Y \in G} \min_{X \in \Omega} F(X, Y).$$

Докажем вначале одно вспомогательное утверждение, имеющее и самостоятельный интерес.

Пусть $G \subset E_m$ — ограниченное замкнутое выпуклое множество,

$$\varphi(X) = \max_{Y \in G} F(X, Y),$$

$$R(X) = \{Y \in G \mid F(X, Y) = \varphi(X)\},$$

$$H(X) = \left\{ \frac{\partial F(X, Y)}{\partial X} \mid Y \in R(X) \right\}.$$

Через $L(X)$ будем обозначать выпуклую оболочку множества $H(X)$:

$$L(X) = \text{co } H(X).$$

Теорема 5.1. Если при любом фиксированном X из некоторой окрестности $S_\delta(X_0)$ точки X_0 функция $F(X, Y)$ является вогнутой по Y на G , то

$$L(X_0) = H(X_0).$$

Для доказательства этой теоремы понадобится следующая

Лемма 5.1. В условиях теоремы 5.1 при любых $Y_1, Y_2 \in R(X_0)$ и $\alpha \in [0, 1]$ точка $Y_\alpha \stackrel{\text{def}}{=} \alpha Y_1 + (1 - \alpha) Y_2$ принадлежит $R(X_0)$ и

$$\frac{\partial F(X_0, Y_\alpha)}{\partial X} = \alpha \frac{\partial F(X_0, Y_1)}{\partial X} + (1 - \alpha) \frac{\partial F(X_0, Y_2)}{\partial X}. \quad (5.1)$$

Доказательство. Из вогнутости $F(X_0, Y)$ по Y следует

$$F(X_0, Y_\alpha) \geq \alpha F(X_0, Y_1) + (1 - \alpha) F(X_0, Y_2). \quad (5.2)$$

С другой стороны,

$$F(X_0, Y_1) = F(X_0, Y_2) = \varphi(X_0),$$

поэтому

$$F(X_0, Y_\alpha) \leq \varphi(X_0) = \alpha F(X_0, Y_1) + (1 - \alpha) F(X_0, Y_2). \quad (5.3)$$

Объединяя (5.2) и (5.3), получим

$$F(X_0, Y_\alpha) = \alpha F(X_0, Y_1) + (1 - \alpha) F(X_0, Y_2), \quad (5.4)$$

или, что то же самое,

$$F(X_0, Y_\alpha) = \varphi(X_0).$$

Значит, $Y_\alpha \in R(X_0)$.

Далее, из вогнутости $F(X, Y)$ по Y при любом $X \in S_\delta(X_0)$ имеем

$$F(X, Y_\alpha) \geq \alpha F(X, Y_1) + (1 - \alpha) F(X, Y_2). \quad (5.5)$$

Зафиксируем $\alpha \in [0, 1]$. Возьмем произвольный вектор $g \in E_n$, $\|g\| = 1$, и найдем производную функции $\Phi(X) \stackrel{df}{=} F(X, Y_\alpha)$ в точке X_0 по направлению g :

$$\begin{aligned} \frac{\partial \Phi(X_0)}{\partial g} &= \lim_{\beta \rightarrow +0} \frac{F(X_0 + \beta g, Y_\alpha) - F(X_0, Y_\alpha)}{\beta} = \\ &= \left(\frac{\partial F(X_0, Y_\alpha)}{\partial X}, g \right). \end{aligned} \quad (5.6)$$

Из (5.4) и (5.5) следует, что при малых $\beta > 0$

$$\begin{aligned} F(X_0 + \beta g, Y_\alpha) - F(X_0, Y_\alpha) &\geq \\ &\geq \alpha [F(X_0 + \beta g, Y_1) - F(X_0, Y_1)] + \\ &\quad + (1 - \alpha) [F(X_0 + \beta g, Y_2) - F(X_0, Y_2)]. \end{aligned}$$

Отсюда

$$\frac{\partial \Phi(X_0)}{\partial g} \geq \alpha \left(\frac{\partial F(X_0, Y_1)}{\partial X}, g \right) + (1 - \alpha) \left(\frac{\partial F(X_0, Y_2)}{\partial X}, g \right). \quad (5.7)$$

Из (5.6) и (5.7) получаем

$$\left(\frac{\partial F(X_0, Y_\alpha)}{\partial X} - \left[\alpha \frac{\partial F(X_0, Y_1)}{\partial X} + (1 - \alpha) \frac{\partial F(X_0, Y_2)}{\partial X} \right], g \right) \geq 0.$$

Поскольку последнее неравенство справедливо для произвольного $g \in E_n$, $\|g\| = 1$, то

$$\frac{\partial F(X_0, Y_\alpha)}{\partial X} - \left[\alpha \frac{\partial F(X_0, Y_1)}{\partial X} + (1 - \alpha) \frac{\partial F(X_0, Y_2)}{\partial X} \right] = 0,$$

что равносильно (5.1).

Лемма доказана.

Доказательство теоремы. В лемме 5.1 по существу установлено, что множество $H(X_0)$ выпуклое. С другой стороны, выпуклая оболочка выпуклого множества совпадает с исходным множеством (см. Приложение II, § 1). Отсюда следует, что $L(X_0) = H(X_0)$.

Теорема доказана.

2. Теорема 5.2. Пусть функция $F(X, Y)$ непрерывна вместе с $\frac{\partial F(X, Y)}{\partial X}$ на $\Omega' \times G$, где $\Omega' \subset E_n$, $G \subset E_m$. Предположим, что Ω' — открытое множество, а $\Omega \subset \Omega'$ и G — ограниченные замкнутые выпуклые множества. Если при каждом фиксированном $X_0 \in \Omega'$ функция $F(X_0, Y)$ вогнута по Y на G , а при каждом фиксированном $Y_0 \in G$ функция $F(X, Y_0)$ выпукла по X на Ω , то

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) = \max_{Y \in G} \min_{X \in \Omega} F(X, Y).$$

Доказательство. Рассмотрим функцию

$$\varphi(X) = \max_{Y \in G} F(X, Y).$$

Пусть $X^* \in \Omega$ — точка минимума $\varphi(X)$ на Ω . В силу теорем 2.2 и 3.1

$$L(X^*) \cap \Gamma^+(X^*) \neq \emptyset.$$

Пусть $Z^* \in L(X^*) \cap \Gamma^+(X^*)$. По теореме 5.1 $Z^* \in H(X^*)$, т. е. существует такая точка $Y^* \in R(X^*)$, что

$$Z^* = \frac{\partial F(X^*, Y^*)}{\partial X}.$$

Так как $Z^* \in \Gamma^+(X^*)$, то

$$\frac{\partial F(X^*, Y^*)}{\partial X} \in \Gamma^+(X^*). \quad (5.8)$$

В силу замечания 2 к теореме 3.1 главы IV и (5.8)

$$F(X^*, Y^*) \geq \min_{X \in \Omega} F(X, Y^*). \quad (5.9)$$

С другой стороны, $Y^* \in R(X^*)$, так что

$$F(X^*, Y^*) = \max_{Y \in G} F(X^*, Y). \quad (5.10)$$

Учитывая (5.9) и (5.10), получим для любых $X \in \Omega$ и $Y \in G$

$$F(X^*, Y) \leq F(X^*, Y^*) \leq F(X, Y^*). \quad (5.11)$$

Из (5.11) следует

$$\max_{Y \in G} F(X^*, Y) \leq \min_{X \in \Omega} F(X, Y^*).$$

Тем более

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) \leq \max_{Y \in G} \min_{X \in \Omega} F(X, Y). \quad (5.12)$$

Противоположное неравенство тривиально. Действительно, для любых $X \in \Omega$ и $Y \in G$

$$F(X, Y) \leq \max_{Y' \in G} F(X, Y').$$

Отсюда

$$\min_{X \in \Omega} F(X, Y) \leq \min_{X \in \Omega} \max_{Y' \in G} F(X, Y').$$

Поскольку последнее неравенство справедливо для всех $Y \in G$, то

$$\max_{Y \in G} \min_{X \in \Omega} F(X, Y) \leq \min_{X \in \Omega} \max_{Y \in G} F(X, Y). \quad (5.13)$$

Объединяя (5.12) и (5.13), получим требуемое.

Теорема доказана.

З а м е ч а н и е. Теорема 5.2 доказана при избыточном предположении (непрерывная дифференцируемость функции $F(X, Y)$ по Y). Однако приведенное доказательство конструктивно в том смысле, что, зная точку X^* , мы можем указать точку Y^* как точку, определяемую условием

$$\frac{\partial F(X^*, Y^*)}{\partial X} \in H(X^*) \cap \Gamma^+(X^*).$$

Пара $[X^*, Y^*]$, для которой выполняется неравенство (5.11), называется седловой точкой функции $F(X, Y)$.

3. Обратимся к следующей минимаксной задаче (см. замечание к теореме 6.1 главы IV):

$$\max_{V \in L} (V, Z - X) \rightarrow \min_{Z \in \Omega} \quad (5.14)$$

Будем предполагать, что Ω и L — ограниченные замкнутые множества в E_n , причем L — выпуклое, а Ω — строго выпуклое множества. По теореме 5.2 о минимаксе

$$\min_{Z \in \Omega} \max_{V \in L} (V, Z - X) = \max_{V \in L} \min_{Z \in \Omega} (V, Z - X). \quad (5.15)$$

Введем функцию

$$\theta(V) = \min_{Z \in \Omega} (V, Z - X).$$

Покажем, что $\theta(V)$ непрерывно дифференцируемая всюду на E_n , за исключением, быть может, точки $V = 0$. Положим

$$\hat{R}(V) = \{Z \in \Omega \mid (V, Z - X) = \theta(V)\}.$$

Лемма 5.2. Если Ω — строго выпуклое множество и $V \neq 0$, то $\hat{R}(V)$ состоит из единственной точки $Z(V)$, причем $Z(V)$ непрерывно зависит от V .

Доказательство. Предположим, что $V \neq 0$ и в $\hat{R}(V)$ входят две точки Z_1 и Z_2 :

$$(V, Z_1 - X) = (V, Z_2 - X) = \min_{Z \in \Omega} (V, Z - X).$$

Поскольку Ω — строго выпуклое множество, то точка $Z' \stackrel{\text{df}}{=} \frac{1}{2}(Z_1 + Z_2)$ является внутренней точкой Ω . В связи с этим найдется $\rho > 0$ такое, что

$$(Z' + Z) \in \Omega \quad \text{для всех } Z, \|Z\| < \rho.$$

Положим $Z_0 = \frac{-\rho V}{2\|V\|}$. Тогда

$$\begin{aligned} \min_{Z \in \Omega} (V, Z - X) &\leq (V, Z' + Z_0 - X) = \\ &= \frac{1}{2}(V, Z_1 - X) + \frac{1}{2}(V, Z_2 - X) - \frac{\rho}{2\|V\|}(V, V) = \\ &= \min_{Z \in \Omega} (V, Z - X) - \frac{1}{2}\rho\|V\|, \end{aligned}$$

что невозможно. Тем самым доказано, что множество $\hat{R}(V)$ состоит из единственной точки; обозначим ее через $Z(V)$.

Покажем теперь, что вектор-функция $Z(V)$ непрерывна в любой точке $V_0 \neq 0$. Допустим противное. Тогда найдется последовательность $\{V'_k\}$, сходящаяся к V_0 и такая, что $\{Z(V'_k)\}$ не сходится к $Z(V_0)$. Поскольку $Z(V'_k) \in \Omega$ при всех k , то существует подпоследовательность $\{Z(V'_{k_j})\}$, сходящаяся к некоторой точке Z^* , отличной от $Z(V_0)$. Очевидно, $Z^* \in \Omega$.

Заметим, что

$$(V'_{k_j}, Z(V'_{k_j}) - X) = \min_{Z \in \Omega} (V'_{k_j}, Z - X). \quad (5.16)$$

Переходя в равенстве (5.16) к пределу при $k_j \rightarrow \infty$, получим

$$(V_0, Z^* - X) = \min_{Z \in \Omega} (V_0, Z - X).$$

Таким образом, $Z^* \in \hat{R}(V_0)$, что невозможно, ибо $\hat{R}(V_0)$ состоит из единственной точки $Z(V_0)$.

Лемма доказана.

Приступим к доказательству непрерывной дифференцируемости функции $\theta(V)$, $V \neq 0$.

Поскольку

$$-\theta(V) = \max_{Z \in \Omega} (V, X - Z),$$

то в силу теоремы 2.1 функция $\theta(V)$ дифференцируема в любой точке $V \in E_n$ по любому направлению $g \in E_n$, $\|g\| = 1$. При этом

$$\frac{\partial \theta(V)}{\partial g} = - \frac{\partial (-\theta(V))}{\partial g} = \min_{Z \in \hat{R}(V)} (Z - X, g).$$

Если $V \neq 0$, то $\hat{R}(V)$ состоит из единственной точки $Z(V)$. Значит, для всех $g \in E_n$, $\|g\| = 1$, и $V \neq 0$ будет

$$\frac{\partial \theta(V)}{\partial g} = (Z(V) - X, g).$$

Учитывая, что функция $Z(V)$ непрерывна по V , заключаем (см. Приложение III, § 3), что при $V \neq 0$ функция $\theta(V)$ непрерывно дифференцируема и

$$\frac{\partial \theta(V)}{\partial V} = Z(V) - X. \quad (5.17)$$

Таким образом, если $0 \notin L$, то минимаксная задача (5.14) в силу (5.15) сведена к более простой задаче максимизации на L непрерывно дифференцируемой функции $\theta(V)$, причем градиент этой функции определяется формулой (5.17).

§ 6*. Разыскание седловых точек на многогранниках

1. Пусть $\Omega \subset E_n$, $G \subset E_m$ — ограниченные и замкнутые множества. Предположим, что на $\Omega \times G$ задана непрерывная функция $F(X, Y)$. Напомним еще раз определение седловой точки.

Определение. Точка $[X^*, Y^*] \in \Omega \times G$ называется *седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$* , если

$$F(X^*, Y) \leq F(X^*, Y^*) \leq F(X, Y^*) \quad (6.1)$$

для всех $X \in \Omega$, $Y \in G$.

Лемма 6.1. *Существование седловой точки функции $F(X, Y)$ на множестве $\Omega \times G$ эквивалентно выполнению соотношения*

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) = \max_{Y \in G} \min_{X \in \Omega} F(X, Y). \quad (6.2)$$

Доказательство. Прежде всего заметим, что все максимумы и минимумы достигаются в силу непрерывности $F(X, Y)$ и замкнутости и ограниченности множеств Ω и G .

Вначале докажем, что из существования седловой точки вытекает (6.2). Из (6.1) имеем

$$\max_{Y \in G} F(X^*, Y) \leq \min_{X \in \Omega} F(X, Y^*).$$

Отсюда

$$\min_{X \in \Omega} \max_{Y \in G} F(X, Y) \leq \max_{Y \in G} \min_{X \in \Omega} F(X, Y).$$

Сопоставляя это неравенство с очевидным неравенством

$$\max_{Y \in G} \min_{X \in \Omega} F(X, Y) \leq \min_{X \in \Omega} \max_{Y \in G} F(X, Y),$$

получаем (6.2).

Пусть теперь имеет место (6.2). Положим

$$\varphi(X) = \max_{Y \in G} F(X, Y), \quad \Phi(Y) = \min_{X \in \Omega} F(X, Y).$$

Из (6.2) и непрерывности функций $\varphi(X)$ и $\Phi(Y)$ следует существование точек $X^* \in \Omega$ и $Y^* \in G$ таких, что

$$\min_{X \in \Omega} \varphi(X) = \varphi(X^*) = \max_{Y \in G} \Phi(Y) = \Phi(Y^*). \quad (6.3)$$

Покажем, что $[X^*, Y^*]$ — седловая точка функции $F(X, Y)$ на множестве $\Omega \times G$.

Действительно, из (6.3) имеем для любых $[X, Y] \in \Omega \times G$

$$\begin{aligned} F(X^*, Y) &\leq \max_{Y' \in G} F(X^*, Y') = \varphi(X^*) = \Phi(Y^*) = \\ &= \min_{X \in \Omega} F(X, Y^*) \leq F(X^*, Y^*). \end{aligned} \quad (6.4)$$

Аналогично

$$\begin{aligned} F(X, Y^*) &\geq \min_{X' \in \Omega} F(X', Y^*) = \Phi(Y^*) = \varphi(X^*) = \\ &= \max_{Y \in G} F(X^*, Y) \geq F(X^*, Y^*). \end{aligned} \quad (6.5)$$

Объединяя (6.4) и (6.5), получим (6.1).

Лемма доказана.

2. Пусть $\Omega' \subset E_n$ и $G' \subset E_m$ — открытые множества, и пусть на $\Omega' \times G'$ задана дважды непрерывно дифференцируемая функция $F(X, Y)$.

В E_n рассмотрим множество

$$\Omega = \{X \mid (A_i, X) + a_i \leq 0, i \in [1 : N_1]\},$$

а в E_m — множество

$$G = \{Y \mid (B_j, Y) + b_j \leq 0, j \in [1 : N_2]\}.$$

Ясно, что Ω и G — выпуклые множества.

Предполагаем, что Ω и G не пусты, ограничены и что $\Omega \subset \Omega'$, $G \subset G'$.

Без ограничения общности можно считать, что

$$\|A_i\| = \|B_j\| = 1, \quad i \in [1 : N_1], \quad j \in [1 : N_2]. \quad (6.6)$$

Предположим, что функция $F(X, Y)$ является строго выпуклой по X и строго вогнутой по Y на $\Omega \times G$, т. е.

существуют такие $m_1 > 0$ и $m_2 > 0$, что для всех $[X, Y] \in \Omega \times G$ имеют место неравенства

$$\left(V, \frac{\partial^2 F(X, Y)}{\partial X^2} V \right) \geq m_1 \|V\|^2, \quad (6.7)$$

$$-\left(W, \frac{\partial^2 F(X, Y)}{\partial Y^2} W \right) \geq m_2 \|W\|^2 \quad (6.8)$$

для любых $V \in E_n$, $W \in E_m$.

Требуется найти седловую точку функции $F(X, Y)$ на множестве $\Omega \times G$, т. е. точку, удовлетворяющую неравенствам (6.1). Существование седловой точки следует из теоремы 5.2 и леммы 6.1.

Пусть $[X, Y] \in \Omega \times G$. Рассмотрим конусы

$$\Gamma_1^+(X) = \left\{ g \in E_n \mid g = - \sum_{i \in Q_1(X)} \alpha_i A_i, \alpha_i \geq 0 \right\},$$

$$\Gamma_2^+(Y) = \left\{ q \in E_m \mid q = - \sum_{j \in Q_2(Y)} \beta_j B_j, \beta_j \geq 0 \right\},$$

где

$$Q_1(X) = \{i \in [1 : N_1] \mid (A_i, X) + a_i = 0\},$$

$$Q_2(Y) = \{j \in [1 : N_2] \mid (B_j, Y) + b_j = 0\}.$$

Если $Q_1(X) = \emptyset$, то по определению полагаем $\Gamma_1^+(X) = \{0\}$; если $Q_2(Y) = \emptyset$, то полагаем $\Gamma_2^+(Y) = \{0\}$.

Как показано в § 2 главы V, конус $\Gamma_1^+(X)$ является сопряженным к конусу возможных направлений множества Ω в точке X , а конус $\Gamma_2^+(Y)$ является сопряженным к конусу возможных направлений множества G в точке Y .

Введем функции

$$d_1(X, Y) = \min_{z \in \Gamma_1^+(X)} \left\| Z - \frac{\partial F(X, Y)}{\partial X} \right\| \stackrel{df}{=} \left\| Z_1(X, Y) - \frac{\partial F(X, Y)}{\partial X} \right\|,$$

$$d_2(X, Y) = \min_{z \in \Gamma_2^+(Y)} \left\| Z + \frac{\partial F(X, Y)}{\partial Y} \right\| \stackrel{df}{=} \left\| Z_2(X, Y) + \frac{\partial F(X, Y)}{\partial Y} \right\|.$$

Отметим, что для любого $[X, Y] \in \Omega \times G$ точка $[Z_1(X, Y), Z_2(X, Y)] \in E_n \times E_m$ единственна.

Если $d_1(X, Y) > 0$, то направление

$$g(X, Y) = \frac{1}{d_1(X, Y)} \left(Z_1(X, Y) - \frac{\partial F(X, Y)}{\partial X} \right)$$

является направлением наискорейшего спуска функции $f_1(Z) = F(Z, Y)$ на множестве Ω в точке $Z = X$.

Если же $d_2(X, Y) > 0$, то направление

$$q(X, Y) = \frac{1}{d_2(X, Y)} \left(Z_2(X, Y) + \frac{\partial F(X, Y)}{\partial Y} \right)$$

есть направление наискорейшего спуска функции $f_2(Z) = -F(X, Z)$ на множестве G в точке $Z = Y$.

Из определения седловых точек (см. (6.1)) и теоремы 3.1 главы IV следует

Теорема 6.1. *Для того чтобы точка $[X^*, Y^*] \in \Omega \times G$ была седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$, необходимо и достаточно, чтобы*

$$d_1(X^*, Y^*) = d_2(X^*, Y^*) = 0.$$

Геометрически это условие означает, что

$$\frac{\partial F(X^*, Y^*)}{\partial X} \in \Gamma_1^+(X^*), \quad -\frac{\partial F(X^*, Y^*)}{\partial Y} \in \Gamma_2^+(Y^*).$$

Пусть $\varepsilon > 0$ фиксировано. Рассмотрим множества

$$Q_{1\varepsilon}(X) = \{i \in [1 : N_1] \mid -\varepsilon \leq (A_i, X) + a_i \leq 0\},$$

$$Q_{2\varepsilon}(Y) = \{j \in [1 : N_2] \mid -\varepsilon \leq (B_j, Y) + b_j \leq 0\}$$

и конусы

$$\Gamma_{1\varepsilon}^+(X) = \left\{ g \in E_n \mid g = - \sum_{i \in Q_{1\varepsilon}(X)} \alpha_i A_i, \alpha_i \geq 0 \right\},$$

$$\Gamma_{2\varepsilon}^+(Y) = \left\{ q \in E_m \mid q = - \sum_{j \in Q_{2\varepsilon}(Y)} \beta_j B_j, \beta_j \geq 0 \right\}.$$

Положим

$$\begin{aligned} d_{1\varepsilon}(X, Y) &= \min_{Z \in \Gamma_{1\varepsilon}^+(X)} \left\| Z - \frac{\partial F(X, Y)}{\partial X} \right\| \stackrel{df}{=} \\ &\stackrel{df}{=} \left\| Z_{1\varepsilon}(X, Y) - \frac{\partial F(X, Y)}{\partial X} \right\| \stackrel{df}{=} \|g_\varepsilon(X, Y)\|, \end{aligned} \quad (6.9)$$

$$\begin{aligned} d_{2\varepsilon}(X, Y) &= \min_{Z \in \Gamma_{2\varepsilon}^+(Y)} \left\| Z + \frac{\partial F(X, Y)}{\partial Y} \right\| \stackrel{df}{=} \\ &\stackrel{df}{=} \left\| Z_{2\varepsilon}(X, Y) + \frac{\partial F(X, Y)}{\partial Y} \right\| \stackrel{df}{=} \|q_\varepsilon(X, Y)\|. \end{aligned} \quad (6.10)$$

Нетрудно проверить, что для всех $i \in Q_{1\varepsilon}(X)$ и $j \in Q_{2\varepsilon}(Y)$ будет

$$(A_i, g_\varepsilon(X, Y)) \leq 0, \quad (B_j, g_\varepsilon(X, Y)) \leq 0. \quad (6.11)$$

Определение. Точка $[X^*, Y^*] \in \Omega \times G$ называется ε -седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$, если

$$d_{1\varepsilon}(X^*, Y^*) = d_{2\varepsilon}(X^*, Y^*) = 0. \quad (6.12)$$

Лемма 6.2. Точка $Z_{1\varepsilon}(X, Y)$ может быть представлена в виде

$$Z_{1\varepsilon}(X, Y) = - \sum_{i \in \tilde{Q}_{1\varepsilon}(X, Y)} \alpha_i(X, Y) A_i,$$

где $\alpha_i(X, Y) \geq 0$ и

$$\tilde{Q}_{1\varepsilon}(X, Y) = \{i \in Q_{1\varepsilon}(X) \mid (A_i, g_\varepsilon(X, Y)) = 0\}.$$

Доказательство. Утверждение леммы очевидно, если $Z_{1\varepsilon}(X, Y) = 0$. Пусть $Z_{1\varepsilon}(X, Y) \neq 0$.

Если $g_\varepsilon(X, Y) = 0$, то утверждение леммы также очевидно, ибо тогда $\tilde{Q}_{1\varepsilon}(X, Y) = Q_{1\varepsilon}(X)$. Пусть теперь $g_\varepsilon(X, Y) \neq 0$. Из (6.9) и (6.11) немедленно имеем

$$(Z_{1\varepsilon}(X, Y), g_\varepsilon(X, Y)) = 0, \quad (6.13)$$

$$(A_i, g_\varepsilon(X, Y)) < 0 \quad (6.14)$$

для всех $i \in \hat{Q}_{1\varepsilon}(X, Y) \stackrel{\text{def}}{=} Q_{1\varepsilon}(X) \setminus \tilde{Q}_{1\varepsilon}(X, Y)$.

Действительно, если бы (6.13) оказалось нарушенным, то мы бы сразу нашли точку $Z'_{1\varepsilon} \in \Gamma_{1\varepsilon}^+(X)$ такую, что

$$\left\| Z'_{1\varepsilon} - \frac{\partial F(X, Y)}{\partial X} \right\| < d_{1\varepsilon}(X, Y),$$

что невозможно в силу (6.9).

Неравенство же (6.14) следует из (6.11) и определения множества $\tilde{Q}_{1\varepsilon}(X, Y)$.

Множество $\tilde{Q}_{1\varepsilon}(X, Y)$ не пусто, ибо $Z_{1\varepsilon}(X, Y) \neq 0$.

Пусть $Z_{1\varepsilon}(X, Y) = - \sum_{i \in \tilde{Q}_{1\varepsilon}(X, Y)} \alpha_i(X, Y) A_i$, $\alpha_i(X, Y) \geq 0$.

Перепишем $Z_{1\varepsilon}(X, Y)$ в виде

$$Z_{1\varepsilon}(X, Y) = C_1 + C_2,$$

где

$$C_1 = - \sum_{i \in \tilde{Q}_{1\varepsilon}(X, Y)} \alpha_i(X, Y) A_i, \quad C_2 = - \sum_{i \in \hat{Q}_{1\varepsilon}(X, Y)} \alpha_i(X, Y) A_i.$$

Покажем, что $C_2 = 0$. Действительно, из (6.13) имеем

$$(Z_{1e}(X, Y), g_e(X, Y)) = (C_1, g_e(X, Y)) + (C_2, g_e(X, Y)) =$$

$$= (C_2, g_e(X, Y)) = - \sum_{i \in \hat{Q}_{1e}(X, Y)} \alpha_i(X, Y) (A_i, g_e(X, Y)) = 0. \quad (6.15)$$

Так как в силу (6.14) $(A_i, g_e(X, Y)) < 0$ и $\alpha_i(X, Y) \geq 0$, то (6.15) возможно лишь при $\alpha_i(X, Y) = 0$ для всех $i \in \hat{Q}_{1e}(X, Y)$.

Лемма доказана.

Аналогично доказывается

Лемма 6.3. Точка $Z_{2e}(X, Y)$ может быть представлена в виде

$$Z_{2e}(X, Y) = - \sum_{i \in \hat{Q}_{2e}(X, Y)} \beta_i(X, Y) B_i,$$

где $\beta_i(X, Y) \geq 0$ и

$$\tilde{Q}_{2e}(X, Y) = \{j \in Q_{2e}(Y) \mid (q_e(X, Y), B_j) = 0\}.$$

3. Зафиксируем $\varepsilon > 0$ и опишем метод последовательных приближений для нахождения ε -седловой точки функции $F(X, Y)$ на множестве $\Omega \times G$. Введем в рассмотрение функцию

$$d_e(X, Y) = \frac{1}{2} [d_{1e}^2(X, Y) + d_{2e}^2(X, Y)].$$

Ясно, что условие (6.12) эквивалентно условию

$$d_e(X^*, Y^*) = 0.$$

В качестве начального приближения выберем произвольную точку $[X_0, Y_0] \in \Omega \times G$. Пусть точка $[X_k, Y_k] \in \Omega \times G$ уже найдена.

