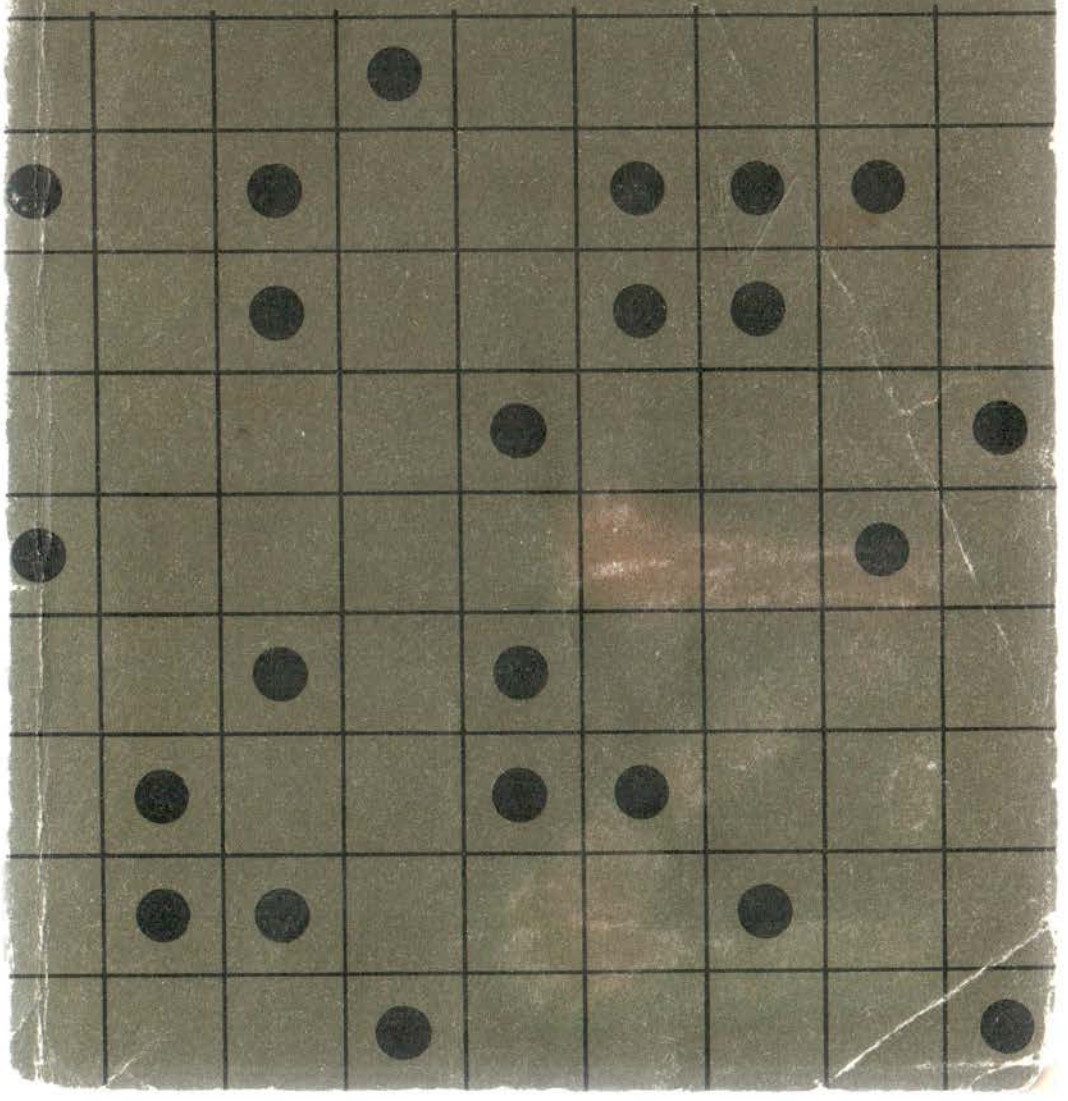
А1630
В-74

Вопросы теории и элементы программного обеспечения минимаксных задач



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

ВОПРОСЫ ТЕОРИИ И ЭЛЕМЕНТЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ МИНИМАКСНЫХ ЗАДАЧ

Под редакцией В. Ф. Демьянова и В. Н. Малоземова

ИЗДАТЕЛЬСТВО ЛЕНИНГРАДСКОГО УНИВЕРСИТЕТА
ЛЕНИНГРАД 1977

Печатается по постановлению
Редакционно-издательского совета
Ленинградского университета