Если $d_e(X_k, Y_k) = 0$, то точка $[X_k, Y_k]$ является ε -седловой, и процесс прекращается. Если же $d_e(X_k, Y_k) > 0$, то рассмотрим лучи

$$X = X_k(\alpha) \stackrel{df}{=} X_k + \alpha g_k, \quad Y = Y_k(\alpha) \stackrel{df}{=} Y_k + \alpha q_k, \quad \alpha \geq 0,$$

где $g_k = g_e(X_k, Y_k)$, $q_k = q_e(X_k, Y_k)$.

Найдем $\alpha_k \geq 0$ такое, что

$$d_e(X_k(\alpha_k), Y_k(\alpha_k)) = \min_{\substack{\alpha \geq 0 \\ X_k(\alpha) \in \Omega, Y_k(\alpha) \in G}} d_e(X_k(\alpha), Y_k(\alpha)).$$

Положим теперь

$$X_{k+1} = X_k(\alpha_k), \quad Y_{k+1} = Y_k(\alpha_k).$$

Ясно, что

$$[X_{k+1}, Y_{k+1}] \in \Omega \times G, \quad d_e(X_{k+1}, Y_{k+1}) \leq d_e(X_k, Y_k). \quad (6.16)$$

Далее продолжаем аналогично.

Таким образом, строим последовательность

$$\{[X_k, Y_k]\}, \quad [X_k, Y_k] \in \Omega \times G, \quad k = 0, 1, 2, \dots$$

Если эта последовательность содержит конечное число элементов, то последняя полученная точка по построению является ε -седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$. В противном случае справедлива

Теорема 6.2. Любая предельная точка построенной выше последовательности $\{[X_k, Y_k]\}$ является ε -седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$.

Доказательство. Существование предельных точек следует из ограниченности и замкнутости множества $\Omega \times G$. Пусть

$$d^* = \lim_{k \rightarrow \infty} d_e(X_k, Y_k).$$

Ясно, что для всех k в силу (6.16)

$$d_e(X_k, Y_k) \geq d^*. \quad (6.17)$$

Пусть

$$[X_{k_s}, Y_{k_s}] \xrightarrow{k_s \rightarrow \infty} [X^*, Y^*].$$

Докажем, что $d_e(X_{k_s}, Y_{k_s}) \xrightarrow{k_s \rightarrow \infty} 0$. Тем самым в силу (6.16) будет доказано, что

$$d_e(X_k, Y_k) \xrightarrow{k \rightarrow \infty} 0. \quad (6.18)$$

Допустим противное

$$d_e(X_{k_s}, Y_{k_s}) \geq 2a^2 > 0, \quad a > 0.$$

Тогда либо $d_{1e}(X_{k_s}, Y_{k_s}) \geq a$, либо $d_{2e}(X_{k_s}, Y_{k_s}) \geq a$.

Пусть, например, для всех k_s

$$d_{1e}(X_{k_s}, Y_{k_s}) \geq a$$

(если неравенство $d_{1e}(X_{k_s}, Y_{k_s}) \geq a$ встречается бесконечное число раз, то, прореживая при необходимости

последовательность $\{[X_{k_s}, Y_{k_s}]\}$, можно добиться выполнения этого неравенства для всех k_s .

Без ограничения общности можем также считать, что

$$Q_{1\varepsilon}(X_{k_s}, Y_{k_s}) = Q' \subset [1 : N_1],$$

$$Q_{2\varepsilon}(X_{k_s}, Y_{k_s}) = Q'' \subset [1 : N_2].$$

Отметим, что

$$Q_{1\varepsilon}(X^*, Y^*) \supset Q', \quad Q_{2\varepsilon}(X^*, Y^*) \supset Q''.$$

Учитывая леммы 6.2, 6.3 и определение множеств $\tilde{Q}_{1\varepsilon}$ и $\tilde{Q}_{2\varepsilon}$, получим для $\alpha > 0$

$$Z_{1\varepsilon}(X_{k_s}, Y_{k_s}) \in \Gamma_{1\varepsilon}^+(X_{k_s}(\alpha)),$$

$$Z_{2\varepsilon}(X_{k_s}, Y_{k_s}) \in \Gamma_{2\varepsilon}^+(Y_{k_s}(\alpha)),$$

и потому

$$d_{1\varepsilon}(X_{k_s}(\alpha), Y_{k_s}(\alpha)) \leq \left\| Z_{1\varepsilon}(X_{k_s}, Y_{k_s}) - \frac{\partial F(X_{k_s}(\alpha), Y_{k_s}(\alpha))}{\partial X} \right\|, \quad (6.19)$$

$$d_{2\varepsilon}(X_{k_s}(\alpha), Y_{k_s}(\alpha)) \leq \left\| Z_{2\varepsilon}(X_{k_s}, Y_{k_s}) + \frac{\partial F(X_{k_s}(\alpha), Y_{k_s}(\alpha))}{\partial Y} \right\|. \quad (6.20)$$

При $\alpha > 0$ таких, что

$$\alpha d_{1\varepsilon}(X_k, Y_k) \leq \varepsilon, \quad \alpha d_{2\varepsilon}(X_k, Y_k) \leq \varepsilon,$$

будет $X_k(\alpha) \in \Omega$, $Y_k(\alpha) \in G$.

Это легко доказать, если учесть (6.11), (6.6) и равенства

$$\alpha d_{1\varepsilon}(X_k, Y_k) = \| \alpha g_k \|, \quad \alpha d_{2\varepsilon}(X_k, Y_k) = \| \alpha q_k \|.$$

Более того, в силу непрерывной дифференцируемости функции $F(X, Y)$ и ограниченности $\Omega \times G$ существует $\alpha_0 > 0$ такое, что для всех k окажется $X_k(\alpha) \in \Omega$, $Y_k(\alpha) \in G$ при $\alpha \in [0, \alpha_0]$.

Воспользуемся теперь определением $d_\varepsilon(X, Y)$, равенством $\|X\|^2 = (X, X)$ и формулами (6.9), (6.10), (6.19),

(6.20), (6.7) и (6.8). Получим

$$\begin{aligned}
 d_e(X_{k_s}(\alpha), Y_{k_s}(\alpha)) &\leq \\
 &\leq d_e(X_{k_s}, Y_{k_s}) + \alpha \left[- \left(g_{k_s}, \frac{\partial^2 F(X_{k_s}, Y_{k_s})}{\partial X^2} g_{k_s} \right) - \right. \\
 &- \left(g_{k_s}, \frac{\partial^2 F(X_{k_s}, Y_{k_s})}{\partial X \partial Y} q_{k_s} \right) + \left(q_{k_s}, \frac{\partial^2 F(X_{k_s}, Y_{k_s})}{\partial Y \partial X} g_{k_s} \right) + \\
 &+ \left. \left(q_{k_s}, \frac{\partial^2 F(X_{k_s}, Y_{k_s})}{\partial Y^2} q_{k_s} \right) \right] + o_{k_s}(\alpha) \leq d_e(X_{k_s}, Y_{k_s}) - \\
 &- \alpha [m_1 d_{1e}^2(X_{k_s}, Y_{k_s}) + m_2 d_{2e}^2(X_{k_s}, Y_{k_s})] + o_{k_s}(\alpha),
 \end{aligned}$$

где $\frac{o_{k_s}(\alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$ равномерно по k_s .

Так как, по предположению, $d_{1e}(X_{k_s}, Y_{k_s}) \geq a$, то при $\alpha \in [0, \alpha_1]$, где $\alpha_1 \in (0, \alpha_0]$ не зависит от k_s , будет

$$d_e(X_{k_s}(\alpha), Y_{k_s}(\alpha)) \leq d_e(X_{k_s}, Y_{k_s}) - \frac{1}{2} m_1 a^2 \alpha.$$

Зафиксируем любое $\alpha' \in (0, \alpha_1]$. Имеем для всех k_s

$$\begin{aligned}
 d_e(X_{k_s+1}, Y_{k_s+1}) &\leq d_e(X_{k_s}(\alpha'), Y_{k_s}(\alpha')) \leq \\
 &\leq d_e(X_{k_s}, Y_{k_s}) - \frac{1}{2} m_1 a^2 \alpha',
 \end{aligned}$$

что противоречит (6.17). Итак, доказано, что

$$d_e(X_{k_s}, Y_{k_s}) \xrightarrow{k_s \rightarrow \infty} 0. \quad (6.21)$$

Так как при больших k_s

$$Q_{1e}(X_{k_s}) \subset Q_{1e}(X^*), \quad Q_{2e}(Y_{k_s}) \subset Q_{2e}(Y^*),$$

то отсюда и из (6.21) следует

$$d_e(X^*, Y^*) = 0,$$

т. е. точка $[X^*, Y^*]$ является ε -седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$. Теорема доказана.

З а м е ч а н и е. Тот факт, что $\varepsilon > 0$, был существенно использован при доказательстве теоремы 6.2, ибо лишь при этом условии найдется $\alpha_0 > 0$ (а потому и $\alpha_1 > 0$), не зависящее от k_s .

4. Теперь опишем метод последовательных приближений для нахождения седловых точек функции $F(X, Y)$ на $\Omega \times G$.

Зафиксируем $\varepsilon_1 > 0$, $a > 0$ и возьмем начальное приближение $[X_0, Y_0] \in \Omega \times G$. Пользуясь описанным в п. 3 методом с $\varepsilon = \varepsilon_1$ в конечное (в силу (6.18)) число шагов найдем точку $[X_1, Y_1] \in \Omega \times G$, для которой $d_{\varepsilon_1}(X_1, Y_1) \leq \leq a\varepsilon_1$. Положим $\varepsilon_2 = \frac{1}{2} \varepsilon_1$. Снова воспользуемся методом п. 3 с начальным приближением $[X_1, Y_1]$ и $\varepsilon = \varepsilon_2$. В конечное число шагов получим точку $[X_2, Y_2] \in \Omega \times G$, для которой

$$d_{2\varepsilon}(X_2, Y_2) \leq a\varepsilon_2.$$

Далее поступаем аналогично.

В результате строим последовательность

$$\{[X_k, Y_k]\}, \quad [X_k, Y_k] \in \Omega \times G, \\ d_{\varepsilon_k}(X_k, Y_k) \leq a\varepsilon_k, \quad \varepsilon_k = 2^{1-k}\varepsilon_1, \quad k = 1, 2, \dots \quad (6.22)$$

Имеет место

Теорема 6.3. *Любая предельная точка последовательности $\{[X_k, Y_k]\}$ является седловой точкой функции $F(X, Y)$ на множестве $\Omega \times G$.*

Доказательство. Существование предельных точек очевидно. Пусть

$$[X_{k_s}, Y_{k_s}] \xrightarrow{k_s \rightarrow \infty} [X^*, Y^*] \in \Omega \times G.$$

Утверждается, что

$$d(X^*, Y^*) \stackrel{\text{df}}{=} \frac{1}{2} (d_1^2(X^*, Y^*) + d_2^2(X^*, Y^*)) = 0.$$

Допустим противное: $d(X^*, Y^*) = a > 0$. При достаточно больших k_s будет

$$Q_{1\varepsilon_{k_s}}(X_{k_s}) \subset Q_1(X^*), \quad Q_{2\varepsilon_{k_s}}(Y_{k_s}) \subset Q_2(Y^*),$$

и потому

$$d_{\varepsilon_{k_s}}(X_{k_s}, Y_{k_s}) \geq \frac{a}{2},$$

что противоречит (6.22). Теорема доказана.

Замечание 1. Везде выше предполагалось, что функция $F(X, Y)$ — строго выпукло-вогнутая. Если функция $F(X, Y)$ является выпукло-вогнутой, то вместо нее можно рассмотреть строго выпукло-вогнутую функцию

$$F_1(X, Y) = F(X, Y) + c(X, X) - d(Y, Y),$$

где $c > 0$ и $d > 0$ — произвольные числа. При достаточно малых c и d седловая точка функции $F_1(X, Y)$ будет близка к седловой точке функции $F(X, Y)$.

Замечание 2. Для выпукло-вогнутых функций в силу леммы 3.2 Приложения II и замечания к ней справедливы неравенства

$$F(X_0, Y_0) + \min_{X \in \Omega} \left(\frac{\partial F(X_0, Y_0)}{\partial X}, X - X_0 \right) \leqslant \\ \leqslant \min_{X \in \Omega} \max_{Y \in G} F(X, Y) \leqslant F(X_0, Y_0) + \max_{Y \in G} \left(\frac{\partial F(X_0, Y_0)}{\partial Y}, Y - Y_0 \right).$$

Эти неравенства дают возможность прекратить поиск седловой (или ϵ -седловой) точки при достижении достаточной точности.

Замечание 3. Линейность функций, задающих множества Ω и G , существенна. Если же эти функции нелинейны, то можно множества Ω и G вначале аппроксимировать множествами рассмотренного в п. 2 вида, а затем применить изложенные выше методы.

§ 7. Наилучшее приближение функций нескольких аргументов обобщенными полиномами

1. Пусть $G \subset E_m$ — ограниченное замкнутое множество.

Определение. Система непрерывных на G функций

$$u_0(Y), u_1(Y), \dots, u_n(Y) \quad (7.1)$$

называется *линейно независимой на G* , если обобщенный полином $\Phi_n(A, Y)$

$$\Phi_n(A, Y) = \sum_{k=0}^n a_k u_k(Y), \quad \text{где} \quad A = (a_0, a_1, \dots, a_n),$$

тождественно равен нулю на G тогда и только тогда, когда

$$A = 0.$$

В дальнейшем будем предполагать, что система функций (7.1) линейно независима на множестве G .

Пусть $f(Y)$ — непрерывная на G функция. Положим

$$\rho = \inf_{A \in \Omega} \max_{Y \in G} |f(Y) - \Phi_n(A, Y)|, \quad (7.2)$$

где $\Omega \subset E_{n+1}$ — замкнутое выпуклое множество.

Обобщенный полином $\Phi_n(A^*, Y)$, $A^* \in \Omega$, называется *полиномом наилучшего приближения функции $f(Y)$ на множестве G при ограничении Ω или просто полиномом наилучшего приближения*, если

$$\max_{Y \in G} |f(Y) - \Phi_n(A^*, Y)| = \rho. \quad (7.3)$$

Множество векторов $A^* \in \Omega$, для которых выполняется равенство (7.3), совпадает с множеством решений следующей минимаксной задачи:

$$\max_{Y \in G} (f(Y) - \Phi_n(A, Y))^2 \rightarrow \inf_{A \in \Omega}. \quad (7.4)$$

Действительно, пусть $\Phi_n(A^*, Y)$, $A^* \in \Omega$, — полином наилучшего приближения. Тогда для всех $A \in \Omega$

$$\max_{Y \in G} |\Delta_n(A^*, Y)| \leq \max_{Y \in G} |\Delta_n(A, Y)|, \quad (7.5)$$

где $\Delta_n(A, Y) = f(Y) - \Phi_n(A, Y)$. Возводя неравенство (7.5) в квадрат и учитывая соотношение (см. Приложение III, § 2)

$$(\max_{Y \in G} |\Delta_n(A, Y)|)^2 = \max_{Y \in G} (\Delta_n(A, Y))^2,$$

получим

$$\max_{Y \in G} (\Delta_n(A^*, Y))^2 \leq \max_{Y \in G} (\Delta_n(A, Y))^2.$$

Отсюда следует, что A^* является решением минимаксной задачи (7.4). Аналогично доказывается, что всякое решение задачи (7.4) дает вектор коэффициентов полинома наилучшего приближения.

Таким образом, доказана эквивалентность минимаксных задач (7.2) и (7.4) в смысле множества решений и установлено равенство

$$\rho^2 = \inf_{A \in \Omega} \max_{Y \in G} (f(Y) - \Phi_n(A, Y))^2.$$

В дальнейшем под полиномом наилучшего приближения будем понимать обобщенный полином $\Phi_n(A^*, Y)$,

$A^* \in \Omega$, для которого

$$\max_{Y \in G} (f(Y) - \Phi_n(A^*, Y))^2 = \rho^2.$$

Введем обозначения:

$$F(A, Y) = (\Delta_n(A, Y))^2 = (f(Y) - \Phi_n(A, Y))^2, \\ \varphi(A) = \max_{Y \in G} F(A, Y).$$

Теорема 7.1. *Обобщенный полином наилучшего приближения существует.*

Доказательство. Пусть последовательность

$$A_k \stackrel{df}{=} (a_0^{(k)}, a_1^{(k)}, \dots, a_n^{(k)}) \in \Omega, \quad k = 1, 2, \dots,$$

такова, что

$$\varphi(A_k) \rightarrow \rho^2. \quad (7.6)$$

Для доказательства теоремы достаточно установить, что последовательность $\{A_k\}$ ограничена. Допустим противное. Тогда в числовой последовательности $\{\tau_k\}$

$$\tau_k = \sum_{i=0}^n |a_i^{(k)}|, \quad k = 1, 2, \dots,$$

найдется подпоследовательность, сходящаяся к бесконечности. Не ограничивая общности, будем считать, что вся последовательность $\{\tau_k\}$ сходится к бесконечности

$$\tau_k \xrightarrow{k \rightarrow \infty} \infty. \quad (7.7)$$

В силу (7.6) равномерно по k выполняется неравенство

$$\max_{Y \in G} \left| \sum_{i=0}^n a_i^{(k)} u_i(Y) \right| \leq c, \quad (7.8)$$

где c — некоторая константа.

Положим

$$\bar{A}_k = (\bar{a}_0^{(k)}, \bar{a}_1^{(k)}, \dots, \bar{a}_n^{(k)}),$$

где $\bar{a}_i^{(k)} = \frac{a_i^{(k)}}{\tau_k}$, $i \in [0:n]$. Тогда

$$\sum_{i=0}^n |\bar{a}_i^{(k)}| = 1, \quad k = 1, 2, \dots, \quad (7.9)$$

и в силу (7.8)

$$\max_{Y \in G} \left| \sum_{i=0}^n \bar{a}_i^{(k)} u_i(Y) \right| \leq \frac{c}{\tau_k}. \quad (7.10)$$

Поскольку последовательность $\{\bar{A}_k\}$ ограничена, то из нее можно выделить сходящуюся подпоследовательность. Не ограничивая общности, будем считать, что вся последовательность $\{\bar{A}_k\}$ сходится к точке $\bar{A} = (\bar{a}_0, \bar{a}_1, \dots, \bar{a}_n)$.

Переходя к пределу при $k \rightarrow \infty$ в (7.9) и (7.10) и учитывая (7.7), получим

$$\sum_{i=0}^k |\bar{a}_i| = 1, \\ \sum_{i=0}^k \bar{a}_i u_i(Y) = 0 \quad \text{для всех } Y \in G.$$

Но это противоречит линейной независимости системы функций $u_0(Y), u_1(Y), \dots, u_n(Y)$ на G .

Теорема доказана.

З а м е ч а н и е. Из доказательства теоремы следует, что при любом $A_0 \in E_{n+1}$ множество $\{A \mid \varphi(A) \leq \varphi(A_0)\}$ ограничено. Тем более будет ограниченным множество

$$M(A_0) = \{A \in \Omega \mid \varphi(A) \leq \varphi(A_0)\}$$

при $A_0 \in \Omega$.

2. Функция $F(A, Y)$ выпукла по A на E_{n+1} при любом $Y \in G$. Действительно, пусть $\alpha \in [0, 1]$. Тогда

$$\begin{aligned} F(\alpha A_1 + (1 - \alpha) A_2, Y) &\leq (\alpha \Delta_n(A_1, Y) + (1 - \alpha) \Delta_n(A_2, Y))^2 + \\ &+ \alpha(1 - \alpha)(\Delta_n(A_1, Y) - \Delta_n(A_2, Y))^2 = \\ &= \alpha(\Delta_n(A_1, Y))^2 + (1 - \alpha)(\Delta_n(A_2, Y))^2 = \\ &= \alpha F(A_1, Y) + (1 - \alpha) F(A_2, Y). \end{aligned}$$

Далее для $i \in [0 : n]$

$$\left(\frac{\partial F(A, Y)}{\partial A} \right)_i = -2 \Delta_n(A, Y) u_i(Y). \quad (7.11)$$

Положим

$$R(A) = \{Y \in G \mid F(A, Y) = \varphi(A)\}.$$

Очевидно, что

$$R(A) = \{Y \in G \mid |\Delta_n(A, Y)| = \max_{V \in G} |\Delta_n(A, V)|\}.$$

Теорема 7.2. Для того чтобы полином $\Phi_n(A^*, Y)$, $A^* \in \Omega$, был полиномом наилучшего приближения функции $f(Y)$ на множестве G при ограничении Ω , необходимо и достаточно, чтобы выполнялось равенство

$$\inf_{A \in \Omega} \max_{Y \in R(A^*)} \Delta_n(A^*, Y) [\Phi_n(A^*, Y) - \Phi_n(A, Y)] = 0. \quad (7.12)$$

Доказательство. Эта теорема является очевидным следствием теоремы 2.2. Нужно заметить только, что в силу (7.11)

$$\left(\frac{\partial F(A^*, Y)}{\partial A}, A - A^* \right) = 2\Delta_n(A^*, Y) [\Phi_n(A^*, Y) - \Phi_n(A, Y)].$$

Если $\Omega = E_{n+1}$, то (7.12) можно заменить более простым равенством

$$\inf_{A \in E_{n+1}} \max_{Y \in R(A^*)} \Delta_n(A^*, Y) \Phi_n(A, Y) = 0.$$

Теорема 7.3. Для того чтобы полином $\Phi_n(A^*, Y)$, $A^* \in \Omega$, был полиномом наилучшего приближения функции $f(Y)$ на множестве G при ограничении Ω , необходимо и достаточно, чтобы нашлись r , $1 \leq r \leq n+2$, попарно различных точек $Y_1, \dots, Y_r$ из $R(A^*)$ и r отрицательных чисел $\alpha_1, \dots, \alpha_r$, $\sum_{i=1}^r \alpha_i = 1$, таких, что

$$\inf_{A \in \Omega} \sum_{i=1}^r \alpha_i \Delta_n(A^*, Y_i) [\Phi_n(A^*, Y_i) - \Phi_n(A, Y_i)] = 0. \quad (7.13)$$

Если $\Omega = E_{n+1}$, то (7.13) можно заменить другим равенством

$$\sum_{i=1}^r \alpha_i \Delta_n(A^*, Y_i) \begin{pmatrix} u_0(Y_i) \\ u_1(Y_i) \\ \vdots \\ u_n(Y_i) \end{pmatrix} = 0.$$

Эта теорема является следствием теоремы 3.2.

Теорема 7.4. Существует такое конечное множество

$$G_r = \{Y_i \in G \mid i \in [1:r], \quad 1 \leq r \leq n+2\},$$

что задача наилучшего приближения функции $f(Y)$ на множестве G при ограничении Ω равносильна задаче наилучшего приближения функции $f(Y)$ на множестве G_r при ограничении Ω .

Справедливость этого утверждения следует из теоремы 3.3.

3. Для нахождения полинома наилучшего приближения может быть использован сеточный метод. Для этого на множестве G вводится последовательность сеток G_N , всюду плотная в G при $N \rightarrow \infty$ (см. § 4). Пусть $\Phi_n(A_N^*, Y)$, $A_N^* \in \Omega$, — один из полиномов наилучшего приближения функции $f(Y)$ на сетке G_N при ограничении Ω .

Справедлива следующая

Теорема 7.5. Всякая предельная точка последовательности $\{A_N^*\}$ дает вектор коэффициентов полинома наилучшего приближения функции $f(Y)$ на множестве G при ограничении Ω .

Предварительно докажем две леммы.

Лемма 7.1. Существует такой номер N_0 , что при $N > N_0$ система функций

$$u_0(Y), u_1(Y), \dots, u_n(Y)$$

линейно независима на G_N .

Доказательство. Допустим противное. Тогда для некоторой подпоследовательности номеров N_s , $s = 1, 2, \dots$, будут выполняться равенства

$$\sum_{k=0}^n a_k^{(s)} u_k(Y) = 0; \quad Y \in G_{N_s}; \quad \sum_{k=0}^n |a_k^{(s)}| > 0. \quad (7.14)$$

Пусть $a_{k_s}^{(s)}$ имеет наибольшее по модулю значение среди всех $a_k^{(s)}$, $k \in [0:n]$. Положив

$$\bar{a}_k^{(s)} = \frac{a_k^{(s)}}{a_{k_s}^{(s)}},$$

придем к эквивалентной (7.14) системе равенств

$$\sum_{k=0}^n \bar{a}_k^{(s)} u_k(Y) = 0, \quad Y \in G_{N_s}. \quad (7.15)$$

При этом $|\bar{a}_k^{(s)}| \leq 1$ и один из коэффициентов $\bar{a}_k^{(s)}$, $k \in [0:n]$, равен единице. Переходя, если нужно, к подпоследовательностям, добьемся того, что при некотором k_0 для всех s выполняется равенство $\bar{a}_{k_0}^{(s)} = 1$, и последовательности $\{\bar{a}_k^{(s)}\}$ сходятся при $s \rightarrow \infty$:

$$\bar{a}_k^{(s)} \rightarrow \bar{a}_k, \quad k \in [0:n].$$

Тогда, учитывая непрерывность функций $u_k(Y)$ и плотность сеток G_N в G при $N \rightarrow \infty$, получим на основании (7.15) для всех $Y \in G$

$$u_{k_0}(Y) + \sum_{\substack{k=0 \\ k \neq k_0}}^n \bar{a}_k u_k(Y) = 0,$$

что противоречит линейной независимости системы функций (7.1) на множестве G .

Лемма доказана.

Лемма 7.2. Если последовательность векторов $\{A_N\}$ такова, что равномерно по N

$$\max_{Y \in G_N} |\Phi_n(A_N, Y)| \leq c,$$

то последовательность векторов $\{A_N\}$ ограничена.

Доказательство. Допустим противное. Тогда, так же как при доказательстве теоремы 7.1, найдется последовательность векторов $\bar{A}_N = (\bar{a}_0^{(N)}, \bar{a}_1^{(N)}, \dots, \bar{a}_n^{(N)})$, $N = 1, 2, \dots$, со следующими свойствами:

$$\text{I.} \quad \sum_{i=0}^n |\bar{a}_i^{(N)}| = 1, \quad N = 1, 2, \dots$$

$$\text{II.} \quad \max_{Y \in G_N} \left| \sum_{i=0}^n \bar{a}_i^{(N)} u_i(Y) \right| \leq \frac{c}{\tau_N},$$

где $\tau_N \xrightarrow{N \rightarrow \infty} \infty$.

$$\text{III.} \quad \bar{A}_N \xrightarrow{N \rightarrow \infty} \bar{A} \stackrel{\text{df}}{=} (\bar{a}_0, \bar{a}_1, \dots, \bar{a}_n).$$

Отсюда следует, что для всех $Y \in G$

$$\sum_{i=0}^n \bar{a}_i u_i(Y) = 0,$$

причем $\sum_{i=0}^n |\bar{a}_i| = 1$. Но это противоречит линейной независимости системы функций (7.1) на G .

Лемма доказана.

Доказательство теоремы. Учитывая лемму 7.1, по теореме 7.1 получим, что при достаточно больших $N > N_0$ существует полином наилучшего приближения $\Phi_n(A_N^*, Y)$, $A_N^* \in \Omega$, функции $f(Y)$ на сетке G_N при ограничении Ω . Далее, нетрудно понять, что равномерно по $N > N_0$ выполняется неравенство

$$\max_{Y \in G_N} |\Phi_n(A_N^*, Y)| \leq \bar{c},$$

где $\bar{c}$ — некоторая константа.

Действительно,

$$\begin{aligned} \max_{Y \in G_N} |\Phi_n(A_N^*, Y)| &\leq \\ &\leq \max_{Y \in G_N} |f(Y) - \Phi_n(A_N^*, Y)| + \max_{Y \in G} |f(Y)| = \\ &= \min_{\{A\}} \max_{Y \in G_N} |f(Y) - \Phi_n(A, Y)| + \max_{Y \in G} |f(Y)| \leq \\ &\leq \rho + \max_{Y \in G} |f(Y)| \stackrel{df}{=} \bar{c}. \end{aligned}$$

Учитывая лемму 7.2, заключаем, что последовательность $\{A_N^*\}$ ограничена и, значит, имеет хотя бы одну предельную точку.

Остается сослаться на теорему 4.1 и заметить, что в силу выпуклости $F(A, Y)$ по A при каждом фиксированном $Y \in G$ любая стационарная точка функции $\Phi(A)$ на Ω дает вектор коэффициентов полинома наилучшего приближения функции $f(Y)$ на множестве G при ограничении Ω .

Теорема доказана.

§ 8. Наилучшее приближение функций, заданных на отрезке алгебраическими полиномами

1. Вернемся еще раз к задаче о наилучшем приближении функций, заданных на отрезке, алгебраическими многочленами и покажем, что теорема Чебышева является следствием результатов, полученных в этой главе.

Пусть

$$\rho = \min_{\{A\}} \max_{c \leq t \leq d} |f(t) - P_n(A, t)|, \quad (8.1)$$

где $f(t)$ — непрерывная на отрезке $[c, d]$ функция,

$$A = (a_0, a_1, \dots, a_n) \in E_{n+1}$$

и

$$P_n(A, t) = \sum_{i=0}^n a_i t^i.$$

Будем в дальнейшем предполагать, что $\rho > 0$, так что для любого $A \in E_{n+1}$

$$\max_{c \leq t \leq d} |f(t) - P_n(A, t)| > 0.$$

Положим $Y = (\xi, t) \in E_2$,

$$F(A, Y) = \xi(-f(t) + P_n(A, t)), \\ G = \{(\xi, t) | \xi = \pm 1, t \in [c, d]\}.$$

Очевидно, что для любого $A \in E_{n+1}$

$$\max_{Y \in G} F(A, Y) = \max_{c \leq t \leq d} |f(t) - P_n(A, t)|. \quad (8.2)$$

Таким образом,

$$\rho = \min_{\{A\}} \max_{Y \in G} F(A, Y). \quad (8.3)$$

Минимаксная задача (8.3) эквивалентна (8.1) и более удобна для исследования.

Поскольку

$$\frac{\partial F(A, Y)}{\partial A} = \xi(1, t, \dots, t^n), \quad (8.4)$$

то функция $F(A, Y)$ непрерывна вместе с $\partial F(A, Y)/\partial A$ на множестве $E_{n+1} \times G$. Более того, функция $F(A, Y)$ выпукла (даже линейна) по A при каждом фиксированном $Y \in G$.

Положим

$$Z(Y) = \xi(1, t, \dots, t^n),$$

$$R(A) = \{Y \in G \mid F(A, Y) = \max_{W \in G} F(A, W)\}.$$

Через $L(A)$ будем обозначать выпуклую оболочку, натянутую на точки $Z(Y)$, $Y \in R(A)$.

Теорема 8.1. *Для того чтобы A^* был вектором коэффициентов полинома наилучшего приближения функции $f(t)$ на отрезке $[c, d]$, необходимо и достаточно, чтобы*

$$0 \in L(A^*).$$

Справедливость этого утверждения очевидным образом следует из теоремы 2.2, следствия из теоремы 3.1 и формулы (8.4).

2. Лемма 8.1. *Пусть заданы $n+2$ точки*

$$Y_i = (\xi_i, t_i) \in G, \quad i \in [0 : n+1],$$

такие, что $\xi_i = \pm 1$ и

$$t_0 < t_1 < \dots < t_{n+1}.$$

Следующие два утверждения равносильны:

I. $\text{sign } \xi_i = -\text{sign } \xi_{i+1}, \quad i \in [0 : n].$

II. *Начало координат принадлежит выпуклой оболочке, натянутой на точки $Z(Y_i)$, $i \in [0 : n+1]$.*

Доказательство. Покажем, что из I следует II. Как известно (см. Приложение I, § 1), существуют такие положительные числа $\alpha_i > 0$, $\sum_{i=0}^{n+1} \alpha_i = 1$, что для любого алгебраического полинома $P_n(A, t)$ степени $\leq n$ будет

$$\sum_{i=0}^{n+1} (-1)^i \alpha_i P_n(A, t_i) = 0.$$

Учитывая это и I, получим

$$\sum_{i=0}^{n+1} \alpha_i Z(Y_i) = (\text{sign } \xi_0) \sum_{i=0}^{n+1} (-1)^i \alpha_i \begin{pmatrix} 1 \\ t_i \\ \vdots \\ t_i^n \end{pmatrix} = 0,$$

что равносильно II.

Докажем обратное утверждение. Пусть нашлись неотрицательные числа $\alpha_i \geq 0$, $\sum_{i=0}^{n+1} \alpha_i = 1$ такие, что

$$\sum_{i=0}^{n+1} \alpha_i Z(Y_i) = 0. \quad (8.5)$$

Требуется доказать, что имеет место I.

Заметим, во-первых, что все α_i , $i \in [0 : n+1]$, строго положительны. Действительно, если бы, например, α_{n+1} обращался в нуль, то для любого алгебраического полинома $P_n(A, t)$ степени $\leq n$ в силу (8.5) выполнялось бы равенство

$$\sum_{i=0}^n \alpha_i \xi_i P_n(A, t_i) = 0. \quad (8.6)$$

В частности, (8.6) имело бы место и для интерполяционного полинома $P_n(A_0, t)$:

$$P_n(A_0, t_i) = \alpha_i \xi_i, \quad i \in [0 : n].$$

Подставляя значения $P_n(A_0, t_i)$ в (8.6), получим

$$\sum_{i=0}^n \alpha_i^2 = 0,$$

что противоречит равенству $\sum_{i=0}^n \alpha_i = 1$.

Итак, все $\alpha_i > 0$, $i \in [0 : n+1]$.

Перепишем (8.5) в виде

$$\sum_{i=0}^n \alpha_i \xi_i \begin{pmatrix} 1 \\ t_i \\ \vdots \\ t_i^n \end{pmatrix} = -\alpha_{n+1} \xi_{n+1} \begin{pmatrix} 1 \\ t_{n+1} \\ \vdots \\ t_{n+1}^n \end{pmatrix}.$$

Имеем по формуле Крамера для $i \in [0 : n]$

$$\alpha_i \xi_i = -\alpha_{n+1} \xi_{n+1} \frac{V(t_0, \dots, t_{i-1}, t_{n+1}, t_{i+1}, \dots, t_n)}{V(t_0, \dots, t_n)},$$

где $V(u_0, \dots, u_n)$ — определитель Вандермонда.

Учитывая, что

$$\begin{aligned} V(t_0, \dots, t_{i-1}, t_{n+1}, t_{i+1}, \dots, t_n) = \\ = (-1)^{n-i} V(t_0, \dots, t_{i-1}, t_{i+1}, \dots, t_{n+1}), \end{aligned}$$

получим

$$\alpha_i \xi_i = -(-1)^{n-i} \alpha_{n+1} \xi_{n+1} \frac{V(t_0, \dots, t_{i-1}, t_{i+1}, \dots, t_{n+1})}{V(t_0, \dots, t_n)}. \quad (8.7)$$

Поскольку числа t_i упорядочены по возрастанию, определители в формуле (8.7) положительны. Кроме того, положительны числа α_i , $i \in [0 : n+1]$.

Отсюда следует, что для всех $i \in [0 : n]$

$$\xi_i = -(-1)^{n-i} \xi_{n+1}.$$

Но это равносильно I.

Лемма доказана.

Положим

$$\Delta(A, t) = f(t) - P_n(A, t).$$

Теорема П. Л. Чебышева. Для того чтобы полином $P_n(A^*, t)$ был полиномом наилучшего приближения функции $f(t)$ на отрезке $[c, d]$, необходимо и достаточно, чтобы нашлись $n+2$ точки t_i ,

$$c \leq t_0 < t_1 < \dots < t_{n+1} \leq d,$$

в которых уклонение $\Delta(A^*, t)$ достигало бы наибольшего по абсолютной величине значения с последовательной переменной знака, т. е. чтобы

$$|\Delta(A^*, t_i)| = \max_{c \leq t \leq d} |\Delta(A^*, t)| \quad (8.8)$$

и

$$\text{sign } \Delta(A^*, t_i) = -\text{sign } \Delta(A^*, t_{i+1}). \quad (8.9)$$

Доказательство. Покажем, что эта теорема является следствием теоремы 8.1.

Достаточность. Предположим, что нашлись $n+2$ точки t_i , $i \in [0 : n+1]$, с требуемыми свойствами. Тогда, положив

$$\xi_i = -\text{sign } \Delta(A^*, t_i),$$

закключаем, что точки $Y_i \stackrel{df}{=} (\xi_i, t_i)$ принадлежат $R(A^*)$. Действительно, в силу (8.2) и (8.8)

$$F(A^*, Y_i) = -\xi_i \Delta(A^*, t_i) = |\Delta(A^*, t_i)| = \\ = \max_{c \leq t \leq d} |\Delta(A^*, t)| = \max_{Y \in G} F(A^*, Y).$$

Далее из (8.9) следует, что

$$\text{sign } \xi_i = -\text{sign } \xi_{i+1}.$$

Воспользуемся леммой 8.1. Получим, что начало координат принадлежит выпуклой оболочке, натянутой на точки $Z(Y_i)$, $i \in [0 : n+1]$, $Y_i \in R(A^*)$. Тем более начало координат принадлежит $L(A^*)$ (см. определение $L(A)$).

Отсюда и из теоремы 8.1 следует, что A^* есть вектор коэффициентов полинома наилучшего приближения.

Достаточность доказана.

Необходимость. Пусть

$$0 \in L(A^*) \subset E_{n+1}.$$

В силу леммы 1.1 Приложения II найдутся r попарно различных точек

$$Z(Y_i), \quad i \in [0 : r-1], \quad 1 \leq r \leq n+2, \quad \text{где } Y_i \in R(A^*),$$

и неотрицательные числа $\alpha_i \geq 0$, $\sum_{i=0}^{r-1} \alpha_i = 1$, такие что

$$\sum_{i=0}^{r-1} \alpha_i Z(Y_i) = 0.$$

Заметим, что точки $Y_i = (\xi_i, t_i)$ попарно различны, а поскольку $Y_i \in R(A^*)$, то попарно различными будут и вторые координаты этих точек t_i , $i \in [0 : r-1]$. Далее, так же как при доказательстве леммы 8.1, показывается, что $r = n+2$.

Теперь нам известно следующее: существуют $n+2$ точки Y_i , $i \in [0 : n+1]$, с попарно различными вторыми координатами такие, что

$$\sum_{i=0}^{n+1} \alpha_i Z(Y_i) = 0; \quad \alpha_i \geq 0, \quad \sum_{i=0}^{n+1} \alpha_i = 1.$$

Будем считать, что

$$t_0 < t_1 < \dots < t_{n+1}.$$

Тогда из леммы 8.1 следует

$$\text{sign } \xi_i = -\text{sign } \xi_{i+1}. \quad (8.10)$$

Поскольку $Y_i \in R(A^*)$, то, учитывая (8.2), получим

$$\begin{aligned} \max_{c \leq t \leq d} |f(t) - P_n(A^*, t)| &= \max_{Y \in G} F(A^*, Y) = \\ &= F(A^*, Y_i) = -\xi_i(f(t_i) - P_n(A^*, t_i)). \end{aligned}$$

Отсюда и из (8.10) очевидным образом следуют равенства (8.8) и (8.9).

Теорема доказана.

3. Теорема Валле-Пуссена. Если в $n+2$ точках t_i отрезка $[c, d]$,

$$c \leq t_0 < t_1 < \dots < t_{n+1} \leq d,$$

уклонение $\Delta(A_0, t)$ попеременно меняет знак, то

$$\rho \geq \min_{i \in [0:n+1]} |\Delta(A_0, t_i)|.$$

Доказательство. Положим

$$\xi_i = -\text{sign } \Delta(A_0, t_i)$$

и рассмотрим точки $Y_i = (\xi_i, t_i)$, $i \in [0:n+1]$. В силу леммы 8.1 начало координат пространства E_{n+1} принадлежит выпуклой оболочке, натянутой на точки $Z(Y_i)$, $i \in [0:n+1]$. Отсюда и из (8.4) следует, что выполняется неравенство

$$\min_{\|g\|=1} \max_{i \in [0:n+1]} \left(\frac{\partial F(A_0, Y_i)}{\partial A}, g \right) \geq 0.$$

Учитывая следствие 2 из теоремы 2.3, окончательно имеем

$$\begin{aligned} \rho &\geq \min_{i \in [0:n+1]} F(A_0, Y_i) = \min_{i \in [0:n+1]} [-\xi_i \Delta(A_0, t_i)] = \\ &= \min_{i \in [0:n+1]} |\Delta(A_0, t_i)|. \end{aligned}$$

Теорема доказана.

§ 1. Разделенные разности

Пусть задана таблица значений некоторой функции:

$$y_k = y(t_k), \quad k \in [0 : N].$$

Значения аргумента — *узлы* — предполагаются попарно различными; никаких других ограничений на них не накладывается.

Отношения

$$y[t_s, t_{s+1}] = \frac{y_{s+1} - y_s}{t_{s+1} - t_s}, \quad s \in [0 : N - 1],$$

называются *разделенными разностями первого порядка*.

Разделенные разности второго порядка определяются следующим образом:

$$y[t_s, t_{s+1}, t_{s+2}] = \frac{y[t_{s+1}, t_{s+2}] - y[t_s, t_{s+1}]}{t_{s+2} - t_s}, \quad s \in [0 : N - 2].$$

Вообще *разделенные разности n -го порядка* ($n \leq N$) получаются из разделенных разностей $(n - 1)$ -го порядка с помощью рекуррентного соотношения

$$\begin{aligned} y[t_s, t_{s+1}, \dots, t_{s+n-1}, t_{s+n}] = \\ = \frac{y[t_{s+1}, \dots, t_{s+n}] - y[t_s, \dots, t_{s+n-1}]}{t_{s+n} - t_s}, \quad s \in [0 : N - n]. \end{aligned}$$

Ниже приведена характерного вида таблица, в которой содержатся значения функции вместе со всеми разделенными разностями ($N = 3$).

t	y	Разделенные разности		
		I	II	III
t_0	y_0	$y[t_0, t_1]$ $y[t_1, t_2]$ $y[t_2, t_3]$	$y[t_0, t_1, t_2]$ $y[t_1, t_2, t_3]$	$y[t_0, t_1, t_2, t_3]$
t_1	y_1			
t_2	y_2			
t_3	y_3			

Известно следующее представление разделенных разностей непосредственно через значения функции:

$$y[t_s, t_{s+1}, \dots, t_{s+n}] = \sum_{k=0}^n \frac{y_{s+k}}{(t_{s+k} - t_s) \dots (t_{s+k} - t_{s+n-1}) \dots (t_{s+k} - t_{s+n})}, \quad (1.1)$$

где $s \in [0 : N - n]$ и

$$(t_{s+k} - t_s) \dots (t_{s+k} - t_{s+n-1}) \dots (t_{s+k} - t_{s+n}) = \prod_{i=0, i \neq k}^n (t_{s+k} - t_{s+i}).$$

Укажем основные свойства разделенных разностей.

1. Разделенная разность есть *симметрическая функция* своих аргументов, т. е.

$$\begin{aligned} y[t_s, \dots, t_{s+i}, \dots, t_{s+j}, \dots, t_{s+n}] = \\ = y[t_s, \dots, t_{s+j}, \dots, t_{s+i}, \dots, t_{s+n}], \\ i, j \in [0 : n], \quad i \neq j. \end{aligned}$$

2. Пусть заданы две таблицы

$$y_k = y(t_k), \quad z_k = z(t_k), \quad k \in [0 : N].$$

Составим еще одну таблицу

$$W_k = W(t_k) \stackrel{\text{def}}{=} \alpha y(t_k) + \beta z(t_k),$$

где α и β — произвольные постоянные. Тогда

$$\begin{aligned} W[t_s, t_{s+1}, \dots, t_{s+n}] = \alpha y[t_s, t_{s+1}, \dots, t_{s+n}] + \\ + \beta z[t_s, t_{s+1}, \dots, t_{s+n}]. \end{aligned}$$

3. Пусть $t_0, t_1, \dots, t_{n+1}$ — произвольный набор узлов,

$$y(t) = t^{n+1} \quad \text{и} \quad y_k = y(t_k), \quad k \in [0 : n+1].$$

Тогда

$$y[t_0, t_1, \dots, t_{n+1}] = 1.$$

4. Если в условиях предыдущего пункта $y(t) = Q_n(t)$, где $Q_n(t)$ — произвольный алгебраический полином степени $\leq n$, то

$$y[t_0, t_1, \dots, t_{n+1}] = 0.$$

Учитывая (1.1), этот результат можно переписать так:

$$\sum_{k=0}^{n+1} \frac{Q_n(t_k)}{(t_k - t_0) \dots (t_k - t_{k-1}) \dots (t_k - t_{n+1})} = 0. \quad (1.2)$$

Положим для $k \in [0 : n+1]$

$$\alpha_k = \frac{(-1)^k}{(t_k - t_0) \dots (t_k - t_{k-1}) \dots (t_k - t_{n+1})} \cdot \sum_{i=0}^{n+1} \frac{(-1)^i}{(t_i - t_0) \dots (t_i - t_{i-1}) \dots (t_i - t_{n+1})}.$$

Очевидно, $\sum_{k=0}^{n+1} \alpha_k = 1$. С помощью чисел α_k равенство (1.2) можно переписать в следующем эквивалентном виде:

$$\sum_{k=0}^{n+1} (-1)^k \alpha_k Q_n(t_k) = 0.$$

Если узлы t_k упорядочены по возрастанию

$$t_0 < t_1 < \dots < t_{n+1},$$

то, как легко проверить, числа α_k , $k \in [0 : n+1]$, положительны

$$\alpha_k = \frac{1}{(t_k - t_0) \dots (t_k - t_{k-1}) \dots (t_k - t_{n+1})} \cdot \sum_{i=0}^{n+1} \frac{1}{(t_i - t_0) \dots (t_i - t_{i-1}) \dots (t_i - t_{n+1})} > 0.$$

§ 2. Интерполяционные полиномы

1. Алгебраический полином $L_n(t)$ степени $\leq n$,

$$L_n(t) = b_0 + b_1 t + \dots + b_n t^n,$$

удовлетворяющий соотношениям

$$L_n(t_k) = y_k, \quad k \in [0 : n],$$

называют *интерполяционным полиномом таблицы* (t_k, y_k) , $k \in [0 : n]$.

Если узлы таблицы t_k попарно различны, то интерполяционный полином существует и единствен. При этом

$$b_i = \frac{\Delta_i}{\Delta}, \quad i \in [0 : n],$$

где Δ — определитель Вандермонда,

$$\Delta \stackrel{df}{=} \begin{vmatrix} 1 & t_0 & \dots & t_0^n \\ 1 & t_1 & \dots & t_1^n \\ \dots & \dots & \dots & \dots \\ 1 & t_n & \dots & t_n^n \end{vmatrix} = \prod_{n \geq k > j \geq 0} (t_k - t_j),$$

а определитель Δ_i получается из Δ заменой i -го столбца на столбец $\{y_k\}$, $k \in [0 : n]$.

Если разложить определитель Δ_i по элементам i -го столбца, то придем к следующей формуле:

$$b_i = \frac{\sum_{k=0}^n y_k \Delta_{ki}}{\Delta}, \quad i \in [0 : n], \quad (2.1)$$

где Δ_{ki} есть алгебраическое дополнение элемента определителя Δ , стоящего на пересечении k -й строки и i -го столбца. Заметим, что числа Δ и Δ_{ki} зависят только от узлов интерполяции $\{t_k\}$.

Пусть

$$P_n(t) = a_0 + a_1 t + \dots + a_n t^n$$

и $y_k = P_n(t_k)$, $k \in [0 : n]$. В этом случае интерполяционный полином $L_n(t)$ в силу его единственности совпадает с $P_n(t)$.

Учитывая (2.1), получим

$$a_i = \frac{\sum_{k=0}^n P_n(t_k) \Delta_{ki}}{\Delta}, \quad i \in [0 : n]. \quad (2.2)$$

Отсюда следует, в частности, что если $P_n(t)$ обращается в нуль в $n+1$ точке, то все его коэффициенты равны нулю.

2. В дальнейшем через A будем обозначать вектор коэффициентов полинома $P_n(A, t) = \sum_{i=0}^n a_i t^i$,

$$A = (a_0, a_1, \dots, a_n).$$

Будем говорить, что *последовательность векторов* $A_s = (a_0^{(s)}, a_1^{(s)}, \dots, a_n^{(s)})$, $s = 1, 2, \dots$, *сходится к вектору* $A^* = (a_0^*, a_1^*, \dots, a_n^*)$ *при* $s \rightarrow \infty$, *и писать*

$$A_s \xrightarrow{s \rightarrow \infty} A^*,$$

если имеет место покоординатная сходимость

$$a_i^{(s)} \xrightarrow{s \rightarrow \infty} a_i^*, \quad i \in [0 : n].$$

Лемма 2.1. *Имеется последовательность полиномов $P_n(A_s, t)$, $s = 1, 2, \dots$, степени $\leq n$ и полином $P_n(A^*, t)$. Для того чтобы*

$$A_s \xrightarrow{s \rightarrow \infty} A^*,$$

достаточно, чтобы для некоторого множества попарно различных точек t_k , $k \in [0 : n]$, выполнялось предельное соотношение

$$P_n(A_s, t_k) \xrightarrow{s \rightarrow \infty} P_n(A^*, t_k).$$

Следует из (2.2).

Лемма 2.2 *(о компактности). Пусть задана последовательность полиномов $P_n(A_s, t)$, $s = 1, 2, \dots$, равномерно ограниченная на множестве, состоящем из $(n+1)$ -й точки t_k , $k \in [0 : n]$,*

$$|P_n(A_s, t_k)| \leq M.$$

Тогда найдутся подпоследовательность векторов $\{A_{s_i}\}$ и вектор A^* такие, что

$$A_{s_i} \xrightarrow{i \rightarrow \infty} A^*.$$

Доказательство. Поскольку при каждом $k \in [0 : n]$ числовые последовательности $P_n(A_s, t_k)$, $s = 1, 2, \dots$, ограничены, то найдется подпоследовательность индексов $\{s_i\}$ такая, что при $k \in [0 : n]$

$$P_n(A_{s_i}, t_k) \xrightarrow{i \rightarrow \infty} y_k^*.$$

Пусть $P_n(A^*, t)$ — интерполяционный полином, определяемый из соотношений

$$P_n(A^*, t_k) = y_k^*, \quad k \in [0 : n].$$

Теперь имеем

$$P_n(A_{s_i}, t_k) \xrightarrow{i \rightarrow \infty} P_n(A^*, t_k), \quad k \in [0 : n].$$

В силу леммы 2.1 отсюда следует, что

$$A_{s_i} \xrightarrow{i \rightarrow \infty} A^*.$$

Лемма доказана.

3. Возвращаясь к интерполяционной задаче,

$$L_n(t_k) = y_k, \quad k \in [0 : n],$$

запишем формулу Лагранжа для интерполяционного полинома

$$L_n(t) = \sum_{k=0}^n y_k l_k(t),$$

где

$$l_k(t) = \frac{(t-t_0)(t-t_1) \dots (t-t_{k-1})(t-t_{k+1}) \dots (t-t_n)}{(t_k-t_0)(t_k-t_1) \dots (t_k-t_{k-1})(t_k-t_{k+1}) \dots (t_k-t_n)}.$$

В частности, для любого алгебраического полинома $P_n(A, t)$ степени $\leq n$ справедливо представление

$$P_n(A, t) = \sum_{k=0}^n P_n(A, t_k) l_k(t). \quad (2.3)$$

Лемма 2.3. Пусть задана последовательность алгебраических полиномов $P_n(A_s, t)$, $s = 1, 2, \dots$, равно-

мерно ограниченная на множестве, состоящем из $(n+1)$ -й точки отрезка $[c, d]$,

$$|P_n(A_s, t_k)| \leq M, \quad k \in [0:n], \quad s = 1, 2, \dots$$

Тогда при любом натуральном r производные $P_n^{(r)}(A_s, t)$ ограничены на отрезке $[c, d]$ равномерно по $s = 1, 2, \dots$,

$$|P_n^{(r)}(A_s, t)| \leq M^{(r)}.$$

Доказательство. В силу (2.3)

$$|P_n^{(r)}(A_s, t)| = \left| \sum_{k=0}^n P_n(A_s, t_k) l_k^{(r)}(t) \right| \leq M \sum_{k=0}^n |l_k^{(r)}(t)|.$$

Заключение леммы следует теперь из того, что правая часть полученного неравенства не зависит от s .

Лемма доказана.

Лемма 2.4 (о непрерывной зависимости). Пусть интерполяционный полином $P_n(A^*, t)$ и последовательность интерполяционных полиномов $P_n(A_s, t)$, $s = 1, 2, \dots$, определяются соотношениями

$$\begin{aligned} P_n(A^*, t_k^*) &= y_k^*, \\ P_n(A_s, t_k^{(s)}) &= y_k^{(s)}, \\ k &\in [0:n], \quad s = 1, 2, \dots \end{aligned}$$

Если при $s \rightarrow \infty$ будет

$$t_k^{(s)} \rightarrow t_k^*, \quad y_k^{(s)} \rightarrow y_k^*, \quad k \in [0:n],$$

то $A_s \rightarrow A^*$.

Утверждение леммы непосредственно следует из (2.3).

4. В заключение этого параграфа выпишем еще формулу Ньютона для интерполяционного полинома

$$L_n(t) = y_0 + \sum_{k=1}^n y[t_0, t_1, \dots, t_k] (t - t_0)(t - t_1) \dots (t - t_{k-1}).$$

ПРИЛОЖЕНИЕ II

ВЫПУКЛЫЕ МНОЖЕСТВА И ВЫПУКЛЫЕ ФУНКЦИИ

§ 1. Выпуклые оболочки. Теорема отделимости

1. Пусть E_n обозначает n -мерное евклидово пространство векторов $X = (x_1, \dots, x_n)$. Отдельно запишем представление для нуля пространства E_n и вектора с индексом

$$0 = (0, \dots, 0), \quad X_k = (x_1^{(k)}, \dots, x_n^{(k)}).$$

Введем обозначения

$$(X_1, X_2) = \sum_{i=1}^n x_i^{(1)} x_i^{(2)}, \quad \|X\| = \sqrt{(X, X)}.$$

Известно, что для любых векторов $X_1, X_2 \in E_n$ справедливо неравенство (Коши — Буняковского)

$$|(X_1, X_2)| \leq \|X_1\| \cdot \|X_2\|.$$

По определению соотношения

$$X_s \xrightarrow{s \rightarrow \infty} X^* \quad \text{и} \quad \|X_s - X^*\| \xrightarrow{s \rightarrow \infty} 0$$

равносильны.

Система векторов $X_1, \dots, X_r$, принадлежащих E_n , называется *линейно независимой*, если равенство

$$\sum_{k=1}^r \alpha_k X_k = 0$$

возможно тогда и только тогда, когда все коэффициенты α_k , $k \in [1:r]$, равны нулю. Как известно, при $r \geq n + 1$ любая система векторов $X_1, \dots, X_r$ из E_n линейно зависима. Это значит, что найдутся числа $\beta_1, \dots, \beta_r$, не равные нулю одновременно, $\sum_{k=1}^r |\beta_k| > 0$, для которых

$$\sum_{k=1}^r \beta_k X_k = 0. \quad (1.1)$$

Если $r \geq n + 2$, то одновременно с (1.1) можно добиться выполнения равенства

$$\sum_{k=1}^r \beta_k = 0. \quad (1.2)$$

Действительно, введя в рассмотрение векторы $\bar{X}_k = (x_1^{(k)}, \dots, x_n^{(k)}, 1) \in E_{n+1}$, $k \in [1:r]$, $r \geq n + 2$,

найдем, в силу предыдущего, числа β_k , $\sum_{k=1}^r |\beta_k| > 0$, такие, что

$$\sum_{k=1}^r \beta_k \bar{X}_k = 0. \quad (1.3)$$

Легко понять, что равенство (1.3) эквивалентно совокупности равенств (1.1) и (1.2).

2. Через $S_\delta(X_0)$ будем обозначать δ -окрестность точки X_0 :

$$S_\delta(X_0) = \{X \in E_n \mid \|X - X_0\| < \delta\}, \quad \delta > 0.$$

Множество $G \subset E_n$ называется *открытым*, если для любой точки $X_0 \in G$ найдется такое $\delta > 0$, что

$$S_\delta(X_0) \subset G.$$

Множество $F \subset E_n$ называется *замкнутым*, если из соотношения

$$X_s \xrightarrow{s \rightarrow \infty} X^*, \quad X_s \in F, \quad s = 1, 2, \dots,$$

следует, что $X^* \in F$.