УДК 517.5+519.3

Вопросы теории и элементы программного обеспечения минимаксных задач. Л., Изд-во Ленингр. ун-та, 1977. 192 с.

В монографии излагаются некоторые новые результаты по теории минимакса и аппроксимации функций в основном в области численных методов (в частности, методов выравнивания максимумов и экстремального базиса), приводится несколько схем с ускоренной сходимостью. Основное внимание уделяется разработке методов, более эффективных с точки зрения реализации на ЭВМ. Рассматривается целочисленная минимаксная задача. Приводятся программы (на языке АЛГОЛ-60) решения ряда задач, возникающих в теории минимакса. Каждая программа (процедура) сопровождается контрольным примером и некоторыми рекомендациями по практическому ее использованию.

Книга рассчитана на широкий круг лиц, специализирующихся в области математического программирования и вычислительной математики. Ил. - 11, табл. - 4, библиогр. - 183 назв.

В $\frac{20204-030}{076(02)-77}$ 68-77

© Издательство Ленинградского университета, 1977 г.

О Г Л А В Л Е Н И Е

Предисловие 5

Часть первая. ВОПРОСЫ ТЕОРИИ МИНИМАКСНЫХ ЗАДАЧ

§ I.I. Альтернансная форма условия оптимальности (Малоземов В.Н.) 7

§ I.2. Сходимость алгоритма Ремеза в нелинейном случае (Малоземов В.Н.). 19

§ I.3. Выравнивание максимумов (Малоземов В.Н., Певный А.Б.) 27

§ I.4. Выравнивание максимумов. Теорема о квадратичной скорости сходимости (Васильев Л.В.) 34

§ I.5. Метод обобщенного градиента с полной релаксацией (Васильев Л.В., Демьянов В.Ф.). 40

§ I.6. Метод экстремального базиса (Демьянов В.Ф.) . . 52

§ I.7. О непрерывных методах минимизации выпуклых функций (Демьянов В.Ф., Рубинов А.М.) 65

§ I.8. Метод возможных направлений без управляющих последовательностей (Васильев Л.В., Тарасов В.Н.) 71

§ I.9. Ускорение сходимости при решении нелинейных минимаксных задач (Тарасов В.Н.) 76

§ I.IO. Совместное приближение (Малоземова Л.К.). . . . 81

§ I.II. Аппроксимация с интерполяцией (Даугавет В.А., Малоземов В.Н.). 88

§ I.I2. Аппроксимация с двусторонними ограничениями (Даугавет В.А.) 95

§ I.I3. О положительных полиномах (Малоземов В.Н.) . 100

§ I.I4. О чувствительности величины наилучшего приближения (Певный А.Б.). 104

§ I.I5. Одна задача со связанными ограничениями (Войтон Е.Ф.). 106

§ I.I6. Целочисленная аппроксимация задачи оптимального размещения прямоугольников в трапеции (Уфимцев Г.В., Элиашберг Я.М.) 114

Часть вторая. ЭЛЕМЕНТЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ МИНИМАКСНЫХ ЗАДАЧ