Замыканием произвольного множества $H \subset E_n$ на вается множество точек X , допускающих представление

$$X = \lim_{s \rightarrow \infty} X_s; \quad X_s \in H, \quad s = 1, 2, \dots$$

Точка X_0 называется *внутренней точкой* множества H , если при некотором $\delta > 0$

$$S_\delta(X_0) \subset H.$$

Точка X_0 называется *граничной точкой* множества H , если ее δ -окрестность $S_\delta(X_0)$ при любом $\delta > 0$ содержит хотя бы одну точку, не принадлежащую H , и

хотя бы одну точку, принадлежащую H и отличную от X_0 .

Множество L называется *выпуклым*, если вместе с любыми двумя точками $X_1, X_2 \in L$ оно содержит отрезок, соединяющий эти точки:

$$\alpha X_1 + (1 - \alpha) X_2 \in L, \quad 0 \leq \alpha \leq 1.$$

Замкнутое выпуклое множество L называется *строго выпуклым*, если для любых $X_1, X_2 \in L$ и любого $\alpha \in (0, 1)$ точка $\alpha X_1 + (1 - \alpha) X_2$ является внутренней точкой множества L .

Множество H называется *ограниченным*, если существует такое число $M > 0$, что для всех $X \in H$ будет

$$\|X\| \leq M.$$

3. *Выпуклая оболочка* произвольного множества $G \subseteq E_n$, обозначаемая $\text{co } G$ (или $\text{co}(G)$), определяется следующим образом:

$$\text{co } G = \left\{ X = \sum_{k=1}^r \alpha_k X_k \mid X_k \in G; \alpha_k \geq 0, \sum_{k=1}^r \alpha_k = 1, \right. \\ \left. r - \text{произвольное натуральное число} \right\}. \quad (1.4)$$

Лемма 1.1. *Любой вектор $X \in \text{co } G$ может быть представлен в виде выпуклой комбинации не более чем $(n + 1)$ -го вектора множества G .*

Доказательство. Пусть для некоторого $X \in \text{co } G$ в представлении (1.4) после выбрасывания слагаемых с $\alpha_k = 0$ оказалось, что $r \geq n + 2$:

$$X = \sum_{k=1}^r \alpha_k X_k; \quad X_k \in G; \quad \alpha_k > 0, \quad \sum_{k=1}^r \alpha_k = 1.$$

Согласно замечанию, сделанному в п. 1, найдутся числа β_k , $\sum_{k=1}^r |\beta_k| > 0$ такие, что

$$\sum_{k=1}^r \beta_k X_k = 0, \quad \sum_{k=1}^r \beta_k = 0.$$

Положим $\varepsilon = \min_{\{\beta_k > 0\}} \frac{\alpha_k}{\beta_k} > 0$ и $\bar{\alpha}_k = \alpha_k - \varepsilon \beta_k$, $k \in [1 : r]$.

Очевидно,

$$X = \sum_{k=1}^r \bar{\alpha}_k X_k, \quad \sum_{k=1}^r \bar{\alpha}_k = 1.$$

Рассматривая отдельно положительные и неположительные β_k , найдем, что $\bar{\alpha}_k \geq 0$, $k \in [1 : r]$; при этом хотя бы одно $\bar{\alpha}_k$ обращается в нуль.

Таким образом, вектор X представлен как выпуклая комбинация не более чем $(r-1)$ -го вектора из G . Повторяя описанный процесс, придем к представлению, в котором $r \leq n+1$.

Лемма доказана.

Лемма 1.2. Если $G \subset E_n$ — ограниченное замкнутое множество, то со G есть ограниченное замкнутое выпуклое множество.

Доказательство. Если все компоненты векторов из G ограничены по модулю числом M , то все компоненты векторов из со G ограничены по модулю тем же числом M . Действительно, при всех $v \in [1 : n]$ в силу (1.4) имеем

$$|x_v| = \left| \sum_{k=1}^r \alpha_k x_v^{(k)} \right| \leq \sum_{k=1}^r \alpha_k |x_v^{(k)}| \leq M \sum_{k=1}^r \alpha_k = M.$$

Отсюда следует ограниченность со G .

Докажем выпуклость со G . Пусть $X \in \text{со } G$ и $Y \in \text{со } G$, причем

$$X = \sum_{k=1}^{r_1-1} \alpha_k X_k, \quad Y = \sum_{k=r_1}^{r_2} \alpha_k X_k; \quad X_k \in G;$$

$$\sum_{k=1}^{r_1-1} \alpha_k = \sum_{k=r_1}^{r_2} \alpha_k = 1; \quad \alpha_k \geq 0, \quad k \in [1 : r_2].$$

Покажем, что $\alpha X + (1-\alpha)Y \in \text{со } G$ при всех $\alpha \in [0, 1]$. Имеем

$$\alpha X + (1-\alpha)Y = \sum_{k=1}^{r_1-1} \alpha \alpha_k X_k + \sum_{k=r_1}^{r_2} (1-\alpha) \alpha_k X_k = \sum_{k=1}^{r_2} \bar{\alpha}_k X_k,$$

где

$$\bar{\alpha}_k = \begin{cases} \alpha \alpha_k, & \text{если } k \in [1 : r_1 - 1], \\ (1-\alpha) \alpha_k, & \text{если } k \in [r_1 : r_2]. \end{cases}$$

Очевидно, $\bar{\alpha}_k \geq 0$, $k \in [1 : r_2]$, и

$$\sum_{k=1}^{r_2} \bar{\alpha}_k = \alpha \sum_{k=1}^{r_1-1} \alpha_k + (1-\alpha) \sum_{k=r_1}^{r_2} \alpha_k = 1.$$

Значит, по определению точка $\alpha X + (1-\alpha)Y$ при всех $\alpha \in [0, 1]$ принадлежит $\text{co } G$, чем доказана выпуклость множества $\text{co } G$.

Осталось доказать замкнутость $\text{co } G$.

Пусть $X_s \xrightarrow{s \rightarrow \infty} X^*$, $X_s \in \text{co } G$, $s = 1, 2, \dots$. Покажем, что $X^* \in \text{co } G$. Представим каждый вектор X_s в виде выпуклой комбинации не более чем $(n+1)$ -го вектора из G :

$$X_s = \sum_{k=1}^{r_s} \alpha_{sk} X_{sk}; \quad \alpha_{sk} \geq 0, \quad \sum_{k=1}^{r_s} \alpha_{sk} = 1; \quad (1.5)$$

$$X_{sk} \in G, \quad k \in [1 : r_s], \quad s = 1, 2, \dots$$

Пополняя, если нужно, представление (1.5) произвольными векторами из G с нулевыми коэффициентами, добьемся того, чтобы $r_s = n+1$, $s = 1, 2, \dots$

Пользуясь тем, что G — ограниченное замкнутое множество и $0 \leq \alpha_{sk} \leq 1$, $k \in [1 : n+1]$, $s = 1, 2, \dots$, выделим такую подпоследовательность $\{s_i\}$, что для всех $k \in [1 : n+1]$ будет

$$\begin{aligned} X_{s_i k} &\xrightarrow{s_i \rightarrow \infty} X_{0k}, \\ \alpha_{s_i k} &\xrightarrow{s_i \rightarrow \infty} \alpha_{0k}. \end{aligned}$$

На основании (1.5) получим

$$X^* = \sum_{k=1}^{n+1} \alpha_{0k} X_{0k},$$

причем

$$X_{0k} \in G, \quad \alpha_{0k} \geq 0, \quad \sum_{k=1}^{n+1} \alpha_{0k} = 1.$$

Значит, $X^* \in \text{co } G$.

Лемма доказана.

Нетрудно проверить, что если G — выпуклое множество, то

$$\text{co } G = G.$$

Действительно, по определению выпуклой оболочки $G \subset \text{co } G$.

Покажем, что $\text{co } G \subset G$. Пусть $X \in \text{co } G$, $X = \sum_{k=1}^r \alpha_k X_k$;

$X_k \in G$, $\alpha_k > 0$, $\sum_{k=1}^r \alpha_k = 1$. Если $r = 1$ или $r = 2$, то очевидно, $X \in G$.

В общем случае следует воспользоваться определением выпуклого множества и соотношением

$$X = \alpha_r X_r + (1 - \alpha_r) \left\{ \sum_{k=1}^{r-1} \frac{\alpha_k}{1 - \alpha_r} X_k \right\}, \quad \sum_{k=1}^{r-1} \frac{\alpha_k}{1 - \alpha_r} = 1.$$

4. Лемма 1.3. Пусть $\Omega \subset E_n$ — некоторое замкнутое выпуклое множество. Тогда найдется точка $X^* \in \Omega$ такая, что для всех $X \in \Omega$ будет

$$(X, X^*) \geq (X^*, X^*).$$

Доказательство. Рассмотрим функцию $F(X) = (X, X)$. Нетрудно показать, что $F(X)$ достигает на Ω своего наименьшего значения в некоторой точке $X^* \in \Omega$:

$$\min_{X \in \Omega} F(X) = F(X^*) \geq 0. \quad (1.6)$$

Докажем, что для любого $X \in \Omega$

$$(X, X^*) \geq (X^*, X^*). \quad (1.7)$$

Действительно, в силу (1.6) и выпуклости Ω при $0 < \alpha < 1$ будем иметь

$$(X^* + \alpha(X - X^*), X^* + \alpha(X - X^*)) \geq (X^*, X^*).$$

Отсюда

$$2(X - X^*, X^*) + \alpha(X - X^*, X - X^*) \geq 0.$$

Теперь легко доказать от противного, что

$$(X - X^*, X^*) \geq 0.$$

Последнее же неравенство равносильно (1.7).

Лемма доказана.

Замечание. В силу (1.6) точка X^* имеет наименьшую норму среди всех точек $X \in \Omega$. В связи с этим X^* называют ближайшей к началу координат точкой множества Ω .

Рассмотрим теперь два замкнутых выпуклых множества Ω_1 и Ω_2 , причем одно из них предполагается ограниченным. Пусть

$$\rho = \inf_{X \in \Omega_1, Y \in \Omega_2} \|X - Y\|.$$

Лемма 1.4. *Существуют такие точки $X_0 \in \Omega_1$ и $Y_0 \in \Omega_2$, что*

$$\|X_0 - Y_0\| = \rho.$$

При этом для любых $X \in \Omega_1$ и $Y \in \Omega_2$ будет

$$(X - Y, X_0 - Y_0) \geq (X_0 - Y_0, X_0 - Y_0).$$

Доказательство. Введем множество $\Omega = \Omega_1 - \Omega_2$:

$$\Omega = \{X - Y \mid X \in \Omega_1, Y \in \Omega_2\}.$$

Нетрудно показать, что Ω является замкнутым выпуклым множеством. Остается воспользоваться леммой 1.3 и замечанием к ней.

Лемма доказана.

Теорема 1.1 (теорема отделмости). *Пусть $\Omega \subset E_n$ — замкнутое выпуклое множество и $X_0 \notin \Omega$. Тогда найдутся вектор $g_0 \in E_n$, $\|g_0\| = 1$, и число $a > 0$ такие, что для любого $X \in \Omega$ будет*

$$(X - X_0, g_0) \leq -a.$$

Доказательство. Рассмотрим множество Ω_0 :

$$\Omega_0 = \{X - X_0 \mid X \in \Omega\}.$$

Нетрудно проверить, что Ω_0 является замкнутым выпуклым множеством и что $0 \notin \Omega_0$. В силу леммы 1.3 найдется точка $Z^* \in \Omega_0$ ($Z^* \neq 0$) такая, что для всех $Z \in \Omega_0$ будет

$$(Z, Z^*) \geq (Z^*, Z^*).$$

Положим $g_0 = -Z^*/\|Z^*\|$, $a = \|Z^*\| > 0$. Тогда для любого $Z \in \Omega_0$ получим

$$(Z, g_0) = -\frac{1}{\|Z^*\|} (Z, Z^*) \leq -\frac{1}{\|Z^*\|} (Z^*, Z^*) = -a,$$

что равносильно утверждению теоремы.

Теорема доказана.

Следствие. Если $\Omega \subset E_n$ — замкнутое выпуклое множество и X_0 — граничная точка множества Ω , то существует вектор $g_0 \in E_n$, $\|g_0\| = 1$ такой, что

$$(X - X_0, g_0) \leq 0$$

для всех $X \in \Omega$.

Доказательство. Так как X_0 — граничная точка Ω , то найдется последовательность точек $\{X_i\}$ со свойствами

$$X_i \notin \Omega, \|X_0 - X_i\| \xrightarrow{i \rightarrow \infty} 0.$$

Поскольку $X_i \notin \Omega$, то по теореме 1.1 имеем для всех $X \in \Omega$

$$(X - X_i, g_i) \leq -a_i, \quad (1.8)$$

где

$$a_i \stackrel{\text{df}}{=} \min_{Z \in \Omega} \|Z - X_i\| = \|Z_i^* - X_i\| > 0,$$

$$g_i = -\frac{Z_i^* - X_i}{\|Z_i^* - X_i\|}, \quad \|g_i\| = 1.$$

Заметим, что по определению a_i

$$0 < a_i \leq \|X_0 - X_i\|.$$

Отсюда следует соотношение

$$a_i \xrightarrow{i \rightarrow \infty} 0.$$

Не ограничивая общности, будем считать, что последовательность $\{g_i\}$ сходится к вектору g_0 , $\|g_0\| = 1$. Перейдем теперь к пределу при $i \rightarrow \infty$ в неравенстве (1.8). Получим для всех $X \in \Omega$

$$(X - X_0, g_0) \leq 0,$$

что и требовалось доказать.

§ 2. Выпуклые конусы

1. Множество $\Gamma \subset E_n$ называется *конусом*, если вместе с вектором X оно содержит вектор λX , $\lambda \geq 0$. Примерами конусов могут служить $\{0\}$ и E_n .

Пусть $\Omega \subset E_n$ — замкнутое выпуклое множество и $X^* \in \Omega$. Рассмотрим конус

$$\Gamma(X^*) = \{V = \lambda(X - X^*) \mid \lambda > 0, X \in \Omega\}.$$

Очевидно, $0 \in \Gamma(X^*)$. Замыкание конуса $\Gamma(X^*)$ называется *конусом возможных направлений множества Ω в точке X^** и обозначается $\bar{\Gamma}(X^*)$. По определению замыкания (см. § 1, п. 2) точка V принадлежит $\bar{\Gamma}(X^*)$ тогда и только тогда, когда

$$V = \lim_{s \rightarrow \infty} \lambda_s (X_s - X^*), \quad (2.1)$$

где $\lambda_s > 0$, $X_s \in \Omega$, $s = 1, 2, \dots$

Лемма 2.1. *Если Ω — замкнутое выпуклое множество и $X^* \in \Omega$, то $\bar{\Gamma}(X^*)$ является замкнутым выпуклым конусом.*

Доказательство. В силу (2.1) ясно, что $\bar{\Gamma}(X^*)$ является конусом и что этот конус замкнутый. Докажем его выпуклость. Для этого достаточно установить выпуклость $\Gamma(X^*)$.

Пусть $V_1, V_2 \in \Gamma(X^*)$, $V_1 = \lambda_1 (X_1 - X^*)$, $V_2 = \lambda_2 (X_2 - X^*)$, где $\lambda_1, \lambda_2 > 0$, $X_1, X_2 \in \Omega$. Покажем, что при $0 \leq \alpha \leq 1$

$$\alpha V_1 + (1 - \alpha) V_2 \in \Gamma(X^*).$$

Положим $\bar{\lambda} = \alpha \lambda_1 + (1 - \alpha) \lambda_2 > 0$,

$$\alpha_0 = \frac{\alpha \lambda_1}{\bar{\lambda}}, \quad 0 \leq \alpha_0 \leq 1.$$

Поскольку Ω — выпуклое множество, то

$$X_0 \stackrel{\text{def}}{=} \alpha_0 X_1 + (1 - \alpha_0) X_2 \in \Omega.$$

Теперь имеем

$$\begin{aligned} \alpha V_1 + (1 - \alpha) V_2 &= \alpha \lambda_1 (X_1 - X^*) + (1 - \alpha) \lambda_2 (X_2 - X^*) = \\ &= \bar{\lambda} [\alpha_0 (X_1 - X^*) + (1 - \alpha_0) (X_2 - X^*)] = \\ &= \bar{\lambda} (X_0 - X^*) \in \Gamma(X^*). \end{aligned}$$

Лемма доказана.

2. Пусть $\Gamma \subset E_n$ — конус. Введем понятие *сопряженного конуса Γ^+* :

$$\Gamma^+ = \{W \in E_n \mid (W, V) \geq 0 \text{ для всех } V \in \Gamma\}.$$

Заметим, что если $\Gamma = E_n$, то $\Gamma^+ = \{0\}$. Действительно, $0 \in \Gamma^+$. С другой стороны, если бы $W_0 \in \Gamma^+$, $\|W_0\| > 0$, то, положив $V = -W_0 \in \Gamma$, получили бы $(W_0, V) = -\|W_0\|^2 < 0$, что противоречит условию $W_0 \in \Gamma^+$.

Другой пример. Пусть

$$\Gamma = \{X = (x_1, \dots, x_n) \mid x_i \geq 0, i \in [1: n]\}.$$

Нетрудно проверить, что в этом случае $\Gamma^+ = \Gamma$.

Лемма 2.2. Если $\Gamma \subset E_n$ — конус, то $\Gamma^+ \subset E_n$ является замкнутым выпуклым конусом.

Доказательство легко следует из определения сопряженного конуса.

Теорема 2.1. Пусть $\Gamma \subset E_n$ — замкнутый выпуклый конус, $G \subset E_n$ — ограниченное замкнутое выпуклое множество. Для того чтобы Γ и G не имели общих точек, т. е.

$$\Gamma \cap G = \emptyset, \quad (2.2)$$

необходимо и достаточно, чтобы нашелся вектор $W_0 \in \Gamma^+$ такой, что

$$\max_{X \in G} (W_0, X) < 0. \quad (2.3)$$

Доказательство. Достаточность. Поскольку $W_0 \in \Gamma^+$, то для любого $V \in \Gamma$

$$(W_0, V) \geq 0.$$

С другой стороны, в силу (2.3) для любого $X \in G$

$$(W_0, X) < 0.$$

Отсюда следует, что Γ и G не могут иметь общих точек.

Необходимость. Положим $\Omega = \Gamma - G$:

$$\Omega = \{V - X \mid V \in \Gamma, X \in G\}.$$

Нетрудно проверить, что Ω является замкнутым выпуклым множеством, не содержащим начала координат 0. В силу леммы 1.3 найдется вектор $W_0 \in \Omega$, $W_0 \neq 0$, такой, что для всех $W \in \Omega$ будут выполняться неравенства

$$(W, W) \geq (W_0, W_0), \quad (2.4)$$

$$(W, W_0) \geq (W_0, W_0). \quad (2.5)$$

Пусть $W_0 = V_0 - X_0$, $V_0 \in \Gamma$, $X_0 \in G$. Докажем, что

$$(W_0, V_0) = 0. \quad (2.6)$$

Поскольку $V_0 \in \Gamma$, то $\lambda V_0 \in \Gamma$ при всех $\lambda > 0$, и потому

$$W_\lambda \stackrel{\text{df}}{=} \lambda V_0 - X_0 = (\lambda - 1)V_0 + W_0 \in \Omega.$$

Теперь имеем

$$\begin{aligned}(W_\lambda, W_\lambda) &= ((\lambda - 1)V_0 + W_0, (\lambda - 1)V_0 + W_0) = \\ &= (\lambda - 1)[(\lambda - 1)(V_0, V_0) + 2(W_0, V_0)] + (W_0, W_0).\end{aligned}$$

Предположение о том, что $(W_0, V_0) \neq 0$, приводит нас к заключению, что при некотором положительном λ (большем единицы, если $(W_0, V_0) < 0$, и меньшем единицы, если $(W_0, V_0) > 0$) будет выполняться неравенство

$$(W_\lambda, W_\lambda) < (W_0, W_0).$$

Но это противоречит (2.4). Итак, доказана справедливость равенства (2.6).

Подставим в (2.5) $W = V - X_0$, $V \in \Gamma$:

$$(V - X_0, W_0) \geq (W_0, V_0 - X_0).$$

Отсюда, учитывая (2.6), получим для всех $V \in \Gamma$

$$(W_0, V) \geq 0,$$

т. е. $W_0 \in \Gamma^+$. Теперь подставим в (2.5) $W = V_0 - X$, $X \in G$:

$$(V_0 - X, W_0) \geq (W_0, W_0).$$

Учитывая (2.6), получим для всех $X \in G$

$$(W_0, X) \leq -(W_0, W_0).$$

В частности,

$$\max_{X \in G} (W_0, X) < 0.$$

Теорема доказана.

Лемма 2.3. Если $\Gamma \subset E_n$ — замкнутый выпуклый конус, то

$$\Gamma^{++} = \Gamma. \quad (2.7)$$

Доказательство. Пусть $V_0 \in \Gamma$. Тогда для любого $W \in \Gamma^+$ будет

$$(W, V_0) \geq 0.$$

Значит, $V_0 \in \Gamma^{++}$ и, следовательно, $\Gamma \subset \Gamma^{++}$.

Обратное включение будем доказывать от противного. Пусть $V_0 \in \Gamma^{++}$, но $V_0 \notin \Gamma$. Положим $G = \{V_0\}$.

В силу теоремы 2.1 найдется вектор $W_0 \in \Gamma^+$ такой, что $(W_0, V_0) < 0$. Значит, $V_0 \notin \Gamma^{++}$, что противоречит предположению.

Лемма доказана.

Следствие. Если $\Gamma \subset E_n$ — выпуклый конус и $\bar{\Gamma}$ — его замыкание, то

$$\Gamma^{++} = \bar{\Gamma}. \quad (2.8)$$

Действительно, $\bar{\Gamma}$ — замкнутый выпуклый конус, и поэтому в силу (2.7)

$$\bar{\Gamma}^{++} = \bar{\Gamma}.$$

Остается заметить, что $\bar{\Gamma}^+ = \Gamma^+$.

Пусть $\Gamma_i \subset E_n$, $i \in [1 : s]$, — произвольные конусы и

$$\sum_{i=1}^s \Gamma_i \stackrel{\text{df}}{=} \left\{ V = \sum_{i=1}^s V_i \mid V_i \in \Gamma_i, i \in [1 : s] \right\}.$$

Имеет место равенство

$$\left(\sum_{i=1}^s \Gamma_i \right)^+ = \bigcap_{i=1}^s \Gamma_i^+. \quad (2.9)$$

Справедливость этого равенства легко следует из определения сопряженного конуса и того факта, что $0 \in \Gamma_i$, $i \in [1 : s]$.

На основании (2.7), (2.8) и (2.9) заключаем, что если $\Gamma_i \subset E_n$, $i \in [1 : s]$, — замкнутые выпуклые конусы, то

$$\left(\bigcap_{i=1}^s \Gamma_i \right)^+ = \overline{\sum_{i=1}^s \Gamma_i^+}.$$

Действительно,

$$\left(\bigcap_{i=1}^s \Gamma_i \right)^+ = \left(\bigcap_{i=1}^s (\Gamma_i^+)^+ \right)^+ = \left(\sum_{i=1}^s \Gamma_i^+ \right)^{++} = \overline{\sum_{i=1}^s \Gamma_i^+}.$$

3. Пусть G — произвольное множество из E_n . Через $K(G)$ будем обозначать выпуклый конус, натянутый на G :

$$K(G) = \left\{ v = \sum_{k=1}^r \alpha_k V_k \mid V_k \in G, \alpha_k \geq 0, k \in [1 : r], \right. \\ \left. r - \text{любое натуральное число} \right\}.$$

Лемма 2.4. *Любая точка $V \in K(G)$ может быть представлена в виде*

$$V = \sum_{k=1}^r \alpha_k V_k; \quad V_k \in G, \quad \alpha_k > 0, \quad k \in [1:r], \quad (2.10)$$

где $1 \leq r \leq n$ и векторы V_k , $k \in [1:r]$, линейно независимы (см. § 1, п. 1).

Доказательство. Пусть $V \in K(G)$. Среди всех представлений вектора V в виде (2.10) выберем представление с наименьшим r . Докажем, что в этом случае векторы $V_1, \dots, V_r$ линейно независимы. Допустим противное. Тогда найдутся числа β_k , $\sum_{k=1}^r |\beta_k| > 0$, для которых

$$\sum_{k=1}^r \beta_k V_k = 0.$$

Положим

$$\varepsilon = \min_{\{k | \beta_k > 0\}} \frac{\alpha_k}{\beta_k}, \quad \bar{\alpha}_k = \alpha_k - \varepsilon \beta_k, \quad k \in [1:r].$$

Очевидно,

$$V = \sum_{k=1}^r \bar{\alpha}_k V_k, \quad (2.11)$$

причем все коэффициенты $\bar{\alpha}_k$ неотрицательны и хотя бы один из них обращается в нуль. Выбрасывая слагаемые с $\bar{\alpha}_k = 0$, получим на основании (2.11) представление для V вида (2.10), в котором число слагаемых меньше r . Но это противоречит предположению. Итак, векторы $V_1, \dots, V_r$ линейно независимы.

Неравенство $1 \leq r \leq n$ следует из того, что в E_n не может быть более чем n линейно независимых векторов.

Лемма доказана.

Лемма 2.5. *Пусть $G \subset E_n$ — ограниченное замкнутое множество и начало координат не принадлежит выпуклой оболочке множества G : $0 \notin \text{co}G$. Тогда $K(G)$ — замкнутый выпуклый конус.*

Доказательство. Очевидно, что $K(G)$ — выпуклый конус. Докажем его замкнутость.

Пусть

$$V_i \xrightarrow{i \rightarrow \infty} V^*; \quad V_i \in K(G), \quad i = 1, 2, \dots \quad (2.12)$$

Требуется доказать, что $V^* \in K(G)$.

Нетрудно проверить справедливость следующего утверждения: для того чтобы $V \in K(G)$, необходимо и достаточно, чтобы $V = \lambda \bar{V}$, где $\lambda \geq 0$ и $\bar{V} \in \text{co } G$.

В связи с этим соотношение (2.12) можно переписать в виде

$$\lambda_i \bar{V}_i \xrightarrow{i \rightarrow \infty} V^*; \quad \lambda_i \geq 0, \quad \bar{V}_i \in \text{co } G. \quad (2.13)$$

Поскольку $0 \notin \text{co } G$, то найдется $c > 0$ такое, что

$$\|\bar{V}_i\| \geq c$$

равномерно по i . Отсюда и из (2.13) следует, что последовательность $\{\lambda_i\}$ ограничена.

Выделим теперь такую подпоследовательность индексов $\{i_s\}$, для которой

$$\lambda_{i_s} \xrightarrow{i_s \rightarrow \infty} \bar{\lambda}, \quad (2.14)$$

$$\bar{V}_{i_s} \xrightarrow{i_s \rightarrow \infty} \bar{V}. \quad (2.15)$$

В силу леммы 1.2 множество $\text{co } G$ замкнутое, поэтому $\bar{V} \in \text{co } G$. Учитывая (2.13), (2.14) и (2.15), получим

$$V^* = \bar{\lambda} \bar{V}; \quad \bar{\lambda} \geq 0, \quad \bar{V} \in \text{co } G.$$

Значит, $V^* \in K(G)$.

Лемма доказана.

4. Предположение $0 \notin \text{co } G$ в лемме 2.5 существенно, о чем свидетельствует следующий пример:

Пусть $G = \{X = (x_1, x_2) | (x_1 - 1)^2 + x_2^2 \leq 1\}$. В этом случае $0 \in \text{co } G$. Конус $K(G)$ имеет вид

$$K(G) = \{0\} \cup \{X = (x_1, x_2) | x_1 > 0\}.$$

Получили, что $K(G)$ не является замкнутым множеством.

Однако если G состоит из конечного числа точек, то требование $0 \notin \text{co } G$ в лемме 2.5 может быть опущено. Справедлива

Лемма 2.6. Если множество G состоит из конечного числа точек, то $K(G)$ является замкнутым выпуклым конусом.

Доказательство. Пусть

$$V_s \xrightarrow{s \rightarrow \infty} V^*; \quad V_s \in K(G), \quad s = 1, 2, \dots \quad (2.16)$$

Требуется доказать, что $V^* \in K(G)$.