§ 2.1. Решение линейных задач аппроксимации при ограничениях (Даугавет В.А.) 128
§ 2.2. Построение алгебраического полинома наилучшего приближения (Малоземов В.Н.) 135
§ 2.3. Ускоренное выравнивание (Рыскелдиев У.А.). . . . 138
§ 2.4. Нахождение разбиений с равными уклонениями (Певный А.Б.). 141
§ 2.5. Решение задачи дробно-рациональной аппроксимации (Белых В.М.) 145
§ 2.6. Решение задачи синтеза электрического фильтра (Белых В.М.) 151
§ 2.7. Определение расстояния от конуса до многогогранника (Даугавет В.А.) 161
§ 2.8. Нахождение направления наискорейшего спуска в минимаксных задачах (Васильев Л.В.). 164
§ 2.9. Обобщенный градиентный спуск с растяжением пространства (Васильев А.А.). 168
§ 2.10. Процедура выравнивания максимумов (Васильев Л.В.) 172
§ 2.11. Вычисление ограничителя шага спуска (Васильев Л.В.) 175
§ 2.12. Контрольный пример для процедур rest, mmax, lp, down (Васильев Л.В.) 177
Указатель литературы 180

ПРЕДИСЛОВИЕ

Интерес к минимаксным задачам, и в частности к задачам равномерной аппроксимации, объясняется прежде всего тем, что они образуют достаточно широкий и важный, с практической точки зрения, класс экстремальных задач с негладкой целевой функцией. К этому классу мы относим также задачи на минимум выпуклых функций, ибо любая выпуклая функция может быть представлена как максимум линейных функций.

Прогресс в теории минимакса обусловлен развитием общей теории экстремальных задач [5, 21, 22, 44, 49, 50, 53, 57, 62, 63, 67, 79, 80, 89, 97, 98, 101, 102, 105, 109, 114, 164]. Однако в рамках общей теории минимакс приобрел относительную самостоятельность. За последние годы появились монографии, сборники и обзорные статьи, частично или полностью посвященные минимаксным задачам [19, 28, 34, 38-40, 59, 60, 66, 100, 121-123, 128, 129, 132, 140, 153, 163, 173, 174]. Получены более общие необходимые [6, 71, 139] и достаточные [75] условия минимакса, исследованы альтернансные свойства решений нелинейных минимаксных задач с ограничениями [30, 32, 76]. Существенно продвинуто решение вопроса о дифференцируемости по направлениям функции максимума и функции последовательного максимина, найдены формулы для производных по направлениям второго и более высокого порядка (см. монографию [34] и указатель литературы в ней).

Появились новые направления исследований - минимаксные задачи со связанными ограничениями [4, 13, 14, 24, 29, 41, 65, 86, 87, 104, 106, 107, 142], задачи оптимального управления с минимаксным критерием оптимальности [12, 16, 27, 36]. Расширяются приложения минимаксной теории, особенно к задачам проектирования и синтеза [2, 8, 64].

Большое внимание уделяется методам решения минимаксных задач [23, 25, 26, 31, 45, 47, 51, 52, 54, 58, 68, 70, 72, 83, 96, 119, 135, 154, 155, 161, 165, 166, 168-170, 180], в частности нахождению направления наискорейшего спуска [9, 10, 15, 88], и методам, обладающим высокой скоростью сходимости в окрестности стационарной точки [7, 33, 38, 74, 147, 156, 175]. Получены полезные оцен-

ки и установлена скорость сходимости некоторых методов [42, 85]. Одной из основных задач в этой области в настоящее время являются разработка программного обеспечения минимакса и его широкое внедрение в практику. Первые шаги в этом направлении уже сделаны [3, 78, 81, 94, 95].

Данная книга примыкает к монографии [39] и может рассматриваться как ее продолжение. По крайней мере основные результаты из [39] считаются известными.

Книга состоит из двух частей. Первая часть посвящена в основном методам решения минимаксных задач и характеризации в альтернансных терминах точки локального минимума функции максимума. Вторая часть содержит некоторые элементы программного обеспечения минимаксных задач в виде АЛГОЛ-процедур. Особенно широко представлены процедуры по решению задач равномерной аппроксимации.

Авторы выражают глубокую благодарность И.К.Даугавету и И.В.Романовскому за многочисленные замечания и предложения, способствовавшие улучшению книги.

Часть первая

ВОПРОСЫ ТЕОРИИ МИНИМАКСНЫХ ЗАДАЧ

§ 1.1. Альтернансная форма условия оптимальности

Рассматривается задача

$$\max_{Y\in G} F(X,Y) \to \min_{X\in\Omega}, \tag{1}$$

где $\Omega=\{X\in E_n \mid h(X,\beta)\le 0 \quad \forall\,\beta\in\omega\}$. Предполагается, что G и ω – непустые метрические компакты; $h(X,\beta)$ – функция, непрерывная вместе с $h'_X(X,\beta)$ на $E_n\times\omega$ и выпуклая по X при каждом фиксированном $\beta\in\omega$; $F(X,Y)$ – функция, непрерывная вместе с $F'_X(X,Y)$ на $\Omega'\times G$, где $\Omega'\subset E_n$ – некоторое открытое множество, содержащее Ω (чаще всего $\Omega'=E_n$). Кроме того, считается выполненным условие Слейтера: при некотором $\bar X\in E_n$

$$h(\bar X,\beta)<0 \quad \forall\,\beta\in\omega. \tag{2}$$

Если ω – дискретное множество, $\omega=1:s$, и $h(X,j)=(A_j,X)+b_j$ при $j\in s_1+1:s$, то формулировка условия Слейтера немного меняется: существует точка $\bar X\in E_n$, в которой

$$\begin{aligned} &h(\bar X,j)<0 \quad \forall j\in 1:s_1,\\ &h(\bar X,j)\le 0 \quad \forall j\in s_1+1:s. \end{aligned} \tag{2'}$$

В этом параграфе изучаются альтернансные свойства решений задачи (1).

1°. Установим вначале несколько вспомогательных предложений.

Определение 1. Система n-мерных векторов

$$U_0, U_1, \dots, U_m \tag{3}$$

называется p-независимой, где $p\in 0:m$, если из соотношений $\sum_{i=0}^{m}\alpha_i U_i=\mathbf{0}$, $\sum_{i=0}^{p}\alpha_i=0$ следует $\alpha_0=\alpha_1=\dots=\alpha_m=0$.

Очевидно, что m-независимость совпадает с аффинной независимостью системы (3), а 0-независимость равносильна линейной независимости векторов $U_1, \ldots, U_m$.

Пусть $p \geqslant 1$. Тогда для p-независимости системы (3) необходимо и достаточно, чтобы при любом $k \in 0:p$ векторы $U_0 - U_k, \ldots, U_{k-1} - U_k, U_{k+1} - U_k, \ldots, U_p - U_k, U_{p+1}, \ldots, U_m$ были линейно-независимы. Отсюда, в частности, следует, что p-независимая система содержит не более $n+1$ элементов (n-мерных векторов).

Пример. В E_2 векторы (1,0), (0,1), (0, -1) аффинно-независимы, но не являются 0-независимыми, а векторы (1,0), (1,0), (0,1) 0-независимы, но не являются аффинно-независимыми.

Лемма 1. Допустим, что векторы $U_0, U_1, \ldots, U_n$ p-независимы и существуют положительные числа $\alpha_0, \alpha_1, \ldots, \alpha_n$ такие, что $\Sigma_{i=0}^{n} \alpha_i U_i = \mathbf{0}$. Тогда определители Δ_i, составленные из столбцов $U_0, \ldots, U_{i-1}, U_{i+1}, \ldots, U_n$, отличны от нуля и имеют чередующиеся знаки: $\operatorname{sign} \Delta_{i-1} = -\operatorname{sign} \Delta_i$ $\forall i \in 1:n$.