В силу леммы 2.4 справедливо представление

$$V_s = \sum_{k=1}^{r_s} \lambda_{ks} V_{ks}; \quad \lambda_{ks} > 0, \quad V_{ks} \in G,$$

причем система векторов $V_{1s}, \dots, V_{r_s s}$ линейно независима. Поскольку множество G конечно, то среди наборов $\{V_{ks}\}_{k=1}^{r_s}$, $s = 1, 2, \dots$, хотя бы один набор встретится бесконечное число раз. Поэтому, не ограничивая общности, будем писать

$$V_s = \sum_{k=1}^r \lambda_{ks} V'_k, \quad \lambda_{ks} > 0, \quad V'_k \in G. \quad (2.17)$$

Существенно, что система векторов $V'_1, \dots, V'_r$ линейно независима.

Покажем, что все последовательности $\{\lambda_{ks}\}$, $k \in [1:r]$, ограничены. Допустим противное. Тогда для некоторой подпоследовательности индексов $\{s_i\}$ будем иметь

$$\gamma_{s_i} \stackrel{\text{df}}{=} \sum_{k=1}^r |\lambda_{ks_i}| \xrightarrow{s_i \rightarrow \infty} \infty. \quad (2.18)$$

Из (2.17) следует равенство

$$\sum_{k=1}^r \lambda'_{ks_i} V'_k = V_{s_i} / \gamma_{s_i}, \quad (2.19)$$

где $\lambda'_{ks_i} = \lambda_{ks_i} / \gamma_{s_i}$. Очевидно, для всех s_i

$$\sum_{k=1}^r |\lambda'_{ks_i}| = 1. \quad (2.20)$$

Не ограничивая общности, можно считать, что все последовательности $\{\lambda'_{ks_i}\}$, $k \in [1:r]$, сходятся

$$\lambda'_{ks_i} \xrightarrow{s_i \rightarrow \infty} \lambda'_{k0}.$$

Переходя к пределу при $s_i \rightarrow \infty$ в равенствах (2.19) и (2.20), получим в силу (2.16) и (2.18)

$$\sum_{k=1}^r \lambda'_{k0} V'_k = 0, \quad \sum_{k=1}^r |\lambda'_{k0}| = 1,$$

что противоречит линейной независимости системы $V'_1, \dots, V'_r$. Итак, все последовательности $\{\lambda_{ks}\}$, $k \in [1:r]$, в (2.17) ограничены.

Очевидно, найдется такая подпоследовательность индексов $\{\bar{s}_i\}$, для которой

$$\lambda_{k\bar{s}_i} \xrightarrow{\bar{s}_i \rightarrow \infty} \lambda_{k0}, \quad k \in [1:r].$$

Переходя в равенстве (2.17) к пределу по подпоследовательности индексов $\{\bar{s}_i\}$, получим

$$V^* = \sum_{k=1}^r \lambda_{k0} V'_k, \quad \lambda_{k0} \geq 0, \quad V'_k \in G,$$

откуда следует, что $V^* \in K(G)$.

Лемма доказана.

§ 3. Выпуклые функции

1. Функция $F(X)$ называется *выпуклой на выпуклом множестве* $\Omega \subset E_n$, если для любых X_1 и X_2 из Ω и любого $\alpha \in [0, 1]$ выполняется неравенство

$$F(\alpha X_1 + (1 - \alpha) X_2) \leq \alpha F(X_1) + (1 - \alpha) F(X_2). \quad (3.1)$$

Если при $X_1 \neq X_2$ и $\alpha \in (0, 1)$ неравенство (3.1) строгое, то функция $F(X)$ называется *строго выпуклой на Ω* . Неравенство (3.1) можно переписать в эквивалентном виде

$$F(X_2 + \alpha(X_1 - X_2)) \leq F(X_2) + \alpha[F(X_1) - F(X_2)].$$

Функция $F(X)$ называется *вогнутой на выпуклом множестве* $\Omega \subset E_n$, если для любых X_1 и X_2 из Ω и любого $\alpha \in [0, 1]$ выполняется неравенство

$$F(\alpha X_1 + (1 - \alpha) X_2) \geq \alpha F(X_1) + (1 - \alpha) F(X_2).$$

Из определений следует, что если функция $F(X)$ выпукла на Ω , то функция $(-F(X))$ вогнута на Ω .

Примеры.

I. Покажем, что функция $F_1(X) = (X, X)$ выпукла на E_n . Действительно, при всех $\alpha \in [0, 1]$ имеем

$$\begin{aligned} F_1(\alpha X_1 + (1 - \alpha) X_2) &\leq (\alpha X_1 + (1 - \alpha) X_2, \alpha X_1 + (1 - \alpha) X_2) + \\ &\quad + \alpha(1 - \alpha)(X_1 - X_2, X_1 - X_2) = \\ &= \alpha(X_1, X_1) + (1 - \alpha)(X_2, X_2) = \\ &= \alpha F_1(X_1) + (1 - \alpha) F_1(X_2), \end{aligned}$$

что и требовалось доказать. Нетрудно понять, что функция $F_1(X)$ является строго выпуклой на E_n .

II. Покажем, что функция $F_2(X) = (A, X)^2$ выпукла на E_n . Действительно, при всех $\alpha \in [0, 1]$ имеем

$$\begin{aligned} F_2(\alpha X_1 + (1 - \alpha) X_2) &\leq (A, \alpha X_1 + (1 - \alpha) X_2)^2 + \\ &\quad + \alpha(1 - \alpha)(A, X_1 - X_2)^2 = \alpha(A, X_1)^2 + (1 - \alpha)(A, X_2)^2 = \\ &= \alpha F_2(X_1) + (1 - \alpha) F_2(X_2), \end{aligned}$$

что и требовалось доказать.

2. Отметим два свойства выпуклых функций, которые используются в основном тексте.

Лемма 3.1. Пусть $\Omega \subset E_n$ — выпуклое множество и $F_i(X)$, $i \in [0 : N]$, — выпуклые на Ω функции. Тогда функции

$$\varphi(X) = \max_{i \in [0 : N]} F_i(X)$$

и

$$\Phi(X) = \sum_{i=0}^N \lambda_i F_i(X),$$

где все λ_i , $i \in [0 : N]$, неотрицательны, являются выпуклыми на Ω .

Доказательство. Докажем, например, выпуклость $\varphi(X)$ на Ω . По определению выпуклых функций имеем для $X_1, X_2 \in \Omega$ и $\alpha \in [0, 1]$

$$\begin{aligned} F_i(\alpha X_1 + (1 - \alpha) X_2) &\leq \alpha F_i(X_1) + (1 - \alpha) F_i(X_2) \leq \\ &\leq \alpha \max_{i \in [0 : N]} F_i(X_1) + (1 - \alpha) \max_{i \in [0 : N]} F_i(X_2). \end{aligned}$$

Отсюда следует:

$$\begin{aligned} \max_{i \in [0 : N]} F_i(\alpha X_1 + (1 - \alpha) X_2) &\leq \\ &\leq \alpha \max_{i \in [0 : N]} F_i(X_1) + (1 - \alpha) \max_{i \in [0 : N]} F_i(X_2), \end{aligned}$$

или, что то же самое,

$$\varphi(\alpha X_1 + (1 - \alpha) X_2) \leq \alpha \varphi(X_1) + (1 - \alpha) \varphi(X_2).$$

Но это и требовалось доказать.

Лемма 3.2. Пусть функция $F(X)$ непрерывно дифференцируема на открытом множестве $\Omega' \subset E_n$. Если к тому же функция $F(X)$ является выпуклой на некотором выпуклом множестве $\Omega \subset \Omega'$, то для всех $X_1, X_2 \in \Omega$ будет выполняться неравенство

$$\left(\frac{\partial F(X_1)}{\partial X}, X_2 - X_1 \right) \leq F(X_2) - F(X_1),$$

где

$$\frac{\partial F(X)}{\partial X} = \left(\frac{\partial F(X)}{\partial x_1}, \dots, \frac{\partial F(X)}{\partial x_n} \right).$$

Доказательство. Нетрудно проверить, что

$$\left(\frac{\partial F(X_1)}{\partial X}, X_2 - X_1 \right) = \lim_{\alpha \rightarrow +0} \frac{F(X_1 + \alpha(X_2 - X_1)) - F(X_1)}{\alpha}. \quad (3.2)$$

В силу выпуклости $F(X)$ на Ω имеем для $0 < \alpha < 1$

$$F(X_1 + \alpha(X_2 - X_1)) - F(X_1) \leq \alpha [F(X_2) - F(X_1)].$$

Поделив обе части последнего неравенства на $\alpha > 0$, перейдем в нем к пределу при $\alpha \rightarrow +0$. Получим на основании (3.2)

$$\left(\frac{\partial F(X_1)}{\partial X}, X_2 - X_1 \right) \leq F(X_2) - F(X_1).$$

Лемма доказана.

Замечание. Если в условиях леммы функция $F(X)$ является вогнутой на Ω , то

$$\left(\frac{\partial F(X_1)}{\partial X}, X_2 - X_1 \right) \geq F(X_2) - F(X_1)$$

для всех $X_1, X_2 \in \Omega$.

ПРИЛОЖЕНИЕ III

НЕПРЕРЫВНЫЕ И НЕПРЕРЫВНО ДИФФЕРЕНЦИРУЕМЫЕ ФУНКЦИИ

§ 1. Непрерывные функции

Функция $F(X)$, заданная на множестве $G \subset E_n$, называется *непрерывной в точке* $X_0 \in G$, если по любому $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех $X \in G$, для которых $\|X - X_0\| < \delta$, будет

$$|F(X) - F(X_0)| < \varepsilon.$$

Функция $F(X)$ называется *непрерывной на* G , если она непрерывна в каждой точке этого множества.

Если $G \subset E_n$ — ограниченное замкнутое множество, то любая непрерывная на G функция $F(X)$ достигает наибольшего и наименьшего на G значения, т. е. существуют такие точки $X_1, X_2 \in G$, что

$$F(X_1) = \sup_{X \in G} F(X), \quad F(X_2) = \inf_{X \in G} F(X).$$

Функция $F(X)$, непрерывная на ограниченном замкнутом множестве $G \subset E_n$, является *равномерно непрерывной* на G . Последнее означает, что по любому $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех $X_1, X_2 \in G$, для которых $\|X_1 - X_2\| < \delta$, будет

$$|F(X_1) - F(X_2)| < \varepsilon.$$

Вектор-функция $\Phi(X) = (F_1(X), \dots, F_r(X))$ называется *непрерывной на множестве* G , если непрерывными на G будут все $F_i(X)$, $i \in [1:r]$.

Рассмотрим функцию двух переменных $F(X, Y)$, где

$$X = (x_1, \dots, x_n) \in \Omega, \quad Y = (y_1, \dots, y_m) \in G.$$

Введем обозначение

$$\begin{aligned} \Omega \times G &= \{W = (x_1, \dots, x_n, y_1, \dots, y_m) \in \\ &\in E_{n+m} \mid X = (x_1, \dots, x_n) \in \Omega, \quad Y = (y_1, \dots, y_m) \in G\}. \end{aligned}$$

Будем говорить, что функция $F(X, Y)$, $X \in \Omega$, $Y \in G$, непрерывна по совокупности переменных, если функция $F(W)$ непрерывна на $\Omega \times G$.

Заметим, что если $\Omega \subset E_n$ и $G \subset E_m$ — ограниченные замкнутые множества, то ограниченным и замкнутым будет также множество $\Omega \times G \subset E_{n+m}$.

§ 2. Некоторые равенства и неравенства для непрерывных функций

1. В этом пункте рассматриваются функции, непрерывные на ограниченном замкнутом множестве $G \subset E_n$. Имеют место следующие соотношения.

I.

$$\max_{X \in G} F(X) = - \min_{X \in G} (-F(X)), \quad (2.1)$$

$$\min_{X \in G} F(X) = - \max_{X \in G} (-F(X)). \quad (2.2)$$

Доказательство. Обозначим через $\bar{X} \in G$ точку, в которой достигается максимум функции $F(X)$ на G , а через $X_0 \in G$ точку, в которой достигается минимум функции $-F(X)$ на G . Тогда

$$\max_{X \in G} F(X) = F(\bar{X}) = -(-F(\bar{X})) \leq - \min_{X \in G} (-F(X)). \quad (2.3)$$

С другой стороны,

$$- \min_{X \in G} (-F(X)) = -(-F(X_0)) = F(X_0) \leq \max_{X \in G} F(X). \quad (2.4)$$

Объединяя (2.3) и (2.4), получим (2.1).

Равенство (2.2) доказывается аналогично.

II.

$$\max_{X \in G} \lambda F(X) = \lambda \max_{X \in G} F(X), \quad (2.5)$$

$$\min_{X \in G} \lambda F(X) = \lambda \min_{X \in G} F(X). \quad (2.6)$$

Здесь $\lambda \geq 0$ — константа.

Доказательство. Поскольку при $\lambda \geq 0$, $Z \in G$ будет

$$\lambda F(Z) \leq \lambda \max_{X \in G} F(X),$$

то

$$\max_{Z \in G} \lambda F(Z) = \max_{X \in G} \lambda F(X) \leq \lambda \max_{X \in G} F(X). \quad (2.7)$$

С другой стороны,

$$\lambda \max_{X \in G} F(X) = \lambda F(\bar{X}) \leq \max_{X \in G} \lambda F(X). \quad (2.8)$$

Объединяя (2.7) и (2.8), получим (2.5).
Равенство (2.6) доказывается аналогично.

III.

$$\max_{X \in G} [F(X) + C] = \max_{X \in G} F(X) + C,$$

$$\min_{X \in G} [F(X) + C] = \min_{X \in G} F(X) + C.$$

Здесь C — произвольная константа.

Доказательство предоставляется читателю.

IV.

$$|\max_{X \in G} F(X)| \leq \max_{X \in G} |F(X)|, \quad (2.9)$$

$$|\min_{X \in G} F(X)| \leq \max_{X \in G} |F(X)|. \quad (2.10)$$

Доказательство. Имеем

$$|\max_{X \in G} F(X)| = |F(\bar{X})| \leq \max_{X \in G} |F(X)|.$$

Тем самым доказано неравенство (2.9). Неравенство (2.10) является следствием (2.9) и (2.2). Действительно,

$$|\min_{X \in G} F(X)| = |-\max_{X \in G} (-F(X))| \leq \max_{X \in G} | -F(X) | = \max_{X \in G} |F(X)|.$$

V.

$$\max_{X \in G} [F_1(X) + F_2(X)] \leq \max_{X \in G} F_1(X) + \max_{X \in G} F_2(X), \quad (2.11)$$

$$\max_{X \in G} [F_1(X) + F_2(X)] \geq \max_{X \in G} F_1(X) + \max_{X \in Q} F_2(X), \quad (2.12)$$

где

$$Q = \{X \in G \mid F_1(X) = \max_{Z \in G} F_1(Z)\}.$$

Доказательство. Неравенство (2.11) очевидно. Докажем неравенство (2.12). Для этого надо заметить, что функция $F_1(X)$ на Q постоянна, и воспользоваться

соотношением III. Получим

$$\begin{aligned} \max_{X \in G} [F_1(X) + F_2(X)] &\geq \max_{X \in Q} [F_1(X) + F_2(X)] = \\ &= \max_{X \in G} F_1(X) + \max_{X \in Q} F_2(X). \end{aligned}$$

VI.

$$\begin{aligned} \min_{X \in G} [F_1(X) + F_2(X)] &\geq \min_{X \in G} F_1(X) + \min_{X \in G} F_2(X), \\ \min_{X \in G} [F_1(X) + F_2(X)] &\leq \min_{X \in G} F_1(X) + \min_{X \in Q_1} F_2(X), \end{aligned}$$

где $Q_1 = \{X \in G \mid F_1(X) = \min_{Z \in G} F_1(Z)\}$.

Доказательство этих неравенств предоставляется читателю.

VII.

$$|\max_{X \in G} F_1(X) - \max_{X \in G} F_2(X)| \leq \max_{X \in G} |F_1(X) - F_2(X)|, \quad (2.13)$$

$$|\min_{X \in G} F_1(X) - \min_{X \in G} F_2(X)| \leq \max_{X \in G} |F_1(X) - F_2(X)|. \quad (2.14)$$

Доказательство. В силу (2.11) имеем

$$\begin{aligned} \max_{X \in G} F_1(X) &\leq \max_{X \in G} F_2(X) + \max_{X \in G} |F_1(X) - F_2(X)|, \\ \max_{X \in G} F_2(X) &\leq \max_{X \in G} F_1(X) + \max_{X \in G} |F_2(X) - F_1(X)|. \end{aligned}$$

Отсюда очевидным образом следует (2.13).

Для доказательства неравенства (2.14) надо воспользоваться (2.13) и (2.2).

VIII.

$$\begin{aligned} |\min_{X \in G} \max_{i \in [0: N]} F_i(X) - \min_{X \in G} \max_{i \in [0: N]} \Phi_i(X)| &\leq \\ &\leq \max_{X \in G} \max_{i \in [0: N]} |F_i(X) - \Phi_i(X)|. \end{aligned}$$

Доказательство. Нетрудно понять, что функции

$$\varphi_1(X) = \max_{i \in [0: N]} F_i(X) \quad \text{и} \quad \varphi_2(X) = \max_{i \in [0: N]} \Phi_i(X)$$

непрерывны на G . Учитывая (2.14) и (2.13), получим

$$\begin{aligned} |\min_{X \in G} \max_{i \in [0: N]} F_i(X) - \min_{X \in G} \max_{i \in [0: N]} \Phi_i(X)| &\leq \\ &\leq \max_{X \in G} |\varphi_1(X) - \varphi_2(X)| \leq \max_{X \in G} \max_{i \in [0: N]} |F_i(X) - \Phi_i(X)|, \end{aligned}$$

что и требовалось доказать.

IX.

$$(\max_{X \in G} |F(X)|)^2 = \max_{X \in G} (F(X))^2, \quad (2.15)$$

$$(\min_{X \in G} |F(X)|)^2 = \min_{X \in G} (F(X))^2. \quad (2.16)$$

Доказательство. Имеем

$$(\max_{X \in G} |F(X)|)^2 = (|F(X')|)^2 \leq \max_{X \in G} (F(X))^2. \quad (2.17)$$

С другой стороны, для всех $X \in G$

$$(F(X))^2 = (|F(X)|)^2 \leq (\max_{Z \in G} |F(Z)|)^2.$$

Отсюда следует, что

$$\max_{X \in G} (F(X))^2 \leq (\max_{X \in G} |F(X)|)^2. \quad (2.18)$$

Объединяя (2.17) и (2.18), получим (2.15). Равенство (2.16) доказывается аналогично.

Заметим, в частности, что в силу (2.15)

$$\max_{X \in G} |F(X)| = \sqrt{\max_{X \in G} (F(X))^2}.$$

X. Если $G = \bigcup_{k=1}^r G_k$, то

$$\max_{X \in G} F(X) = \max_{k \in [1:r]} \sup_{X \in G_k} F(X), \quad (2.19)$$

$$\min_{X \in G} F(X) = \min_{k \in [1:r]} \inf_{X \in G_k} F(X). \quad (2.20)$$

Доказательство. Поскольку для всех $k \in [1:r]$

$$\max_{X \in G} F(X) \geq \sup_{X \in G_k} F(X),$$

то

$$\max_{X \in G} F(X) \geq \max_{k \in [1:r]} \sup_{X \in G_k} F(X). \quad (2.21)$$

С другой стороны, если $\bar{X}$ — точка, в которой достигается максимум $F(X)$ на G , то $\bar{X}$ принадлежит одному из G_k , например G_1 . Поэтому

$$\max_{X \in G} F(X) = F(\bar{X}) \leq \sup_{X \in G_1} F(X) \leq \max_{k \in [1:r]} \sup_{X \in G_k} F(X). \quad (2.22)$$

Объединяя (2.21) и (2.22), получим (2.19).

Равенство (2.20) доказывается аналогично.

2. Пусть $F(t)$ — непрерывная на отрезке $[c, d]$ функция. Тогда имеют место следующие соотношения.

XI.

$$\begin{aligned} \max_{c \leq t \leq d} F(t) &= \max_{-d \leq z \leq -c} F(-z), \\ \max_{c \leq t \leq d} F(t) &= \max_{\alpha \leq z \leq \beta} F\left(c + \frac{d-c}{\beta-\alpha}(z-\alpha)\right). \end{aligned} \quad (2.23)$$

Докажем для примера равенство (2.23). Зафиксируем $t_0 \in [c, d]$ и положим

$$z_0 = \alpha + \frac{\beta-\alpha}{d-c}(t_0 - c).$$

Очевидно, $z_0 \in [\alpha, \beta]$. Теперь имеем

$$F(t_0) = F\left(c + \frac{d-c}{\beta-\alpha}(z_0 - \alpha)\right) \leq \max_{\alpha \leq z \leq \beta} F\left(c + \frac{d-c}{\beta-\alpha}(z - \alpha)\right).$$

Поскольку последнее неравенство справедливо для произвольного $t_0 \in [c, d]$, то

$$\max_{c \leq t \leq d} F(t) \leq \max_{\alpha \leq z \leq \beta} F\left(c + \frac{d-c}{\beta-\alpha}(z - \alpha)\right). \quad (2.24)$$

Докажем обратное неравенство. Пусть $\bar{z} \in [\alpha, \beta]$. Положим

$$\bar{t} = c + \frac{d-c}{\beta-\alpha}(\bar{z} - \alpha).$$

Очевидно, $\bar{t} \in [c, d]$. Теперь имеем

$$F\left(c + \frac{d-c}{\beta-\alpha}(\bar{z} - \alpha)\right) = F(\bar{t}) \leq \max_{c \leq t \leq d} F(t).$$

Поскольку последнее неравенство справедливо для произвольного $\bar{z} \in [\alpha, \beta]$, то

$$\max_{\alpha \leq z \leq \beta} F\left(c + \frac{d-c}{\beta-\alpha}(z - \alpha)\right) \leq \max_{c \leq t \leq d} F(t). \quad (2.25)$$

Объединяя (2.24) и (2.25), получим (2.23).

§ 3. Непрерывно дифференцируемые функции

1. Лемма 3.1. Пусть функция $F(X)$ непрерывна и имеет непрерывные частные производные первого порядка на открытом множестве $\Omega' \subset E_n$. Тогда для произвольных $X \in \Omega'$ и $g \in E_n$, $\|g\| = 1$, при достаточно

малых α справедлива формула

$$F(X + \alpha g) = F(X) + \alpha \left(\frac{\partial F(X)}{\partial X}, g \right) + o(X, g; \alpha),$$

где

$$\frac{\partial F(X)}{\partial X} = \left(\frac{\partial F(X)}{\partial x_1}, \dots, \frac{\partial F(X)}{\partial x_n} \right)$$

и $o(X, g; \alpha)$ — величина, для которой равномерно по g , $\|g\|=1$, выполняется соотношение

$$\frac{o(X, g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0.$$

Доказательство. Положим при фиксированных X и g

$$h(\alpha) = F(X + \alpha g). \quad (3.1)$$

Функция одной переменной $h(\alpha)$ имеет непрерывную производную в окрестности точки $\alpha=0$, причем, как нетрудно проверить,

$$h'(\alpha) = \left(\frac{\partial F(X + \alpha g)}{\partial X}, g \right). \quad (3.2)$$

Запишем для $h(\alpha)$ формулу Ньютона — Лейбница

$$h(\alpha) = h(0) + \int_0^\alpha h'(t) dt. \quad (3.3)$$

Эту формулу можно переписать в другом виде

$$h(\alpha) = h(0) + \alpha h'(0) + \int_0^\alpha [h'(t) - h'(0)] dt. \quad (3.4)$$

Учитывая (3.4), (3.1) и (3.2), получим

$$\begin{aligned} F(X + \alpha g) = F(X) + \alpha \left(\frac{\partial F(X)}{\partial X}, g \right) + \\ + \int_0^\alpha \left(\frac{\partial F(X + tg)}{\partial X} - \frac{\partial F(X)}{\partial X}, g \right) dt. \end{aligned}$$

Введем обозначение

$$o(X, g; \alpha) = \int_0^{\alpha} \left(\frac{\partial F(X + tg)}{\partial X} - \frac{\partial F(X)}{\partial X}, g \right) dt.$$

Для завершения доказательства леммы осталось установить, что равномерно по g , $\|g\| = 1$,

$$\frac{o(X, g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0.$$

Это немедленно следует из неравенства

$$|o(X, g; \alpha)| \leq |\alpha| \max_{t \in [-|\alpha|, |\alpha|]} \left\| \frac{\partial F(X + tg)}{\partial X} - \frac{\partial F(X)}{\partial X} \right\|$$

и непрерывности вектор-функции $\partial F(X)/\partial X$ в окрестности точки X .

Лемма доказана.

Замечание 1. Пусть $X \in \Omega'$ и $X + tg \in \Omega'$ при $t \in [0, \alpha]$, $\alpha > 0$. Из (3.3) по теореме о среднем для интегралов следует, что при некотором τ , $0 < \tau < \alpha$, будет

$$h(\alpha) = h(0) + \alpha h'(\tau). \quad (3.5)$$

Учитывая (3.1), (3.2) и (3.5), получим

$$F(X + \alpha g) = F(X) + \alpha \left(\frac{\partial F(X + \tau g)}{\partial X}, g \right),$$

где $0 < \tau < \alpha$.

Замечание 2. Если $M \subset \Omega'$ — ограниченное замкнутое множество, то найдется такое $\alpha_0 > 0$, что для всех $\alpha \in [0, \alpha_0]$ будем иметь

$$F(X + \alpha g) = F(X) + \alpha \left(\frac{\partial F(X)}{\partial X}, g \right) + o(X, g; \alpha),$$

где $\frac{o(X, g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$ равномерно по $X \in M$ и $g \in E_n$, $\|g\| = 1$.

2. Пусть функция $F(X)$ задана на некотором открытом множестве $\Omega' \subset E_n$.

Говорят, что функция $F(X)$ дифференцируема в точке $X \in \Omega'$, если существует такой вектор $V(X)$, что для

всех $g \in E_n$, $\|g\|=1$, будет

$$F(X + \alpha g) = F(X) + \alpha(g, V(X)) + o(X, g; \alpha), \quad (3.6)$$

где $\frac{o(X, g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$ равномерно по g .

Функция $F(X)$ называется *непрерывно дифференцируемой* на Ω' , если она дифференцируема в каждой точке этого множества и вектор-функция $V(X)$, фигурирующая в (3.6), непрерывна на Ω' .

В лемме 3.1 было, по существу, установлено, что если функция $F(X)$ непрерывна и имеет непрерывные частные производные первого порядка на открытом множестве Ω' , то она непрерывно дифференцируема на Ω' . Приведем еще одно достаточное условие непрерывной дифференцируемости.

Лемма 3.2. Пусть для всех $X \in \Omega'$ и $g \in E_n$, $\|g\|=1$, выполняется соотношение

$$\frac{\partial F(X)}{\partial g} \stackrel{\text{df}}{=} \lim_{\alpha \rightarrow +0} \frac{F(X + \alpha g) - F(X)}{\alpha} = (g, V(X)), \quad (3.7)$$

где вектор-функция $V(X)$ непрерывна на Ω' . Тогда функция $F(X)$ является непрерывно дифференцируемой на Ω' , причем

$$\frac{\partial F(X)}{\partial X} = V(X).$$

Доказательство. Зафиксируем g , $\|g\|=1$, и $X_0 \in \Omega'$, и рассмотрим функцию одной переменной

$$h(\alpha) = F(X_0 + \alpha g). \quad (3.8)$$

Поскольку $X_0 \in \Omega'$, то найдется $\delta > 0$ такое, что

$$X_0 + \alpha g \in \Omega' \quad \text{при} \quad \alpha \in (-\delta, \delta).$$

Докажем, что функция $h(\alpha)$ имеет непрерывную производную на $(-\delta, \delta)$. Действительно, в силу (3.7)

$$\begin{aligned} \lim_{\beta \rightarrow +0} \frac{1}{\beta} [h(\alpha + \beta) - h(\alpha)] &= \\ &= \lim_{\beta \rightarrow +0} \frac{1}{\beta} [F((X_0 + \alpha g) + \beta g) - F(X_0 + \alpha g)] = \\ &= (g, V(X_0 + \alpha g)). \end{aligned} \quad (3.9)$$

С другой стороны,

$$\begin{aligned} \lim_{\beta \rightarrow -0} \frac{h(\alpha + \beta) - h(\alpha)}{\beta} &= \\ &= - \lim_{\beta \rightarrow -0} \frac{F((X_0 + \alpha g) + (-\beta)(-g)) - F(X_0 + \alpha g)}{-\beta} = \\ &= - \lim_{\gamma \rightarrow +0} \frac{F((X_0 + \alpha g) + \gamma(-g)) - F(X_0 + \alpha g)}{\gamma} = \\ &= -(-g, V(X_0 + \alpha g)) = (g, V(X_0 + \alpha g)). \quad (3.10) \end{aligned}$$

Таким образом, функция $h(\alpha)$ в любой точке интервала $(-\delta, \delta)$ имеет производную слева и производную справа, причем эти односторонние производные совпадают. Значит, $h(\alpha)$ имеет обычную производную в каждой точке интервала $(-\delta, \delta)$, причем в силу (3.9) и (3.10)

$$h'(\alpha) = (g, V(X_0 + \alpha g)). \quad (3.11)$$

Учитывая непрерывность вектор-функции $V(X)$ на Ω' , заключаем, что $h'(\alpha)$ непрерывна на $(-\delta, \delta)$.