Доказательство. По условию система линейных уравнений

$$\begin{aligned} &\Sigma_{i=0}^{n} z_i U_i = \mathbf{0}, \\ &\Sigma_{i=0}^{p} z_i = 1 \end{aligned} \tag{4}$$

имеет единственное решение $z_i = \alpha_i / \Sigma_{i=0}^{p} \alpha_i$. Поэтому определитель этой системы Δ отличен от нуля. По формуле Крамера

$$0 < z_i = (-1)^{n+i} \frac{\Delta_i}{\Delta}, \quad i \in 0:n,$$

откуда и следует требуемое.

Лемма 1 допускает обращение.

Лемма 2. Если для n-мерных векторов $U_0, U_1, \ldots, U_n$ определители Δ_i отличны от нуля и имеют чередующиеся знаки, то система $\Sigma_{i=0}^{n} \alpha_i U_i = \mathbf{0}$ имеет положительное решение, а векторы $U_0, U_1, \ldots, U_n$ p-независимы при любом $p \in 0:n$.

Доказательство. Зафиксируем $p \in 0:n$ и рассмотрим систему (4). Ее определитель $\Delta = \Sigma_{i=0}^{p} (-1)^{n+i} \Delta_i \neq 0$. Отсюда следуют и p-независимость векторов $U_0, U_1, \ldots, U_n$ и существование положительного решения:

$$z_i = \frac{(-1)^{n+i} \Delta_i}{\Delta} > 0, \quad i \in 0:n.$$

Лемма доказана.

В дальнейшем существенную роль будет играть следующая теорема.

Теорема 1. Линейная однородная система

$$\Sigma_{i=0}^{r-1} \alpha_i V_i = \mathbf{0}, \tag{5}$$

где $r \in 1:n+1$ и V_i – n-мерные векторы, имеет положительное решение, и столбцы $V_0, V_1, \ldots, V_{r-1}$ p-независимы ($p \in 0:r-1$) тогда и только тогда, когда:

I. Любой минор порядка r матрицы

$$B = \begin{pmatrix} v_1^{(0)} & \ldots & v_1^{(p)} & v_1^{(p+1)} & \ldots & v_1^{(r-1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ v_n^{(0)} & \ldots & v_n^{(p)} & v_n^{(p+1)} & \ldots & v_n^{(r-1)} \\ 1 & \ldots & 1 & 0 & \ldots & 0 \end{pmatrix},$$

не содержащий ее последнюю строку, равен нулю.

II. Существует ненулевой минор Δ порядка r такой, что все определители Δ_i, которые получают из Δ исключением i-го столбца и последней строки, отличны от нуля и имеют чередующиеся знаки: $\operatorname{sign} \Delta_{i-1} = -\operatorname{sign} \Delta_i$ $\forall i \in 1:r-1$.

Доказательство. Необходимость. Допустим, что система (5) имеет положительное решение. Тогда очевидно выполняется условие I. Далее, поскольку векторы $V_0, V_1, \ldots, V_{r-1}$ p-независимы, то существует ненулевой минор Δ порядка r матрицы B. Пусть для определенности Δ образуют 1-я, ..., $(r-1)$-я и $(n+1)$-я строки. Тогда для $(r-1)$-мерных векторов $U_i = (v_1^{(i)}, \ldots, v_{r-1}^{(i)})$, $i \in 0:r-1$, выполнены условия леммы 1 (с заменой n на $r-1$), откуда на основании леммы 1 и следует II.

Достаточность. Допустим для определенности, что ненулевой минор порядка r образуют 1-я, ..., $(r-1)$-я и $(n+1)$-я строки матрицы B. Тогда, во-первых, векторы $V_0, V_1, \ldots, V_{r-1}$ p-независимы, а, во-вторых, в силу II и леммы 2 система

$$\Sigma_{i=0}^{r-1} \alpha_i v_j^{(i)} = 0, \quad j \in 1:r-1,$$

имеет положительное решение $\alpha_0^*, \alpha_1^*, \ldots, \alpha_{r-1}^*$. Нетрудно проверить, что и для остальных $j \in r:n$

$$\Sigma_{i=0}^{r-1} \alpha_i^* v_j^{(i)} = 0. \tag{6}$$

Действительно, зафиксируем $j \in r:n$ и рассмотрим матрицу

$$\begin{pmatrix} v_1^{(0)} & \dots & v_1^{(r-1)} \\ \dots & \dots & \dots \\ v_{r-1}^{(0)} & \dots & v_{r-1}^{(r-1)} \\ v_j^{(0)} & \dots & v_j^{(r-1)} \end{pmatrix}.$$