Запишем для функции $h(\alpha)$ формулу Ньютона — Лейбница

$$\begin{aligned} h(\alpha) &= h(0) + \int_0^\alpha h'(t) dt = \\ &= h(0) + \alpha h'(0) + \int_0^\alpha [h'(t) - h'(0)] dt. \end{aligned}$$

В силу (3.8) и (3.11) эту формулу можно переписать иначе

$$F(X_0 + \alpha g) = F(X_0) + \alpha(g, V(X_0)) + o(X_0, g; \alpha), \quad (3.12)$$

где

$$o(X_0, g, \alpha) = \int_0^\alpha (g, V(X_0 + tg) - V(X_0)) dt.$$

Нетрудно понять, что

$$\frac{o(X_0, g; \alpha)}{\alpha} \xrightarrow{\alpha \rightarrow 0} 0$$

равномерно по g , $\|g\| = 1$. Отсюда следует, что функция $F(X)$ дифференцируема в произвольной точке

$X_0 \in \Omega'$. Учитывая непрерывность функции $V(X)$, входящей в формулу (3.12), заключаем, что $F(X)$ непрерывно дифференцируема на Ω' .

Формула

$$\frac{\partial F(X_0)}{\partial X} = V(X_0)$$

следует из соотношения (3.12), если в нем в качестве g взять поочередно все координатные орты.

Лемма доказана.

3. Лемма 3.3. Пусть функция $F(X)$ непрерывна и имеет непрерывные частные производные до l -го порядка включительно на открытом множестве $\Omega' \subset E_n$. Тогда для фиксированных $X = (x_1, \dots, x_n) \in \Omega'$ и $g = (g_1, \dots, g_n)$, $\|g\| = 1$, при достаточно малых α справедлива формула

$$F(X + \alpha g) = F(X) + \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k F(X)}{\partial g^k} + o(X, g; \alpha^l),$$

где

$$\frac{\partial^k F(X)}{\partial g^k} = \sum_{j_1, \dots, j_k=1}^n \frac{\partial^k F(X)}{\partial x_{j_1} \dots \partial x_{j_k}} g_{j_1} \dots g_{j_k}$$

и $o(X, g; \alpha^l)$ — величина, для которой равномерно по g , $\|g\| = 1$, выполняется соотношение

$$\frac{o(X, g; \alpha^l)}{\alpha^l} \xrightarrow{\alpha \rightarrow 0} 0.$$

Доказательство. Функция одной переменной

$$h(\alpha) = F(X + \alpha g) \quad (3.13)$$

в окрестности точки $\alpha = 0$ имеет непрерывные производные до l -го порядка включительно, поэтому при достаточно малых α справедлива формула Маклорена

$$h(\alpha) = h(0) + \sum_{k=1}^{l-1} \frac{\alpha^k}{k!} h^{(k)}(0) + R_l(\alpha), \quad (3.14)$$

где остаточный член $R_l(\alpha)$ может быть представлен в виде

$$R_l(\alpha) = \frac{\alpha^l}{l!} h^{(l)}(\tau), \quad 0 < |\tau| < |\alpha|, \quad (3.15)$$

или

$$R_l(\alpha) = \frac{1}{(l-1)!} \int_0^\alpha (\alpha - t)^{l-1} h^{(l)}(t) dt.$$

Положив

$$R_l(\alpha) = \frac{\alpha^l}{l!} h^{(l)}(0) + o(X, g; \alpha^l),$$

где

$$o(X, g; \alpha^l) = \frac{1}{(l-1)!} \int_0^\alpha (\alpha - \tau)^{l-1} [h^{(l)}(t) - h^{(l)}(0)] dt, \quad (3.16)$$

будем иметь в силу (3.14)

$$h(\alpha) = h(0) + \sum_{k=1}^l \frac{\alpha^k}{k!} h^{(k)}(0) + o(X, g; \alpha^l). \quad (3.17)$$

Теперь заметим, что на основании (3.13)

$$h'(0) = \sum_{j=1}^n \frac{\partial F(X)}{\partial x_j} g_j = \left(\frac{\partial F(X)}{\partial X}, g \right) \stackrel{df}{=} \frac{\partial F(X)}{\partial g},$$

$$h''(0) = \sum_{j_1, j_2=1}^n \frac{\partial^2 F(X)}{\partial x_{j_1} \partial x_{j_2}} g_{j_1} g_{j_2} \stackrel{df}{=} \frac{\partial^2 F(X)}{\partial g^2},$$

.....

$$h^{(l)}(0) = \sum_{j_1, \dots, j_l=1}^n \frac{\partial^l F(x)}{\partial x_{j_1} \dots \partial x_{j_l}} g_{j_1} \dots g_{j_l} \stackrel{df}{=} \frac{\partial^l F(X)}{\partial g^l}.$$

Подставляя эти формулы в (3.17), получим

$$F(X + \alpha g) = F(X) + \sum_{k=1}^l \frac{\alpha^k}{k!} \frac{\partial^k F(X)}{\partial g^k} + o(X, g; \alpha^l).$$

Для завершения доказательства леммы осталось установить, что равномерно по $g, \|g\|=1$,

$$\frac{o(X, g; \alpha^l)}{\alpha^l} \xrightarrow{\alpha \rightarrow 0} 0. \quad (3.18)$$

В силу (3.16) имеем

$$|o(X, g; \alpha^l)| \leq$$

$$\leq \frac{|\alpha|^l}{l!} \max_{t \in [-|\alpha|, |\alpha|]} \sum_{j_1, \dots, j_l=1}^n \left| \frac{\partial^l F(X + tg)}{\partial x_{j_1} \dots \partial x_{j_l}} - \frac{\partial^l F(X)}{\partial x_{j_1} \dots \partial x_{j_l}} \right|.$$

Теперь (3.18) следует из непрерывности на Ω' всех частных производных l -го порядка функции $F(X)$.

Лемма доказана.

З а м е ч а н и е. Введем обозначение

$$\frac{\partial^2 F(X)}{\partial X^2} = \begin{pmatrix} \frac{\partial^2 F(X)}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 F(X)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F(X)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 F(X)}{\partial x_n \partial x_n} \end{pmatrix}.$$

Тогда будем иметь

$$\frac{\partial^2 F(X)}{\partial g^2} \stackrel{df}{=} \sum_{j_1, j_2=1}^n \frac{\partial^2 F(X)}{\partial x_{j_1} \partial x_{j_2}} g_{j_1} g_{j_2} = \left(\frac{\partial^2 F(X)}{\partial X^2} g, g \right). \quad (3.19)$$

Пусть $X \in \Omega'$ и $X + \gamma g \in \Omega'$ при $\gamma \in [0, \alpha]$. Положив в (3.14) $l=2$, получим на основании (3.15) и (3.19)

$$F(X + \alpha g) =$$

$$= F(X) + \alpha \left(\frac{\partial F(X)}{\partial X}, g \right) + \frac{\alpha^2}{2} \left(\frac{\partial^2 F(X + \tau g)}{\partial X^2} g, g \right),$$

где $0 < \tau < \alpha$.

ПРИЛОЖЕНИЕ IV

НАХОЖДЕНИЕ БЛИЖАЙШЕЙ К НАЧАЛУ КООРДИНАТ ТОЧКИ МНОГОГРАННИКА. ИТЕРАТИВНЫЕ МЕТОДЫ

1. Пусть в n -мерном евклидовом пространстве E_n задано конечное множество точек $H = \{Z_i\}_{i=1}^s$. Обозначим через L выпуклую оболочку, натянутую на H :

$$L = \left\{ Z = \sum_{i=1}^s \alpha_i Z_i \mid \alpha_i \geq 0, \sum_{i=1}^s \alpha_i = 1 \right\}.$$

Очевидно, L является ограниченным замкнутым выпуклым множеством.

Через Z^* будем обозначать ближайшую к началу координат точку множества L :

$$(Z^*, Z^*) = \min_{Z \in L} (Z, Z).$$

Нетрудно показать, что точка $Z^* \in L$ существует и единственна. Кроме того, для любого $Z \in L$ выполняется неравенство (см. лемму 1.3 из Приложения II)

$$(Z, Z^*) \geq (Z^*, Z^*). \quad (1)$$

Нашей целью является описание двух методов последовательных приближений для нахождения точки Z^* .

2. Лемма 1. Для любых $V, Z \in L$ выполняется неравенство

$$(V, Z) \geq \min_{i \in [1:s]} (Z_i, Z). \quad (2)$$

Доказательство. Пусть

$$V = \sum_{i=1}^s \alpha'_i Z_i; \quad \alpha'_i \geq 0, \quad \sum_{i=1}^s \alpha'_i = 1.$$

Тогда

$$(V, Z) = \sum_{i=1}^s \alpha'_i (Z_i, Z) \geq \left[\min_{i \in [1:s]} (Z_i, Z) \right] \sum_{i=1}^s \alpha'_i = \min_{i \in [1:s]} (Z_i, Z).$$

Лемма доказана.

Для любой точки $Z \in L$ введем величину $\delta(Z)$:

$$\delta(Z) = (Z, Z) - \min_{i \in [1:s]} (Z_i, Z).$$

В силу (2)

$$\delta(Z) \geq 0.$$

Лемма 2. Для любого $Z \in L$ выполняется неравенство

$$\|Z - Z^*\| \leq \min \{ \sqrt{\delta(Z)}, \|Z\| \}.$$

Доказательство. Неравенство

$$\|Z - Z^*\| \leq \sqrt{\delta(Z)} \quad (3)$$

эквивалентно следующему неравенству:

$$(Z - Z^*, Z - Z^*) - (Z, Z) + \min_{i \in [1:s]} (Z_i, Z) \leq 0,$$

или после простейших преобразований

$$-2(Z^*, Z) + (Z^*, Z^*) + \min_{i \in [1:s]} (Z_i, Z) \leq 0.$$

Справедливость последнего неравенства следует из (1) и (2).

Действительно,

$$\begin{aligned} -2(Z^*, Z) + (Z^*, Z^*) + \min_{i \in [1:s]} (Z_i, Z) &= \\ &= \{(Z^*, Z^*) - (Z, Z^*)\} + \{ \min_{i \in [1:s]} (Z_i, Z) - (Z^*, Z) \} \leq 0. \end{aligned}$$

Неравенство (3) точное. В ситуации, изображенной на рис. 35, неравенство (3) обращается в равенство.

Докажем теперь, что

$$\|Z - Z^*\| \leq \|Z\|. \quad (4)$$

Неравенство (4) следует из (1).

Действительно,

$$2(Z, Z^*) \geq 2(Z^*, Z^*) \geq (Z^*, Z^*).$$

Отсюда

$$(Z, Z) - 2(Z, Z^*) + (Z^*, Z^*) \leq (Z, Z),$$

что эквивалентно (4). Заметим, что неравенство (4) также точное. Оно обращается в равенство, когда $Z^* = 0$, т. е. когда начало координат принадлежит L . Объединяя (3) и (4), получим требуемое.

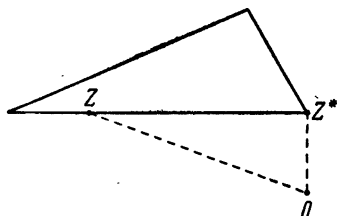


Рис. 35.

Лемма доказана.

Следствие 1. Если последовательность точек $\{V_k\}$, $V_k \in L$, $k=0, 1, 2, \dots$, такова, что $\delta(V_k) \xrightarrow{k \rightarrow \infty} 0$, то

$$V_k \xrightarrow{k \rightarrow \infty} Z^*.$$

Очевидным образом следует из (3).

Следствие 2. Если последовательность точек $\{V_k\}$, $V_k \in L$, $k=0, 1, 2, \dots$, такова, что

$$\|V_{k+1}\| \leq \|V_k\|,$$

и существует такая подпоследовательность $\{V_{k_j}\}$, что

$$\delta(V_{k_j}) \xrightarrow{k_j \rightarrow \infty} 0, \quad \text{то} \quad V_k \xrightarrow{k \rightarrow \infty} Z^*.$$

Доказательство. Поскольку числовая последовательность $\{\|V_k\|\}$ не возрастает и ограничена снизу, то она имеет предел

$$\|V_k\| \xrightarrow{k \rightarrow \infty} \mu.$$

С другой стороны, в силу (3)

$$\|V_{k_j}\| \xrightarrow{k_j \rightarrow \infty} \|Z^*\|.$$

Таким образом, $\|V_k\| \xrightarrow{k \rightarrow \infty} \|Z^*\|$.

В силу единственности Z^* (если $\bar{Z} \in L$ и $\|\bar{Z}\| = \|Z^*\|$, то $\bar{Z} = Z^*$) любая сходящаяся подпоследовательность последовательности $\{V_k\}$ имеет пределом Z^* . Отсюда следует, что вся последовательность $\{V_k\}$ сходится к Z^* .

Утверждение доказано.

Теорема 1. Для того чтобы точка $\bar{Z} \in L$ была ближайшей к началу координат точкой множества L , необходимо и достаточно, чтобы

$$\delta(\bar{Z}) = 0.$$

Доказательство. Достаточность следует из (3).

Необходимость. Пусть $\bar{Z} = Z^*$. Тогда, во-первых,

$$\delta(Z^*) \geq 0. \quad (5)$$

С другой стороны, в силу (1) для любого $i \in [1:s]$ будет

$$(Z_i, Z^*) \geq (Z^*, Z^*),$$

ибо $Z_i \in H \subset L$.

Отсюда

$$\min_{i \in [1:s]} (Z_i, Z^*) \geq (Z^*, Z^*),$$

или, что то же самое,

$$\delta(Z^*) \leq 0. \quad (6)$$

Объединяя (5) и (6), получим требуемое.

Теорема доказана.

3. Первый метод последовательных приближений.

В качестве начального приближения V_0 возьмем любую точку из L (в частности, в качестве V_0 можно взять точку Z_{i_0} из H ; для которой

$$(Z_{i_0}, Z_{i_0}) = \min_{i \in [1:s]} (Z_i, Z_i)).$$

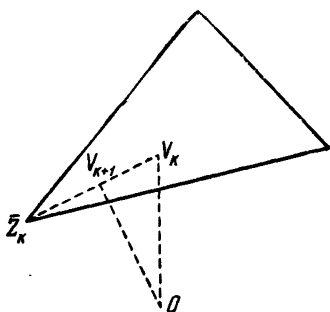


Рис. 36.

Пусть уже найдено k -е приближение $V_k \in L$. Опишем построение V_{k+1} .

Обозначим через $\bar{Z}_k$ точку множества H , для которой

$$(\bar{Z}_k, V_k) = \min_{i \in [1:s]} (Z_i, V_k)$$

(если таких точек несколько, берем любую из них).

В этом случае

$$\delta(V_k) = (V_k - \bar{Z}_k, V_k). \quad (7)$$

Рассмотрим отрезок

$$V_k(t) = V_k + t(\bar{Z}_k - V_k), \quad 0 \leq t \leq 1. \quad (8)$$

Пусть t_k , $0 \leq t_k \leq 1$, определяется соотношением

$$(V_k(t_k), V_k(t_k)) = \min_{t \in [0, 1]} (V_k(t), V_k(t)).$$

Положим $V_{k+1} = V_k(t_k)$ (рис. 36). Очевидно, $V_{k+1} \in L$.

Продолжая описанный процесс, получим последовательность $V_k \in L$, $k = 0, 1, 2, \dots$, причем

$$\|V_{k+1}\| \leq \|V_k\|. \quad (9)$$

Лемма 3. *Справедливо предельное соотношение*

$$\lim_{k \rightarrow \infty} \delta(V_k) = 0.$$

Доказательство. Заметим, во-первых, что в силу (7) и (8)

$$(V_k(t), V_k(t)) = (V_k, V_k) - 2t\delta(V_k) + t^2 \|\bar{Z}_k - V_k\|^2. \quad (10)$$

Будем использовать обозначение $\delta_k = \delta(V_k)$.

Допустим, что утверждение леммы не имеет места. Тогда найдется подпоследовательность $\{V_{k_j}\}$, для которой

$$\delta_{k_j} \geq \varepsilon > 0.$$

Имеем в силу (10)

$$(V_{k_j}(t), V_{k_j}(t)) \leq (V_{k_j}, V_{k_j}) - 2t\varepsilon + t^2 d^2,$$

где

$$d = \max_{Z, V \in L} \|Z - V\| > 0.$$

Отсюда следует, что для $t_0 = \min \left\{ \frac{\varepsilon}{d^2}, 1 \right\}$ (очевидно, $0 < t_0 \leq 1$) будет выполняться неравенство

$$(V_{k_j}(t_0), V_{k_j}(t_0)) \leq (V_{k_j}, V_{k_j}) - t_0 \varepsilon.$$

Учитывая, что по определению $(V_{k_j+1}, V_{k_j+1}) \leq (V_{k_j}(t_0), V_{k_j}(t_0))$, получим равномерно по k_j

$$(V_{k_j+1}, V_{k_j+1}) \leq (V_{k_j}, V_{k_j}) - t_0 \varepsilon. \quad (11)$$

Последнее неравенство немедленно приводит нас к противоречию. Действительно, в силу (9) числовая последовательность $\{\|V_k\|^2\}$ не возрастает и ограничена снизу. Значит, она имеет предел

$$\|V_k\|^2 \xrightarrow{k \rightarrow \infty} \mu,$$

причем

$$\|V_k\|^2 \geq \mu, \quad k = 0, 1, 2, \dots \quad (12)$$

Выберем $\bar{k}_j$ столь большим, чтобы выполнялось неравенство

$$\|V_{\bar{k}_j}\|^2 \leq \mu + \frac{1}{2} t_0 \varepsilon.$$

Тогда получим в силу (11)

$$\|V_{\bar{k}_j+1}\|^2 \leq \|V_{\bar{k}_j}\|^2 - t_0 \varepsilon \leq \mu - \frac{1}{2} t_0 \varepsilon,$$

что противоречит (12).

Лемма доказана.

Теорема 2. *Последовательность $\{V_k\}$, построенная описанным выше методом, сходится к точке Z^* .*

Доказательство очевидным образом следует из леммы 3 и следствия 1 из леммы 2.

Замечание. *Если при некотором k оказалось, что $\delta(V_k) = 0$, т. е. $V_k = Z^*$, то для всех $j = 1, 2, \dots$ будет $\delta(V_{k+j}) = 0$, так что $V_{k+j} = Z^*$.*

Следует из (10).

4. Обозначим через Ξ множество векторов A вида

$$A = (\alpha_1, \dots, \alpha_s); \quad \alpha_i \geq 0, \quad \sum_{i=1}^s \alpha_i = 1.$$

Положим

$$Z(A) = \sum_{i=1}^s \alpha_i Z_i; \quad Z(A) \in L.$$

Введем величину $\Delta(A)$:

$$\Delta(A) = \max_{\{i \mid \alpha_i > 0\}} (Z_i, Z(A)) - \min_{i \in [1:s]} (Z_i, Z(A)). \quad (13)$$

Через $i' = i'(A)$ обозначим индекс, на котором достигается максимум в правой части (13) (если таких индексов несколько, берем любой из них).

Таким образом, $\alpha_{i'} > 0$ и

$$(Z_{i'}, Z(A)) = \max_{\{i \mid \alpha_i > 0\}} (Z_i, Z(A)).$$

Лемма 4. *Для любого вектора $Z(A)$, $A \in \Xi$, справедливы неравенства*

$$\alpha_{i'} \Delta(A) \leq \delta(Z(A)) \leq \Delta(A).$$

Доказательство. Заметим, что

$$(Z(A), Z(A)) = \sum_{i=1}^s \alpha_i (Z_i, Z(A)) \leq \max_{\{i \mid \alpha_i > 0\}} (Z_i, Z(A)).$$

Отсюда следует неравенство $\delta(Z(A)) \leq \Delta(A)$.

Обозначим через $Z_{i'}$, $i'' = i''(A)$, точку множества H , для которой

$$(Z_{i'}, Z(A)) = \min_{i \in [1:s]} (Z_i, Z(A)).$$

В этом случае

$$\Delta(A) = (Z_{i'} - Z_{i'}, Z(A)). \quad (14)$$

Положим $\bar{A} = (\bar{\alpha}_1, \dots, \bar{\alpha}_s) \in \Xi$, где

$$\bar{\alpha}_i = \begin{cases} \alpha_i, & i \neq i', i''; \\ 0, & i = i'; \\ \alpha_{i'} + \alpha_{i''}, & i = i''. \end{cases}$$

Очевидно,

$$Z(\bar{A}) = Z(A) + \alpha_{i'} (Z_{i''} - Z_{i'}). \quad (15)$$

Поскольку $Z(\bar{A}) \in L$, то имеем в силу (2)

$$(Z(\bar{A}), Z(A)) \geq \min_{i \in [1:s]} (Z_i, Z(A)). \quad (16)$$

Учитывая (16), (15) и (14), получим

$$\delta(Z(A)) \geq \alpha_{i'} \Delta(A).$$

Лемма доказана.

Теорема 3. Для того чтобы точка $Z(A)$, $A \in \Xi$, была ближайшей к началу координат точкой множества L , необходимо и достаточно, чтобы $\Delta(A) = 0$.

Доказательство очевидным образом следует из леммы 4 и теоремы 1.

5. Второй метод последовательных приближений. Выберем произвольным образом вектор $A_0 \in \Xi$ и положим

$$V_0 = Z(A_0).$$

Пусть уже найдено k -е приближение $V_k \in L$,

$$V_k = Z(A_k), \quad A_k = (\alpha_1^{(k)}, \dots, \alpha_s^{(k)}) \in \Xi.$$

Опишем построение V_{k+1} . Прежде всего найдем векторы $Z_{i'_k}$ и $Z_{i''_k}$ из H такие, что

$$\begin{aligned}(Z_{i'_k}, V_k) &= \max_{\{i \mid \alpha_i^{(k)} > 0\}} (Z_i, Z(A_k)), \\ (Z_{i''_k}, V_k) &= \min_{i \in [1:s]} (Z_i, Z(A_k)).\end{aligned}$$

В этом случае

$$\Delta_k \stackrel{df}{=} \Delta(A_k) = (Z_{i'_k} - Z_{i''_k}, V_k). \quad (17)$$

Рассмотрим отрезок

$$V_k(t) = V_k + t\alpha_{i'_k}^{(k)}(Z_{i''_k} - Z_{i'_k}), \quad 0 \leq t \leq 1. \quad (18)$$

Пусть t_k , $0 \leq t_k \leq 1$, определяется соотношением

$$\begin{aligned}(V_k(t_k), V_k(t_k)) &= \\ &= \min_{t \in [0, 1]} (V_k(t), V_k(t)).\end{aligned}$$

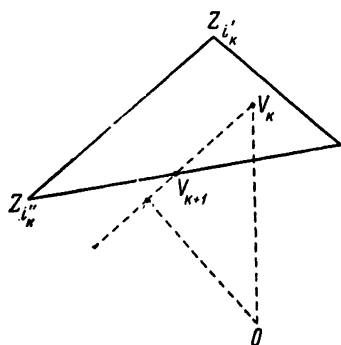


Рис. 37.

Положим $V_{k+1} = V_k(t_k)$ (рис. 37). Нетрудно проверить, что $V_{k+1} = Z(A_{k+1})$, где

$$\begin{aligned}A_{k+1} &= (\alpha_1^{(k+1)}, \dots, \alpha_s^{(k+1)}) \in \Xi, \\ \alpha_i^{(k+1)} &= \begin{cases} \alpha_i^{(k)}, & i \neq i'_k, i''_k; \\ \alpha_{i''_k}^{(k)} - t_k \alpha_{i'_k}^{(k)}, & i = i'_k; \\ \alpha_{i''_k}^{(k)} + t_k \alpha_{i'_k}^{(k)}, & i = i''_k. \end{cases}\end{aligned}$$

В дальнейшем для простоты будем использовать следующие обозначения:

$$\alpha'_k = \alpha_{i'_k}^{(k)}, \quad Z'_k = Z_{i'_k}, \quad \bar{Z}_k = Z_{i''_k}.$$

Продолжая описанный процесс, получим последовательность $\{V_k\}$, $V_k \in L$, $k = 0, 1, 2, \dots$, причем

$$\|V_{k+1}\| \leq \|V_k\|. \quad (19)$$

Лемма 5. Справедливо предельное соотношение

$$\lim_{k \rightarrow \infty} \alpha'_k \Delta_k = 0. \quad (20)$$

Доказательство. Заметим, во-первых, что в силу (17) и (18)

$$(V_k(t), V_k(t)) = (V_k, V_k) - 2t\alpha'_k\Delta_k + t^2(\alpha'_k\|\bar{Z}_k - Z'_k\|)^2. \quad (21)$$

Допустим, что утверждение леммы не имеет места. Тогда найдется подпоследовательность $\{V_{k_j}\}$, для которой

$$\alpha'_{k_j}\Delta_{k_j} \geq \varepsilon > 0.$$

Имеем в силу (21) для всех $t \in [0, 1]$ равномерно по k_j

$$(V_{k_j}(t), V_{k_j}(t)) \leq (V_{k_j}, V_{k_j}) - 2te + t^2d^2,$$

где

$$d = \max_{l, p \in [1: s]} \|Z_l - Z_p\| > 0.$$

Последнее неравенство вместе с (19) немедленно приводит нас к противоречию (см. доказательство леммы 3).

Лемма доказана.

Лемма 6. Справедливо равенство

$$\lim_{k \rightarrow \infty} \Delta_k = 0. \quad (22)$$

Доказательство. Допустим противное:

$$\lim_{k \rightarrow \infty} \Delta_k = \Delta' > 0. \quad (23)$$

Тогда для достаточно больших $k > K_1$ будет

$$\Delta_k \geq \frac{\Delta'}{2}. \quad (24)$$

Учитывая (20), заключаем, что

$$\alpha'_k \xrightarrow{k \rightarrow \infty} 0. \quad (25)$$

Обозначим через $\bar{t}_k$ точку, в которой $(V_k(t), V_k(t))$ достигает глобального минимума. Очевидно (см. (21)),

$$\bar{t}_k = \frac{\Delta_k}{\alpha'_k \|\bar{Z}_k - Z'_k\|^2}.$$

В силу (24) и (25)

$$\bar{t}_k \xrightarrow{k \rightarrow \infty} \infty.$$

Отсюда и из (21) следует, что для достаточно больших $k > K \geq K_1$ минимум $(V_k(t), V_k(t))$ на отрезке $0 \leq t \leq 1$ достигается при $t_k = 1$, так что для этих k

$$V_{k+1} = V_k + \alpha'_k (\bar{Z}_k - Z'_k). \quad (26)$$

Пусть $\{V_{k_j}\}$ — подпоследовательность, для которой

$$\Delta_{k_j} \xrightarrow{k_j \rightarrow \infty} \Delta'. \quad (27)$$

Прореживая, если нужно, подпоследовательность $\{V_{k_j}\}$, можно добиться выполнения следующих условий:

$$\text{I.} \quad \alpha_i^{(k_j)} \xrightarrow{k_j \rightarrow \infty} \alpha_i^*, \quad i \in [1 : s].$$

В этом случае

$$V_{k_j} \xrightarrow{k_j \rightarrow \infty} V^* \stackrel{\text{df}}{=} \sum_{i=1}^s \alpha_i^* Z_i.$$

$$\text{II.} \quad \bar{Z}_{k_j} = \bar{Z} \in H.$$

Это значит, что $\min_{i \in [1 : s]} (Z_i, V_{k_j})$ для всех k_j достигается при одном и том же i :

$$(\bar{Z}, V_{k_j}) = \min_{i \in [1 : s]} (Z_i, V_{k_j}).$$

$$\text{III.} \quad Z'_{k_j} = Z' \in H.$$

Таким образом,

$$(Z', V_{k_j}) = \max_{\{i | \alpha_i^{(k_j)} > 0\}} (Z_i, Z(A_{k_j})).$$

Учитывая (27) и свойства I — III подпоследовательности $\{V_{k_j}\}$, нетрудно показать, что

$$(Z' - \bar{Z}, V^*) = \Delta'. \quad (28)$$

Введем обозначения:

$$\rho^* = \min_{i \in [1 : s]} (Z_i, V^*),$$

$$H_1^* = \{Z_i \in H | (Z_i, V^*) = \rho^*\}, \quad H_2^* = H \setminus H_1^*.$$

Для того чтобы $Z_i \in H_2^*$, необходимо и достаточно, чтобы $(Z_i, V^*) > \rho^*$.