На основании условий I и II ее ранг равен $r-1$, причем первые $r-1$ строк линейно-независимы. Значит, последняя строка является линейной комбинацией первых $r-1$ строк, откуда очевидным образом следует (6). Теорема доказана.

З а м е ч а н и е 1. При $r=1$ условия I, II вырождаются в $V_0 = \mathbf{0}$, а при $r=n+1$ равносильны следующему: определители Δ_i, составленные из столбцов $V_0, \dots, V_{i-1}, V_{i+1}, \dots, V_n$, отличны от нуля и имеют чередующиеся знаки. Разложением по элементам последней строки можно проверить, что при $r=n+1$

$$\Delta \stackrel{\text{def}}{=} \det(B) \neq 0 .$$

З а м е ч а н и е 2. Если система (5) имеет положительное решение и векторы V_i p-независимы, то они автоматически являются аффинно-независимыми. На основании леммы 2 это очевидно для векторов U_i, построенных при доказательстве необходимости, а значит, и для V_i.

2°. Обратимся к задаче (1). По определению Ω - замкнутое выпуклое множество. Введем обозначения:

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

$$R(X) = \{ Y \in G \mid F(X,Y) = \varphi(X) \},$$

$$Q(X) = \{ \beta \in \omega \mid h(X,\beta) = 0 \}.$$

Как известно [39, с. 241], справедливо следующее предложение.

Предложение 1. Для того чтобы минимум функции φ на Ω достигался в точке $X^* \in \Omega$, необходимо, а в случае выпуклости φ на Ω и достаточно, чтобы

$$L(X^*) \cap \Gamma^+(X^*) \neq \emptyset . \qquad (7)$$

Здесь $L(X^*)$ - выпуклая оболочка множества $\{ F'_X(X^*,Y) \mid Y \in R(X^*) \}$, а $\Gamma^+(X^*)$ - конус, сопряженный конусу возможных направлений множества Ω в точке X^*. В нашем случае в силу конкретного задания Ω и выполнения условия Слейтера $\Gamma^+(X^*)$ совпадает с выпуклой конической оболочкой множества $\{ -h'_X(X^*,\beta) \mid \beta \in Q(X^*) \}$ (см., например, [39, с. 180-182, 251]).

Теорема 2. Для того чтобы минимум функции φ на Ω достигался в точке $X^* \in \Omega$, необходимо, а в случае выпуклости φ на Ω и достаточно, чтобы нашлись точки $Y_0, Y_1, \dots, Y_p$ в $R(X^*)$, $\beta_1, \dots, \beta_q$ в $Q(X^*)$ и положительные числа $\alpha_0, \alpha_1, \dots, \alpha_p, \lambda_1, \dots, \lambda_q$, где $p \geqslant 0$, $q \geqslant 0$, такие, что

$$\Sigma_{i=0}^{p} \alpha_i F'_X(X^*, Y_i) + \Sigma_{j=1}^{q} \lambda_j h'_X(X^*, \beta_j) = \mathbf{0}, \qquad (8)$$

и векторы

$$F'_X(X^*,Y_0), \dots, F'_X(X^*,Y_p), h'_X(X^*,\beta_1), \dots, h'_X(X^*,\beta_q) \qquad (9)$$

образовали p-независимую систему.

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Согласно предложению 1 найдется вектор V^*, общий для $L(X^*)$ и $\Gamma^+(X^*)$. По определению $L(X^*)$

$$V^* = \Sigma_{i=0}^{q} \alpha_i F'_X(X^*, Y_i), \qquad (10)$$

где $Y_0, Y_1, \dots, Y_p$ из $R(X^*)$ и $\alpha_0, \alpha_1, \dots, \alpha_p$ - положительные числа, в сумме равные единице. По определению $\Gamma^+(X^*)$

$$V^* = -\Sigma_{j=1}^{q} \lambda_j h'_X(X^*, \beta_j), \qquad (11)$$

где $\beta_1, \dots, \beta_q$ из $Q(X^*)$ и $\lambda_1, \dots, \lambda_q$ - положительные числа. Вычитая (11) из (10), получаем (8).

Покажем, что если векторы (9) p-зависимы, то количество слагаемых в левой части (8) можно уменьшить. Действительно, пусть существуют числа

$$\bar{\alpha}_0, \bar{\alpha}_1, \dots, \bar{\alpha}_p, \bar{\lambda}_1, \dots, \bar{\lambda}_q, \qquad (12)$$

не все равные нулю и такие, что

$$\Sigma_{i=0}^{p} \bar{\alpha}_i F'_X(X^*, Y_i) + \Sigma_{j=1}^{q} \bar{\lambda}_j h'_X(X^*, \beta_j) = \mathbf{0},$$

$$\Sigma_{i=0}^{p} \bar{\alpha}_i = 0 .$$

При этом можно считать, что среди чисел (12) есть хотя бы одно положительное. Тогда каково бы ни было вещественное ε,

$$\Sigma_{i=0}^{p}(\alpha_i-\varepsilon\bar{\alpha}_i)\,F'_x(X^*,Y_i)+\Sigma_{j=1}^{q}(\lambda_j-\varepsilon\bar{\lambda}_j)h'_x(X^*,\beta_j)=\mathbf{0},$$

$$\Sigma_{i=0}^{p}(\alpha_i-\varepsilon\bar{\alpha}_i)=1\,. \qquad (13)$$

Положим

$$\bar{\varepsilon}=\min_{\bar{\alpha}_i>0,\,\bar{\lambda}_j>0}\left\{\frac{\alpha_i}{\bar{\alpha}_i}\,,\ \frac{\lambda_j}{\bar{\lambda}_j}\right\},$$

$$\alpha'_i=\alpha_i-\bar{\varepsilon}\bar{\alpha}_i\quad(i\in 0:p),\qquad \lambda'_j=\lambda_j-\bar{\varepsilon}\bar{\lambda}_j\quad(j\in 1:q).$$

Нетрудно проверить, что $\alpha'_i\geqslant 0$, $\lambda'_j\geqslant 0$ и хотя бы одно из чисел α'_i, λ'_j обращается в нуль. Оставив только ненулевые α'_i, λ'_j, на основании (13) придем к равенству вида (8) с меньшим числом слагаемых. Повторив при необходимости эту процедуру конечное число раз, получим представление нуля (8) с p-независимой системой векторов (9), что и требовалось доказать.

Достаточность очевидным образом следует из предложения 1. Теорема доказана.

Замечание 1. Поскольку система векторов (9) p-независима, то $p+q\leqslant n$. При этом не исключается равенство $p+q=0$.

Замечание 2. Теорема 2 содержит как частный случай (когда G состоит из одного элемента) теорему Куна - Таккера в дифференциальной форме.

3°. Оказывается, что теореме 2 можно придать равносильную альтернансную форму. Предварительно введем определение.

Определение 2. Будем говорить, что точка $X^*\in\Omega$ обладает r-точечным альтернансом ($r\in 1:n+1$), если существуют $Y_0,Y_1,\ldots,Y_p$ из $R(X^*)$ и $\beta_1,\ldots,\beta_q$ из $Q(X^*)$, где $p\geqslant 0$, $q\geqslant 0$, $p+q+1=r$, такие, что:

I. Любой минор порядка r матрицы

$$B=\begin{pmatrix} F'_{x_1}(X^*,Y_0)\ldots F'_{x_1}(X^*,Y_p) & h'_{x_1}(X^*,\beta_1)\ldots h'_{x_1}(X^*,\beta_q)\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ F'_{x_n}(X^*,Y_0)\ldots F'_{x_n}(X^*,Y_p) & h'_{x_n}(X^*,\beta_1)\ldots h'_{x_n}(X^*,\beta_q)\\ 1\quad\ldots\quad 1 & 0\quad\ldots\quad 0 \end{pmatrix},$$

не содержащий ее последнюю строку, равен нулю.