Заметим, что $\bar{Z} \in H_1^*$, а $Z' \in H_2^*$. Это следует из II и (28).

Положим

$$\begin{aligned}\rho' &= \min_{Z_i \in H_2^*} (Z_i, V^*), \quad \rho' > \rho^*; \\ \tau &= \min \{\Delta', \rho' - \rho^*\}.\end{aligned}$$

Заметим, что для $k_i \in H_2^*$ будет

$$(Z_i, V^*) \geq \rho^* + \tau. \quad (29)$$

Выберем $\delta_0 > 0$ так, чтобы из неравенства $\|V - V^*\| < \delta_0$ следовало

$$\max_{i \in [1:s]} |(Z_i, V) - (Z_i, V^*)| < \frac{\tau}{4}. \quad (30)$$

Возьмем любой номер k_j , для которого $\|V_{k_j} - V^*\| < \delta_0$. Тогда

$$H_1(V_{k_j}) \subset H_1^*, \quad (31)$$

где

$$H_1(V_{k_j}) = \{Z_i \in H \mid (Z_i, V_{k_j}) = \min_{l \in [1:s]} (Z_l, V_{k_j})\}.$$

Действительно, пусть $Z_i \in H_1(V_{k_j})$. Тогда в силу (30)

$$\begin{aligned}(Z_i, V^*) &= (Z_i, V_{k_j}) + (Z_i, V^* - V_{k_j}) = \\ &= \min_{l \in [1:s]} (Z_l, V_{k_j}) + (Z_i, V^* - V_{k_j}) = \rho^* + (Z_i, V^* - V_{k_j}) + \\ &+ \left[\min_{l \in [1:s]} (Z_l, V_{k_j}) - \min_{l \in [1:s]} (Z_l, V^*) \right] \leq \rho^* + (Z_i, V^* - V_{k_j}) + \\ &+ \max_{l \in [1:s]} |(Z_l, V_{k_j}) - (Z_l, V^*)| \leq \rho^* + \frac{\tau}{2}.\end{aligned}$$

Учитывая (29), заключаем, что $Z_i \in H_1^*$. Таким образом, соотношение (31) доказано.

Выберем теперь такой индекс $k_j > K$, чтобы выполнялись следующие условия:

$$\text{а) } \Delta_k \geq \Delta' - \frac{\tau}{4} \quad \text{для } k > k_j;$$

$$\text{б) } \|V_{k_j} - V^*\| < \frac{\delta_0}{2};$$

$$\text{в) } \sum_{l=1}^s \alpha'_{k_j+l-1} < \frac{\delta_0}{2d},$$

где

$$d = \max_{l, p \in [1:s]} \|Z_l - Z_p\| > 0.$$

Выполнения неравенства в) можно добиться в силу (25).

Учитывая (26), запишем для $p \in [1:s]$:

$$V_{k_j+p} = V_{k_j} + \sum_{l=1}^p \alpha'_{k_j+l-1} (\bar{Z}_{k_j+l-1} - Z'_{k_j+l-1}).$$

Заметим, во-первых, что на основании в)

$$\|V_{k_j+p} - V_{k_j}\| \leq d \sum_{l=1}^p \alpha'_{k_j+l-1} < \frac{\delta_0}{2},$$

так что в силу б) для $p \in [1:s]$ будет

$$\|V_{k_j+p} - V^*\| < \delta_0.$$

Отсюда следует, что для тех же p

$$H_1(V_{k_j+p}) \subset H_1^*. \quad (32)$$

(Доказательство аналогично доказательству соотношения (31).)

Далее для всех $Z_l \in H_1^*$ и $p \in [1:s]$ выполняется неравенство

$$(Z_l, V_{k_j+p}) \leq \Delta' + \rho^* - \frac{3\tau}{4}. \quad (33)$$

Действительно,

$$\begin{aligned} (Z_l, V_{k_j+p}) &= (Z_l, V^*) + (Z_l, V_{k_j+p} - V^*) \leq \\ &\leq \rho^* + \frac{\tau}{4} \leq \Delta' + \rho^* - \frac{3\tau}{4}. \end{aligned}$$

Теперь легко доказать, что для одного из V_{k_j+p} , $p \in [1:s]$, будет

$$\max_{\{i | \alpha_i^{(k_j+p)} > 0\}} (Z_i, V_{k_j+p}) \leq \Delta' + \rho^* - \frac{3\tau}{4}. \quad (34)$$

Действительно, если для какого-нибудь p неравенство (34) не выполняется, то в силу (33)

$$Z'_{k_j+p} \in H_2^*.$$

Учитывая (26) и (32), заключаем, что в представлении V_{k_j+p+1} вектор Z'_{k_j+p} будет иметь нулевой коэффициент, а для вводимого в представление вектора $\bar{Z}_{k_j+p}$ выполняется неравенство (33). Поскольку множество H_2 состоит не более чем из $(s-1)$ -го вектора, а для всех вводимых в представление векторов выполняется неравенство (33), то найдется $p' \in [1:s]$, для которого будет выполняться неравенство (34).

С другой стороны,

$$\min_{i \in [1:s]} (Z_i, V_{k_j+p'}) \geq \rho^* - \frac{\tau}{4}. \quad (35)$$

Таким образом, в силу (34) и (35) получаем

$$\Delta_{k_j+p'} \stackrel{\text{дф}}{=} \max_{\{i \mid \alpha_i^{(k_j+p')} > 0\}} (Z_i, V_{k_j+p'}) - \\ - \min_{i \in [1:s]} (Z_i, V_{k_j+p'}) \leq \Delta' - \frac{\tau}{2},$$

что противоречит а).

Лемма доказана.

Теорема 4. *Последовательность $\{V_k\}$, построенная выше, сходится к точке Z^* .*

Доказательство очевидным образом следует из лемм 6 и 4 и следствия 2 из леммы 2.

Замечание. Если при некотором k оказалось, что $\Delta_k = 0$, т. е. $V_k = Z^*$, то для $j = 1, 2, \dots$ будет $\Delta_{k+j} = 0$, так что $V_{k+j} = Z^*$.

Следует из (21).

6. Отметим некоторые особенности второго метода последовательных приближений.

Введем гиперплоскость G :

$$G = \{Z \mid (Z, Z^*) = (Z^*, Z^*)\}.$$

Теорема 5. *Если $Z^* \neq 0$, т. е. начало координат не принадлежит L , то, начиная с некоторого номера k , будет*

$$V_k \in G.$$

Доказательство. Заметим, что в силу теоремы 1

$$\min_{i \in [1:s]} (Z_i, Z^*) = (Z^*, Z^*).$$

Положим

$$H_1 = \{Z_i \in H \mid (Z_i, Z^*) = (Z^*, Z^*)\},$$

$$H_2 = H \setminus H_1 = \{Z_i \in H \mid (Z_i, Z^*) > (Z^*, Z^*)\}.$$

Если H_2 — пустое множество, то $V_k \in G$ для всех $k = 0, 1, 2, \dots$. Поэтому в дальнейшем будем предполагать, что H_2 — не пустое множество. Введем обозначение:

$$\tau = \min_{Z_i \in H_2} (Z_i, Z^*) - (Z^*, Z^*) > 0.$$

Поскольку $V_k \xrightarrow{k \rightarrow \infty} Z^*$, то для достаточно больших номеров $k > K_1$ будет

$$\max_{i \in [1:s]} |(Z_i, V_k) - (Z_i, Z^*)| < \frac{\tau}{4}.$$

Нетрудно показать, что для тех же $k > K_1$ выполняются соотношения

$$H_1(V_k) \subset H_1, \quad (36)$$

$$(Z_i, V_k) \leq (Z^*, Z^*) + \frac{\tau}{4}, \quad \text{если } Z_i \in H_1,$$

$$(Z_i, V_k) \geq (Z^*, Z^*) + \frac{3\tau}{4}, \quad \text{если } Z_i \in H_2.$$

Отсюда следует, что если в представлении V_k , $k > K_1$, входит с ненулевым коэффициентом $Z_i \in H_2$, то

$$\Delta_k \geq \frac{\tau}{2}. \quad (37)$$

Пусть

$$V_k = \sum_{\{i \mid Z_i \in H_1\}} \alpha_i^{(k)} Z_i + \sum_{\{i \mid Z_i \in H_2\}} \alpha_i^{(k)} Z_i.$$

В силу определения H_1 и τ

$$(V_k - Z^*, Z^*) = \sum_{\{i \mid Z_i \in H_2\}} \alpha_i^{(k)} (Z_i - Z^*, Z^*) \geq \tau \sum_{\{i \mid Z_i \in H_2\}} \alpha_i^{(k)}.$$

Поскольку левая часть этого неравенства стремится к нулю при $k \rightarrow \infty$, то

$$\sum_{\{i \mid Z_i \in H_2\}} \alpha_i^{(k)} \xrightarrow{k \rightarrow \infty} 0.$$

Выберем столь большое $K > K_1$, чтобы для $k \geq K$ выполнялось неравенство

$$\sum_{\{i \mid Z_i \in H_2\}} \alpha_i^{(k)} \leq \frac{\tau}{2d^2}, \quad (38)$$

где

$$d = \max_{Z_i \in H_1, Z_j \in H_2} \|Z_i - Z_j\| > 0.$$

Обозначим через $\bar{t}_k$ точку, в которой $(V_k(t), V_k(t))$ достигает глобального минимума. Если в представление V_k , $k \geq K$, входит с ненулевым коэффициентом точка $Z_i \in H_2$, то в силу (37) и (38) получим

$$\bar{t}_k = \frac{\Delta_k}{\alpha'_k \|\bar{Z}_k - Z'_k\|^2} \geq \frac{\tau}{2 \left(\sum_{\{i \mid Z_i \in H_2\}} \alpha_i^{(k)} \right) d^2} \geq 1.$$

Отсюда следует, что

$$V_{k+1} = V_k + \alpha'_k (\bar{Z}_k - Z'_k).$$

Теперь нетрудно показать, что в представлении вектора V_{K+s} отсутствуют (имеют нулевые коэффициенты) векторы $Z_i \in H_2$ (см. доказательство леммы 6). Учитывая (36), заключаем, что для всех $k \geq K + s$ будет

$$V_k = \sum_{\{i \mid Z_i \in H_1\}} \alpha_i^{(k)} Z_i; \quad \alpha_i^{(k)} \geq 0, \quad \sum_{\{i \mid Z_i \in H_1\}} \alpha_i^{(k)} = 1,$$

откуда очевидным образом следует, что $(V_k, Z^*) = (Z^*, Z^*)$, т. е. $V_k \in G$.

Теорема доказана.

Теорема 6. При $k \rightarrow \infty$ выполняется предельное соотношение

$$\Delta_k \rightarrow 0.$$

Доказательство. Если $Z^* = 0$, то утверждение теоремы следует из определения Δ_k и того, что $\|V_k\| \xrightarrow{k \rightarrow \infty} 0$.

Поэтому будем считать, что $Z^* \neq 0$. В силу теоремы 5 для $k \geq K + s$ будет

$$V_k = \sum_{\{i \mid Z_i \in H_1\}} \alpha_i^{(k)} Z_i; \quad \alpha_i^{(k)} \geq 0, \quad \sum_{\{i \mid Z_i \in H_1\}} \alpha_i^{(k)} = 1.$$

Кроме того, $H_1(V_k) \subset H_1$. Отсюда следует, что

$$\Delta_k \leq \max_{Z_i \in H_1} (Z_i, V_k) - \min_{Z_i \in H_1} (Z_i, V_k).$$

Правая часть этого неравенства в силу определения H_1 стремится к нулю при $k \rightarrow \infty$. Таким образом, и $\Delta_k \xrightarrow{k \rightarrow \infty} 0$, ибо $\Delta_k \geq 0$.

Теорема доказана.

Положим

$$\tilde{\Delta}_k = \frac{\Delta_k}{\|V_k\|^2}, \quad k = 0, 1, 2, \dots$$

Если $\|V_k\| = 0$, то по определению считаем, что $\tilde{\Delta}_k = \infty$.

Теорема 7. Для того чтобы начало координат принадлежало множеству L , необходимо и достаточно, чтобы при всех $k = 0, 1, 2, \dots$ выполнялось неравенство

$$\tilde{\Delta}_k \geq 1. \quad (39)$$

Доказательство. Необходимость. По условию $Z^ = 0$. Учитывая, что $V_k \in L$, получим в силу лемм 2 и 4*

$$\tilde{\Delta}_k = \frac{\Delta_k}{\|V_k\|^2} \geq \frac{\delta(V_k)}{\|V_k\|^2} \geq 1.$$

Достаточность. Допустим, что $Z^ \neq 0$. Имеем*

$$\tilde{\Delta}_k \leq \frac{\Delta_k}{\|Z^*\|^2}.$$

В силу теоремы 6 $\tilde{\Delta}_k \xrightarrow{k \rightarrow \infty} 0$, что противоречит (39).

Теорема доказана.

Итак, если $Z^* = 0$, то $\|V_k\| \xrightarrow{k \rightarrow \infty} 0$ и при всех $k = 0, 1, 2, \dots$ выполняется неравенство $\tilde{\Delta}_k \geq 1$. Если же $Z^* \neq 0$, то $\|V_k\| \geq \|Z^*\|$ и $\tilde{\Delta}_k \xrightarrow{k \rightarrow \infty} 0$.

Если при некотором k выполнилось неравенство $\tilde{\Delta}_k < 1$, то в силу теоремы 7 начало координат не принадлежит множеству L . Более того, нетрудно показать, что в этом случае гиперплоскость

$$(V_k, Z) - (V_k, \bar{Z}_k) = 0$$

строго отделяет начало координат от множества L .

7. Продолжим рассмотрения, связанные со вторым методом последовательных приближений.

Пусть $V_k = Z(A_k)$, $A_k \in \Xi$.

Положим

$$J_k = \{i \mid \alpha_i^{(k)} > 0\}.$$

Введем множество B_k :

$$B_k = \left\{ Z = Z_{i_0} + \sum_{\substack{i \in J_k \\ i \neq i_0}} \alpha_i (Z_i - Z_{i_0}) \mid \alpha_i \in (-\infty, \infty) \right\}.$$

Здесь i_0 — произвольный фиксированный индекс из J_k .

Очевидно, B_k — выпуклое множество. Обозначим через $\tilde{V}_k$ вектор из B_k с наименьшей нормой:

$$\|\tilde{V}_k\| = \min_{Z \in B_k} \|Z\|.$$

Заметим, что точка $\tilde{V}_k \in B_k$ единственна, хотя неединственным может быть ее представление в виде

$$\tilde{V}_k = Z_{i_0} + \sum_{\substack{i \in J_k \\ i \neq i_0}} \tilde{\alpha}_i (Z_i - Z_{i_0}).$$

Нетрудно показать, что числа $\tilde{\alpha}_i$, $i \in J_k$, $i \neq i_0$, являются решением следующей линейной системы:

$$\left(Z_{i_0} + \sum_{\substack{i \in J_k \\ i \neq i_0}} \alpha_i (Z_i - Z_{i_0}), Z_j - Z_{i_0} \right) = 0, \quad j \in J_k, \quad j \neq i_0. \quad (40)$$

Теорема 8. *Существует бесконечная подпоследовательность векторов $\{\tilde{V}_{k_j}\}$ такая, что при всех k_j будет*

$$\tilde{V}_{k_j} = Z^*.$$

Доказательство. Будем считать, что $Z^* \neq 0$ (если $Z^* = 0$, то доказательство лишь упрощается). Выделим вначале такую подпоследовательность $\{V_{k_j}\}$, чтобы

I. $\alpha_i^{(k_j)} \xrightarrow{k_j \rightarrow \infty} \alpha_i^*$, $i \in [1 : s]$; в этом случае

$$V_{k_j} \xrightarrow{k_j \rightarrow \infty} Z^* = \sum_{i=1}^s \alpha_i^* Z_i.$$

II. $V_{k_j} \in G$ (см. теорему 5).

Положим

$$J^* = \{i \mid \alpha_i^* > 0\}.$$

Очевидно, что при достаточно больших k_j будет

$$J^* \subset J_{k_j}. \quad (41)$$

В дальнейшем только такие k_j и рассматриваются.

Через L^* обозначим выпуклую оболочку, натянутую на точки Z_i , $i \in J^*$. Очевидно, $Z^* \in L^*$.

Через L_{k_j} обозначим выпуклую оболочку, натянутую на точки Z_i , $i \in J_{k_j}$. В силу (41), II и определения множества B_{k_j} имеем

$$L^* \subset L_{k_j} \subset B_{k_j} \subset G.$$

Далее

$$\|Z^*\| = \min_{Z \in G} \|Z\| \leq \min_{Z \in B_{k_j}} \|Z\|.$$

Поскольку $Z^* \in L^*$, то $Z^* \in B_{k_j}$. Значит,

$$\|Z^*\| = \min_{Z \in B_{k_j}} \|Z\|.$$

Учитывая, что точка множества B_{k_j} , имеющая наименьшую норму, единственна, получаем

$$\tilde{V}_{k_j} = Z^*.$$

Теорема доказана.

На основании этой теоремы можно утверждать, что нахождение ближайшей к началу координат точки множества L сводится к решению конечного числа систем линейных уравнений вида (40). Заметим, что решать эти системы целесообразно лишь для тех k , для которых величины $\|V_k\|$ или $\tilde{\Delta}_k$ достаточно малы.

8. Полученные выше результаты допускают обобщение.

Рассмотрим вектор $V = Z(A)$, $A \in \Xi$, и пусть

$$\rho(V) = \min_{i \in [1:s]} (Z_i, V).$$

При фиксированном $c \in [0, 1]$ введем множество

$$F_c \stackrel{\text{def}}{=} F_c(Z(A)) = \{Z_i \in H \mid \alpha_i > 0, (Z_i, V) - \rho(V) \geq c \Delta(A)\}.$$

Положим далее

$$V'_c = \sum_{Z_i \in F_c} \frac{\alpha_i}{\gamma_c} Z_i,$$

где

$$\gamma_c = \sum_{Z_i \in F_c} \alpha_i.$$

Введем величину $\Delta_c(V)$:

$$\Delta_c(V) = (V'_c, V) - \min_{i \in [1:s]} (Z_i, V).$$

Из определения следует, что

$$\Delta(V) \geq \Delta_c(V) \geq c \Delta(V).$$

В частности, $\Delta(V) = \Delta_1(V)$.

Нетрудно проверить, что

$$\delta(V) = \Delta_0(V).$$

Действительно, если $c = 0$, то $\gamma_0 = 1$, $V'_0 = V$, так что $\Delta_0(V) = \delta(V)$.

Лемма 7. Если $0 \leq c_1 \leq c_2 \leq 1$, то для любого вектора $V = Z(A)$, $A \in \Xi$, выполняется неравенство

$$\Delta_{c_1}(V) \leq \Delta_{c_2}(V).$$

Доказательство. Очевидно, что

$$F_{c_1} \supset F_{c_2}, \quad \gamma_{c_1} \geq \gamma_{c_2}. \quad (42)$$

Имеем в силу (42) и определения множества F_{c_2} :

$$\begin{aligned} (V'_{c_2} - V'_{c_1}, V) &= \left(\frac{1}{\gamma_{c_2}} - \frac{1}{\gamma_{c_1}} \right) \sum_{\{i \mid Z_i \in F_{c_2}\}} \alpha_i (Z_i, V) - \\ &\quad - \frac{1}{\gamma_{c_1}} \sum_{\{i \mid Z_i \in F_{c_1} \setminus F_{c_2}\}} \alpha_i (Z_i, V) \geq \left(\frac{1}{\gamma_{c_2}} - \frac{1}{\gamma_{c_1}} \right) (c_2 \Delta(V) + \\ &\quad + \rho(V)) \sum_{\{i \mid Z_i \in F_{c_2}\}} \alpha_i - \frac{1}{\gamma_{c_1}} (c_2 \Delta(V) + \rho(V)) \sum_{\{i \mid Z_i \in F_{c_1} \setminus F_{c_2}\}} \alpha_i = \\ &= (c_2 \Delta(V) + \rho(V)) \left(\sum_{\{i \mid Z_i \in F_{c_2}\}} \frac{\alpha_i}{\gamma_{c_2}} - \sum_{\{i \mid Z_i \in F_{c_1}\}} \frac{\alpha_i}{\gamma_{c_1}} \right) = 0. \end{aligned}$$

Итак,

$$(V'_{c_2} - V'_{c_1}, V) \geq 0,$$

откуда очевидным образом следует утверждение леммы.

Лемма доказана.

Теорема 9. Для того чтобы точка $V = Z(A)$, $A \in \Xi$, была ближайшей к началу координат точкой множества L , достаточно, чтобы при некотором $c \in [0, 1]$, и необходимо, чтобы при всех $c \in [0, 1]$ выполнялось равенство

$$\Delta_c(V) = 0.$$

Следует из соотношений

$$\delta(V) = \Delta_0(V) \leq \Delta_c(V) \leq \Delta_1(V) = \Delta(V)$$

и теорем 1 и 3.

9. В заключение опишем еще один метод нахождения ближайшей к началу координат точки множества L , являющейся обобщением первого и второго методов последовательных приближений.

Зафиксируем некоторое $c \in [0, 1]$. Возьмем произвольным образом вектор $A_0 \in \Xi$ и положим

$$V_0 = Z(A_0).$$

Пусть уже найдено k -е приближение $V_k \in L$,

$$V_k = Z(A_k), A_k = (\alpha_1^{(k)}, \dots, \alpha_s^{(k)}) \in \Xi.$$

Опишем построение V_{k+1} . Найдем вектор $\bar{Z}_k \in H$, для которого

$$(\bar{Z}_k, V_k) = \min_{i \in [1:s]} (Z_i, V_k).$$

(Если таких векторов несколько, берем любой из них.)

Рассмотрим отрезок

$$V_k(t) = V_k + t\gamma_k(\bar{Z}_k - V'_k), \quad 0 \leq t \leq 1,$$

где

$$\gamma_k = \sum_{\{i \mid Z_i \in F_c(V_k)\}} \alpha_i, \quad V'_k = \sum_{\{i \mid Z_i \in F_c(V_k)\}} \frac{\alpha_i}{\gamma_k} Z_i \in L.$$

Заметим, что

$$\Delta_c(V_k) = (V'_k - \bar{Z}_k, V_k).$$

Через t_k , $0 \leq t_k \leq 1$, обозначим число, определяемое соотношением

$$(V_k(t_k), V_k(t_k)) = \min_{0 \leq t \leq 1} (V_k(t), V_k(t)).$$

Положим $V_{k+1} = V_k(t_k)$. Нетрудно показать, что $V_{k+1} = Z(A_{k+1})$, где $A_{k+1} = (\alpha_1^{(k+1)}, \dots, \alpha_s^{(k+1)}) \in \Xi$.

Если $\bar{Z}_k \notin F_c(V_k)$, то

$$\alpha_i^{(k+1)} = \begin{cases} \alpha_i^{(k)}, & \text{если } Z_i \notin F_c(V_k), Z_i \neq \bar{Z}_k, \\ \alpha_i^{(k)} - t_k \alpha_i^{(k)}, & \text{если } Z_i \in F_c(V_k), \\ \alpha_i^{(k)} + t_k \gamma_k, & \text{если } Z_i = \bar{Z}_k. \end{cases}$$

Если же $\bar{Z}_k \in F_c(V_k)$, то

$$\alpha_i^{(k+1)} = \begin{cases} \alpha_i^{(k)}, & \text{если } Z_i \notin F_c(V_k), \\ \alpha_i^{(k)} - t_k \alpha_i^{(k)}, & \text{если } Z_i \in F_c(V_k), Z_i \neq \bar{Z}_k, \\ \alpha_i^{(k)} - t_k \alpha_i^{(k)} + t_k \gamma_k, & \text{если } Z_i = \bar{Z}_k. \end{cases}$$

Продолжая описанный процесс, получим последовательность $\{V_k\}$, $V_k \in L$, $k = 0, 1, 2, \dots$, причем

$$\|V_{k+1}\| \leq \|V_k\|.$$

Если $c = 0$, то $\gamma_k = 1$, $V'_k = V_k$,

$$V_k(t) = V_k + t(\bar{Z}_k - V_k), \quad 0 \leq t \leq 1,$$

и мы приходим к первому методу последовательных приближений.

Пусть $c = 1$. В этом случае

$$F_1(V_k) = \{Z_i \in H \mid \alpha_i^{(k)} > 0, (Z_i - \bar{Z}_k, V_k) = \Delta(V_k)\}.$$

Любая точка множества $F_1(V_k)$ может быть взята в качестве Z'_k (см. п. 5), так что при $c = 1$ получается естественное обобщение второго метода последовательных приближений.

Сформулируем основные результаты, относящиеся к общему методу.

Теорема 10. При $k \rightarrow \infty$ выполняется предельное соотношение

$$V_k \rightarrow Z^*.$$

Теорема 11. Если $Z^* \neq 0$ и $c \in (0, 1]$, то, начиная с некоторого номера k , будет

$$V_k \in G,$$

где $G = \{Z \mid (Z, Z^*) = (Z^*, Z^*)\}$.

Доказательство этих утверждений вполне аналогично доказательству теоремы 4 и 5.

10. Описанные в пп. 3 и 5 методы плохо сходятся, когда начало координат принадлежит многограннику L . В связи с этим рекомендуется делать следующее. Вначале решить задачу линейного программирования в пространстве векторов $W = (\alpha_1, \dots, \alpha_s, u)$: найти $\min u$ при условиях

$$\begin{aligned} \sum_{i=1}^s \alpha_i z_k^{(i)} - u &\leq 0, & k \in [1:n], \\ - \sum_{i=1}^s \alpha_i z_k^{(i)} - u &\leq 0, & k \in [1:n], \\ \sum_{i=1}^s \alpha_i &= 1; \alpha_i \geq 0, i \in [1:s]; & u \geq 0, \end{aligned}$$

где $Z_i = (z_1^{(i)}, \dots, z_n^{(i)})$ — точки из H . Обозначим через $W_0 = (\alpha_1^{(0)}, \dots, \alpha_s^{(0)}, u^{(0)})$ решение этой задачи. Нетрудно проверить, что $0 \in L$ тогда и только тогда, когда $u^{(0)} = 0$. Если $u^{(0)} = 0$, то $Z^* = 0$. Если же $u^{(0)} > 0$, то для нахождения Z^* можно воспользоваться первым или вторым методами последовательных приближений, взяв в качестве начального приближения точку

$$Z_0 = \sum_{i=1}^s \alpha_i^{(0)} Z_i,$$

где вектор $(\alpha_1^{(0)}, \dots, \alpha_s^{(0)})$ составлен из первых s координат вектора W_0 .

Заметим, что во многих случаях достаточно найти первый элемент последовательности $\{V_k\}$, сходящейся к Z^* , для которого

$$\frac{\min_{Z \in H} (Z, V_k)}{(V_k, V_k)} \geq \theta,$$

где $0 < \theta < 1$ (см. § 9 гл. III). Такая точка V_k может быть найдена в конечное число шагов.

ДОБАВЛЕНИЕ

О ЗАДАЧЕ МАНДЕЛЬШТАМА

В этом Добавлении приводится пример численного решения одной непрерывной минимаксной задачи.

1. Постановка задачи. Положим

$$\begin{aligned} Z &= (z_1, z_2, \dots, z_{n+1}) \in E_{n+1}, \\ W_n(Z, t) &= \sum_{k=1}^{n+1} \cos(kt + z_k), \\ \Phi_n(Z) &= \max_{t \in [-\pi, \pi]} |W_n(Z, t)|. \end{aligned}$$

Требуется минимизировать функцию $\Phi_n(Z)$ на E_{n+1} .

2. Предварительные рассуждения. Прежде всего, заметим, что $\Phi_n(z_1, z_2, \dots, z_{n+1})$ — непрерывная на E_{n+1} функция, 2π -периодическая по каждому аргументу. Поскольку

$$\inf_{Z \in E_{n+1}} \Phi_n(Z) = \inf_{Z \in Q_{n+1}} \Phi_n(Z),$$

где $Q_{n+1} = \{Z = (z_1, z_2, \dots, z_{n+1}) \mid -\pi \leq z_k \leq \pi, k \in [1: n+1]\}$ — ограниченное замкнутое множество, то инфимум $\Phi_n(Z)$ на E_{n+1} достигается, т. е. существует точка Z^* , для которой

$$\Phi_n(Z^*) = \inf_{Z \in E_{n+1}} \Phi_n(Z).$$

Далее, справедливы соотношения:

- I. $\max_{Z \in E_{n+1}} \Phi_n(Z) = \Phi_n(0) = (n+1)^2.$
- II. $\Phi_n(-Z) = \Phi_n(Z).$
- III. При любом вещественном λ

$$\Phi_n(Z + \lambda p) = \Phi_n(Z),$$

где $p = (1, 2, \dots, n+1).$

Докажем, например, равенство III. Поскольку максимумы 2π -периодической функции на любых двух

отрезках длиной 2π равны, то

$$\begin{aligned}\Phi_n(Z + \lambda p) &= \max_{t \in [-\pi, \pi]} \left| \sum_{k=1}^{n+1} \cos(kt + z_k + \lambda k) \right| = \\ &= \max_{u \in [-\pi + \lambda, \pi + \lambda]} \left| \sum_{k=1}^{n+1} \cos(ku + z_k) \right| = \Phi_n(Z),\end{aligned}$$

что и требовалось доказать.

Лемма. Если через $Z^{(s)}$ обозначить произвольный $(n+1)$ -мерный вектор, у которого s -я координата равна нулю, то для любого $s \in [1 : n+1]$ будет

$$\min_{\{Z\}} \Phi_n(Z) = \min_{\{Z^{(s)}\}} \Phi_n(Z^{(s)}).$$

Доказательство. Очевидно,

$$\min_{\{Z\}} \Phi_n(Z) \leq \min_{\{Z^{(s)}\}} \Phi_n(Z^{(s)}). \quad (1)$$

Далее, произвольный вектор $Z = (z_1, z_2, \dots, z_{n+1})$ можно представить в виде $Z = \left(Z - \frac{z_s}{s} p\right) + \frac{z_s}{s} p$. Так как s -я координата вектора $Z - \frac{z_s}{s} p$ равна нулю, то в силу III

$$\Phi_n(Z) = \Phi_n\left(Z - \frac{z_s}{s} p\right) \geq \min_{\{Z^{(s)}\}} \Phi_n(Z^{(s)}).$$

Отсюда следует, что

$$\min_{\{Z\}} \Phi_n(Z) \geq \min_{\{Z^{(s)}\}} \Phi_n(Z^{(s)}). \quad (2)$$

Объединяя (1) и (2), получим требуемое.

Лемма доказана.

В дальнейшем будем считать, что $z_1 = 0$.

3. Результаты вычислений. Введем обозначения

$$X = (x_1, \dots, x_n) \in E_n,$$

$$V_n(X, t) = \cos t + \sum_{k=1}^n \cos((k+1)t + x_k),$$

$$\varphi_n(X) = \max_{t \in [-\pi, \pi]} |V_n(X, t)|,$$

$$R(X) = \{t \in [-\pi, \pi] \mid V_n(X, t) = \varphi_n(X)\},$$

и будем решать следующую задачу, эквивалентную исходной: минимизировать функцию $\varphi_n(X)$ на E_n .

В табл. 4 приведены результаты вычислений для $n \in [1:5]$ (в четвертом столбце $t_l^* \in R(X^*)$).

При $n=1$ и $n=2$ вектор X^* является точкой глобального минимума функции $\varphi_n(X)$ на E_n ; при $n \in [3:5]$ гарантируется лишь, что X^* есть точка строгого локального минимума.

Таблица 4

n	X^*	$R(X^*)$	$V_n(X^*, t_l^*)$
1	1,570 796	-2,506 726 -0,634 867	-1,760 173 1,760 173
2	2,859 712 2,146 750	-0,840 986 1,399 331 2,707 688	1,979 806 1,979 806 -1,979 806
3	1,396 481 1,072 719 -1,209 531	-1,991 906 -0,083 577 1,860 299 2,769 081	-2,039 103 2,039 103 2,039 103 -2,039 103
4	1,084 832 2,087 535 -0,250 345 -1,277 773	-1,874 150 -0,953 170 0,833 524 1,545 230 2,277 249	-2,343 516 2,343 516 -2,343 516 2,343 516 -2,343 516
5	1,927 212 2,311 026 -2,514 500 2,102 134 1,076 190	-1,881 850 -1,058 158 -0,481 051 0,249 616 0,855 029 1,988 655	-2,549 372 2,549 372 2,549 372 -2,549 372 2,549 372 2,549 372

4. О методе решения. При $n=1$ легко найти точное решение задачи: $X^* = \pi/2$.

При $n=2$ и $n=3$ разными способами было получено начальное приближение X_0 , которое обладало следующим свойством: в ε -полосе

$$\{t \in [-\pi, \pi] | \varphi_n(X_0) - |V_n(X_0, t)| \leq \varepsilon\},$$

где $\varepsilon = 0,01$, содержалось ровно $n+1$ точек локального максимума функции $|V_n(X_0, t)|$. Обозначим эти точки

через $t_i^{(0)}$, $i \in [1 : n+1]$:

$$-\pi < t_1^{(0)} < t_2^{(0)} < \dots < t_{n+1}^{(0)} \leq \pi.$$

При $n \in [4 : 5]$ для получения хорошего начального приближения рассматривали дискретную минимаксную задачу:

$$\varphi_{nN}(X) \stackrel{\text{df}}{=} \max_{i \in [0 : N]} V_n^2(X, t_i) \rightarrow \min_{X \in E_n},$$

где $t_i = -\pi + i \frac{2\pi}{N}$. Положив $N = 400$ и $\varepsilon = 0,05$, находили ε -стационарную точку X_0 функции $\varphi_{nN}(X)$. Отметим, что множество

$\{t_i | \varphi_{nN}(X_0) - V_n^2(X_0, t_i) \leq \varepsilon;$
 $V_n^2(X_0, t_i) \geq V_n^2(X_0, t_{i-1}); V_n^2(X_0, t_i) \geq V_n^2(X_0, t_{i+1})\}$
 состояло из $n+1$ точек:

$$-\pi < t_1^{(0)} < t_2^{(0)} < \dots < t_{n+1}^{(0)} \leq \pi.$$

На втором этапе при всех $n \in [2 : 5]$ производили выраивание максимумов, которое заключалось в решении следующей нелинейной системы $(2n+1)$ -го порядка:

$$\left. \begin{aligned} \sigma_k V_n(X, y_k) - \sigma_{k+1} V_n(X, y_{k+1}) &= 0, \quad k \in [1 : n], \\ V'_n(X, y_k) &= 0, \quad k \in [1 : n+1]. \end{aligned} \right\} \quad (3)$$

Здесь $\sigma_k = \text{sign } V_n(X_0, t_k^{(0)})$, а неизвестными являются $x_1, \dots, x_n, y_1, \dots, y_{n+1}$.

Для решения системы (3) использовалась специальная модификация метода Ньютона, предложенная Б. Ф. Митчеллом. В качестве начального приближения брали точку, составленную из компонент вектора X_0 и чисел $t_1^{(0)}, \dots, t_{n+1}^{(0)}$.

Обозначим решение системы (3) через $x_1^*, \dots, x_n^*, t_1^*, t_2^*, \dots, t_{n+1}^*$. Тогда вектор $X^* = (x_1^*, \dots, x_n^*)$ является стационарной точкой функции $\varphi_n(X)$, а числа $t_1^*, t_2^*, \dots, t_{n+1}^*$ образуют экстремальный базис $R(X^*)$.

Выполнение условия

$$\sum_{i=1}^{n+1} \alpha_i \frac{\partial |V_n(X^*, t_i^*)|}{\partial X} = 0$$

при некоторых $\alpha_i > 0$, $\sum_{i=1}^{n+1} \alpha_i = 1$, гарантировало то, что X^* является точкой строгого локального минимума функции $\varphi_n(X)$.

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

Принцип выбора оптимальных значений параметров на основе минимизации максимального отклонения был предложен П. Л. Чебышевым и подробно описан в его классическом мемуаре [80].

Предисловие

Задача Мандельштама как пример минимаксной задачи встречается в работе Н. Г. Чеботарева [79]. Теория матричных игр подробно освещена в монографиях С. Карлина [39] и Дж. фон Неймана и О. Моргенштерна [57].

Глава I

Дискретная задача наилучшего приближения функций алгебраическими полиномами рассматривалась Валле-Пуссенем [6]. В частности, Валле-Пуссену принадлежит идея построения теории наилучшего приближения на основе чебышевской интерполяции. Им же разработан алгоритм построения полинома наилучшего приближения, описанный в § 3.

R-алгоритм предложен в [49].

К § 5. Первая геометрическая интерпретация задачи о наилучшем приближении (теорема 5.1) имеется в книге Е. Я. Ремеза [65]. Сведение задачи о наилучшем приближении к задаче о крайней точке пересечения оси с многогранником осуществлено, вероятно, впервые Г. Ш. Рубинштейном [70].

М. Д. Калашников [35] и И. И. Этерман [84] изучали с разных точек зрения полиномы (2.9). В частности, в [84] эти полиномы используются для решения различных задач вычислительной математики.

Глава II

В этой главе рассматривается принципиальный вопрос о переходе от дискретной задачи к непрерывной. Предельные теоремы (§ 3) и метод последовательных чебышевских интерполяций Е. Я. Ремеза (§ 4) можно найти в монографии [65]. Близкие вопросы обсуждались в [64 и 4]. При некоторых дополнительных предположениях теорема 5.1 о сходимости сеточного метода была доказана еще Валле-Пуссенем. Приводимое в книге доказательство предложено в [50]. Относительно теоремы 5.2 см., например, [4].

Понятие плохой обусловленности полиномов было введено Б. А. Самокишем. Ему же принадлежит теорема 6.1 и другие более тонкие результаты о поведении коэффициентов полиномов наилучшего приближения, опубликованные в [72].

Глава III

Принципиальным достижением последних лет в теории минимаксных задач является вывод формулы производной по направлению для функции максимума. Эта формула в некоторых частных случаях была получена в [10 и 61]. В общем случае она установлена в работах [14, 15, 17—19]. Теорема 2.2 доказана в [20].

К § 3. Теоремы 3.1 и 3.2 были, по существу, известны еще Н. Г. Чеботареву [79]. Выражение для направления наискорейшего спуска функции максимума (теорема 3.3) получено в [20].

К § 4. Некоторые из приведенных здесь оценок имеются в [18].

К § 5. Идея метода наискорейшего спуска восходит еще к Коши [44] и Адамару [1]. Детально этот метод был изучен в 40-х годах Левенбергом [47], Л. В. Канторовичем [37] и Карри [40].

Для случая непрерывно дифференцируемых функций и метод покоординатного спуска, и метод наискорейшего спуска приводят к стационарной точке. В минимаксных задачах, как показано в основном тексте, дело обстоит хуже. Пример «заедания» метода наискорейшего спуска взят из [23].

Хороший обзор работ по методу наискорейшего спуска имеется в статье Г. Дж. Келли [41].

К § 6. Первый метод последовательных приближений был опубликован в [18]. Описанная в п. 5 модификация этого метода по идее близка к алгоритму Зойтендейка [30—32] и С. И. Зуховицкого, Р. А. Поляка, М. Е. Примака [33].

К § 7. В одном частном случае минимаксных задач метод, описанный в п. 1, рассматривался в [3].

К § 8. *D*-метод был предложен в [21, 23].

К § 9. Теория линейного программирования подробно излагается в [16, 34 и 85]. О методах минимизации функции одной переменной см. [75]. В [2] имеются программы, оформленные в виде процедур на языке АЛГОЛ-60, которые могут быть использованы для решения минимаксных задач.

О других методах решения дискретных минимаксных задач, в частности о методах второго порядка см. [86].

Глава IV

Сопряженные конусы для формулировки необходимых условий экстремума впервые были использованы А. Я. Дубовицким и А. А. Милутиным [28, 29].

Теорема 3.1. по существу, установлена в [61, 62] (см. также [51]).

К § 5. Некоторые из приведенных здесь оценок имеются в [18].

Глава V

Эта глава содержит результаты применения общей теории, развитой в главах III и IV, к обобщенной задаче математического программирования.

Теория решения задач математического (или оптимального, нелинейного) программирования в настоящее время бурно развивается. Мы не имеем возможности дать здесь подробную библиографию по этому вопросу; отметим лишь монографии [32, 39, 54, 71], работы [12, 48] и обзор Кюнц и Эттли [46], где имеется обширный список литературы.

К § 2. Условие Слейтера было сформулировано в [73].

К § 5. Теорема Куна и Таккера (для случая $N \neq 0$) опубликована в [45]. Различные обобщения этой теоремы приведены, например, в монографии С. Карлина [39].

§§ 6—9 написаны на основе [20, 22].

К § 10. Идея штрафных функций встречается в [41]. Дальнейшие результаты о применении метода штрафных функций к решению минимаксных задач получены в [8, 9].

К § 11. Другие методы решения задач нелинейного программирования разработаны в [67—69, 82, 83 и 60].

Глава VI

В §§ 2—3 полученные в главах III—V результаты обобщаются на непрерывный случай. По этому поводу см. [18, 51, 61].

Теорема 3.3 является частным случаем одной общей теоремы Л. Г. Шнирельмана [81] (см. также [74]).

Необходимым условиям экстремума посвящены работы [7, 11, 26, 28, 29, 36, 54, 59, 63].

К § 4. Теорема 4.1 о сходимости сеточного метода опубликована в [52].

К § 5. Относительно теоремы о минимаксе см., например, [39].

§ 6 написан на основе [24]. О других методах нахождения седловых точек см. [13, 27, 57, 66]. Как известно [16, 57], матричные игры эквивалентны задачам линейного программирования.

К § 7. Идея доказательства теоремы 7.1 о существовании полинома наилучшего приближения заимствована из [38]. Теорема 7.2 установлена в [43, 5, 58]. Вопрос о сходимости сеточного метода в задаче о наилучшем приближении функции нескольких аргументов исследовался в [53].

К § 8. Тот факт, что теорема Чебышева является следствием теоремы 2.2, был замечен еще Н. Г. Чеботаревым [79].

Приложение I

Доказательство приведенных здесь результатов по алгебраической интерполяции можно найти, например, в [56].

Приложение II

При написании этого приложения использовались монографии [39 и 76].

Приложение III

Более подробную информацию по изложенным здесь вопросам можно найти в [77].

Приложение IV

Относительно первого метода последовательных приближений см. [78, 25, 42]. Второй метод предложен в [55].

Добавление

Написано Е. Ф. Войтоном, В. Н. Малоземовым и Б. Ф. Митчеллом.

ЛИТЕРАТУРА

1. Адамар (Hadamard J.), *Mémoire sur le problème d'analyse relatif à l'équilibre des plaques élastiques encastrées*, Мém. prés. Acad. Sci. France 2 33, № 4 (1908).
2. Алгол-процедуры. Методические материалы по программному обеспечению ЭВМ под ред. И. В. Романовского, ВЦ ЛГУ, 1971, вып. 9, Ротапринт.
3. Бирзак Б., Пшеничный Б. Н., О некоторых задачах минимизации негладких функций, Кибернетика, № 6 (1966), 53—57.
4. Буров В. Н., Некоторые эффективные способы решения задачи П. Л. Чебышева о наилучшем приближении, Известия высших учебных заведений, Математика, № 1 (1957), 67—79.
5. Буров В. Н., Об аппроксимации функций полиномами, удовлетворяющими нелинейным связям, ДАН СССР 138, № 3 (1961), 515—517.
6. Валле-Пуссен (Valée Poussin J. Ch.), *Leçons sur l'approximation des fonctions d'une variable réelle*, Paris, 1919.
7. Гермейер Ю. Б., Необходимые условия максимина, Журнал выч. матем. и матем. физики 9, № 2 (1969), 432—438.
8. Гермейер Ю. Б., Приближенное сведение с помощью штрафных функций задачи определения максимина к задаче определения максимума, Журнал выч. матем. и матем. физики 9, № 3 (1969), 730—731.
9. Гермейер Ю. Б., К задаче отыскания максимина с ограничениями, Журнал выч. матем. и матем. физики 10, № 1 (1970), 39—55.
10. Гирсанов И. В., Дифференцируемость решений задач математического программирования, Тезисы докладов Всесоюзной межвузовской конференции по применению методов функционального анализа к решению нелинейных задач, Баку (1965), 43—45.
11. Гирсанов И. В., Лекции по математической теории экстремальных задач, Изд-во Московского университета, 1970.
12. Гольштейн Е. Г., Выпуклое программирование (Элементы теории), «Наука», 1970.
13. Данскин Дж. М., Итеративный метод решения непрерывных игр, Сб. «Матричные игры», Физматгиз, 1961, 123—132.
14. Данскин (Danskin J. M.), The theory of max-min, with applications, SIAM J. on Applied Math. 14 (1966), 641—664.
15. Данскин (Danskin J. M.), The theory of max-min and its application to weapons allocation problems, Springer-Verlag, New York Inc., 1967.
16. Данциг Дж., Линейное программирование, его обобщения и применения, «Прогресс», 1966.

17. Демьянов В. Ф., К минимизации максимального уклонения, Вестник ЛГУ, № 7 (1966), 21—28.
18. Демьянов В. Ф., К решению некоторых минимаксных задач, I, II, Кибернетика, № 6 (1966), 58—66, № 3 (1967), 62—66.
19. Демьянов В. Ф., Метод последовательных приближений для решения одной минимаксной задачи, Тезисы кр. научн. сообщений Международного конгресса математиков, Секция 14, М. (1966), 31.
20. Демьянов В. Ф., Algorithms for some minimax problems, J. Computer and System Sciences 2, № 4 (1968), 342—380.
21. Демьянов В. Ф., К задаче о минимаксе, ДАН СССР 187, № 2 (1969), 255—258.
22. Демьянов В. Ф., К разысканию минимакса на ограниченном множестве, ДАН СССР 191, № 6 (1970), 1216—1219.
23. Демьянов В. Ф., Дифференцирование функции максимина. I, III, Журнал выч. матем. и матем. физики 8, № 6 (1968), 1186—1195; 10, № 1 (1970), 26—38.
24. Демьянов В. Ф., Нахождение седловых точек на многогранниках, ДАН СССР 192, № 1 (1970), 13—15.
25. Демьянов В. Ф., Рубинов А. М., Минимизация гладкого выпуклого функционала на выпуклом множестве, Вестник ЛГУ, № 19 (1964), 5—17.
26. Демьянов В. Ф., Рубинов А. М., Приближенные методы решения экстремальных задач, Изд. ЛГУ, 1968.
27. Дрешер М., Карлин С., Решение выпуклых игр методом неподвижных точек, Сб. «Бесконечные антагонистические игры», Физматгиз, 1963, 180—194.
28. Дубовицкий А. Я., Милютин А. А., Задачи на экстремум при наличии ограничений, ДАН СССР 149, № 4 (1963), 759—762.
29. Дубовицкий А. Я., Милютин А. А., Задачи на экстремум при наличии ограничений, Журнал выч. матем. и матем. физики 5, № 3 (1965), 395—453.
30. Зойтендейк (Zoutendijk G.), Maximizing a function in a convex region, J. Roy. Statist. Soc., Ser. B, 21 (1959), 338—355.
31. Зойтендейк (Zoutendijk G.), Nonlinear programming: A numerical survey, SIAM J. on Control 4 (1966), 194—210.
32. Зойтендейк, Методы возможных направлений, ИЛ, 1963.
33. Зуховицкий С. И., Поляк Р. А., Примак М. Е., Алгоритм для решения задачи выпуклого программирования, ДАН СССР 153, № 5 (1963), 991—994.
34. Зуховицкий С. И., Авдеева Л. И., Линейное и выпуклое программирование, изд. 2-е, «Наука», 1967.
35. Калашников М. Д., О полиномах наилучшего квадратичного приближения в заданной системе точек, ДАН СССР 105 (1955), 634—636.
36. Канторович Л. В., Об одном эффективном методе решения некоторых экстремальных проблем, ДАН СССР 28 (1940), 212—215.
37. Канторович Л. В., О методе наискорейшего спуска, ДАН СССР 56 (1947), 233—236.
38. Канторович Л. В., Акилов Г. П., Функциональный анализ в нормированных пространствах, Физматгиз, 1959.

39. Карлин С., Математические методы в теории игр, программировании и экономике, «Мир», 1964.
40. Карри (Curry H. B.), The method of steepest descent for nonlinear minimization problems, Quart. Appl. Math. 2, № 3 (1944), 258—261.
41. Келли Генри Дж., Метод градиентов, Сб. «Методы оптимизации» под ред. Дж. Лейтмана, «Наука» (1965), 244—306.
42. Козинец Б. Н., Об одном алгоритме обучения линейного персептрона. Сб. «Вычислительная техника и вопросы программирования», вып. 3 (1964), 80—83.
43. Колмогоров А. Н., Замечания по поводу многочленов П. Л. Чебышева, наименее уклоняющихся от заданной функции, УМН 3, № 1 (1948), 216—221.
44. Коши (Cauchy A. L.), Méthode générale pour la résolution des systèmes d'équations simultanées, Compt. rend. Acad. Sci. 25 (1847), 536—538.
45. Кун, Таккер (Kuhn H. M., Tucker A. W.), Nonlinear programming, Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, Berkeley (1951), 481—491.
46. Кюнзи, Эттли (Künzi H. P., Oettli W.), Nichtlineare Optimierung: Neuere Verfahren Bibliographie, Springer-Verlag, Berlin — Heidelberg — New York, 1969.
47. Левенберг (Levenberg K. A.), A method for the solution of certain nonlinear problems in least squares, Quart. Appl. Math. 2 (1944), 164—168.
48. Левитин Е. С., Поляк Б. Т., Методы минимизации при наличии ограничений, Журнал выч. матем. и матем. физики 6, № 5 (1966), 787—823.
49. Малоземов В. Н., К нахождению алгебраического полинома наилучшего приближения, Кибернетика 5 (1969), 125—131.
50. Малоземов В. Н., О сходимости сеточного метода в задаче наилучшей полиномиальной аппроксимации, Вестник ЛГУ, № 19 (1970), 138—140.
51. Малоземов В. Н., К теории нелинейных минимаксных задач, Кибернетика 3 (1970), 121—125.
52. Малоземов В. Н., О сходимости сеточного метода в нелинейных минимаксных задачах, Вестник ЛГУ, № 19 (1971), 35—37.
53. Малоземов В. Н., Наилучшее равномерное приближение функций нескольких аргументов, Журнал выч. матем. и матем. физики 10, № 3 (1970), 575—586.
54. Мангасарян (Mangasarian O. L.), Nonlinear programming, McGraw-Hill Book Company, Inc., New York — Toronto — London, 1970.
55. Митчелл Б. Ф., Демьянов В. Ф., Малоземов В. Н., Нахождение ближайшей к началу координат точки многогранника, Вестник ЛГУ, № 13 (1971), 38—45.
56. Мысовских И. П., Лекции по методам вычислений, Физматгиз, 1962.
57. Нейман Дж. Ф., Моргенштерн О., Теория игр и экономическое поведение, «Наука», 1970.

58. Никольский В. Н., Характеристический признак наименее уклоняющихся элементов в выпуклых множествах, Сб. «Исследования по современным проблемам конструктивной теории функций», Изд. АН Азерб. СССР, Баку (1965), 80—84.
59. Ньюстадт (Neustadt L. W.), An abstract variational theory with applications to a broad class of optimization problems, I, II, SIAM J. on Control 5 (1966), 505—525; 5 (1967), 90—137.
60. Поляк Б. Т., Один общий метод решения экстремальных задач, ДАН СССР 174, № 1 (1967), 33—36.
61. Пшеничный Б. Н., Двойственный метод в экстремальных задачах, I, II, Кибернетика 3 (1965), 89—95; 4 (1965), 64—69.
62. Пшеничный Б. Н., Выпуклое программирование в нормированном пространстве, Кибернетика 5 (1965), 46—54.
63. Пшеничный Б. Н., Необходимые условия экстремума, «Наука», 1969.
64. Ремез Е. Я., О методе последовательных чебышевских интерполяций и о различных вариантах его реализации, Украинский матем. журнал 12, № 2 (1960), 170—180.
65. Ремез Е. Я., Основы численных методов чебышевского приближения, «Наукова думка», Киев, 1969.
66. Робинсон Дж., Итеративный метод решения игр, Сб. «Матричные игры», Физматгиз, (1961), 110—117.
67. Розен (Rosen J. B.), Nonlinear Programming. The gradient projection method, Bull. Amer. Math. Soc. 63 (1957), 25—26.
68. Розен (Rosen J. B.), The gradient projection method for nonlinear programming, Part I. Linear Constraints, Part II. Nonlinear constraints, SIAM J. on Appl. Math. 8 (1960), 181—217; 9 (1961), 514—532.
69. Розен (Rosen J. B.), Gradient Projection as a least squares solution of Kuhn—Tucker conditions, July 1965, Univ. of Wisconsin, TR.
70. Рубинштейн Г. Ш., Задача о крайней точке пересечения оси с многогранником и ее приложение к исследованию конечной системы линейных неравенств, ДАН СССР 100 (1955), 627—630.
71. Рубинштейн Г. Ш., Конечномерные модели оптимизации. Курс лекций, Новосибирский университет, 1970, Ротапринт.
72. Самокиш Б. А., О поведении коэффициентов многочленов, приближающих регулярную на отрезке функцию, Сб. «Методы вычислений», Изд. ЛГУ, вып. 1 (1963), 58—65.
73. Слейтер (Slater M.), Lagrange multipliers revisited. A contribution to nonlinear programming, The RAND Corp. RM-676 (1951), Santa Monica, California.
74. Смирнов В. И., Лебедев Н. А., Конструктивная теория функций комплексного переменного, «Наука», 1964.
75. Уайлд Д. Дж., Методы поиска экстремума, «Наука», 1967.
76. Фаддеев Д. К., Фаддеева В. Н., Вычислительные методы линейной алгебры, Физматгиз, 1960.
77. Фихтенгольц Г. М., Курс дифференциального и интегрального исчисления, т. I, Физматгиз, 1962.
78. Франк, Вулф (Frank M., Wolfe P.), An algorithm for quadratic programming, Naval Res. Logist. Quart. 3 (1956), 93—110.

79. Чеботарев Н. Г., Об одном критерии минимакса, ДАН СССР 39 (1943), 373—376 (см. также Собр. соч., т. 2, Изд. АН СССР, 1949).
80. Чебышев П. Л., Вопросы о наименьших величинах, связанных с приближенным представлением функций, Собр. соч., т. II, Изд. АН СССР (1947), 151—238.
81. Шнирельман Л. Г., О равномерных приближениях, Изв. АН СССР, сер. матем., 2, № 1 (1938), 53—59.
82. Шор Н. З., Применение обобщенного градиентного спуска в блочном программировании, Кибернетика 3 (1967), 53—55.
83. Шор Н. З., Обобщенный градиентный спуск, Труды первой зимней школы по математическому программированию в г. Драгобыче, вып. 3, М. (1969), 578—585.
84. Этерман И. И., Аппроксимация полиномами и решение некоторых задач прикладной математики, Изд. Пензенского политехн. ин-та, 1960.
85. Юдин Д. Б., Гольштейн Е. Г., Линейное программирование, Физматгиз, 1963.
86. Демьянов В. Ф., Малоземов В. Н., К теории нелинейных минимаксных задач, УМН 26, № 3 (1971), 53—104.