II. Существует ненулевой минор Δ порядка r матрицы B такой, что все определители Δ_i, которые получаются из Δ исключением i-го столбца и последней строки, отличны от нуля и имеют чередующиеся знаки: $\operatorname{sign}\Delta_{i-1}=-\operatorname{sign}\Delta_i\quad\forall i\in 1:r-1$.

Оба условия этого определения существенны, когда $r\in 2:n$. Рассмотрим случаи $r=1$ и $r=n+1$. Будем говорить, что $X^*\in\Omega$ обладает одноточечным альтернансом, если существует $Y_0\in R(X^*)$, для которого $F'_x(X^*,Y_0)=\mathbf{0}$. Наличие $(n+1)$-точечного (полного) альтернанса равносильно существованию $Y_0,Y_1,\ldots,Y_p$ из $R(X^*)$ и $\beta_1,\ldots,\beta_q$ из $Q(X^*)$, где $p\geqslant 0$, $q\geqslant 0$, $p+q=n$, таких, что определители Δ_i подматриц, получаемых из B исключением i-го столбца и последней строки, отличны от нуля и имеют чередующиеся знаки.

Справедлива следующая теорема.

Теорема 3. Для того чтобы $X^*\in\Omega$ была точкой минимума функции φ на Ω, необходимо, а в случае выпуклости φ на Ω и достаточно, чтобы X^* обладала r-точечным альтернансом, где $r\in 1:n+1$.

Доказательство непосредственно следует из теорем 2 и 1.

Аналогом известной в теории аппроксимации теоремы Валле-Пуссена является следующая теорема.

Теорема 4. Пусть $X_0\in\Omega$ и нашлись $Y_0,Y_1,\ldots,Y_p$ из G и $\beta_1,\ldots,\beta_q$ из $Q(X_0)$, где $p\geqslant 0$, $q\geqslant 0$, $p+q+1=r$, такие, что матрица

$$B=\begin{pmatrix} F'_x(X_0,Y_0)\ldots F'_x(X_0,Y_p) & h'_x(X_0,\beta_1)\ldots h'_x(X_0,\beta_q)\\ 1\quad\ldots\quad 1 & 0\quad\ldots\quad 0 \end{pmatrix}$$

удовлетворяет условиям I, II определения r-точечного альтернанса и функции $F(X,Y_i)$ при $i\in 0:p$ выпуклы на Ω. Тогда

$$\inf_{X\in\Omega}\varphi(X)\geqslant\min_{i\in 0:p}F(X_0,Y_i). \qquad (14)$$

Доказательство. Положим $\tilde{G}=\{Y_0,Y_1,\ldots,Y_p\}$, $\tilde{F}(X,Y)=F(X,Y)-F(X_0,Y)$ и рассмотрим задачу

$$\tilde{\varphi}(X)\overset{\text{def}}{=}\max_{Y\in\tilde{G}}\tilde{F}(X,Y)\to\min_{X\in\Omega}\,.$$

очевидно, что $\tilde{\varphi}$ - выпуклая на Ω функция, $\tilde{\varphi}(X_0)=0$, $\tilde{G}(X_0)\overset{\text{def}}{=}\{Y\in\tilde{G}\mid\tilde{F}(X_0,Y)=\tilde{\varphi}(X_0)\}=\tilde{G}$ и $\tilde{F}'_x(X_0,Y)=F'_x(X_0,Y)$. По условию X_0 обладает r-точечным альтернансом и в силу теоремы 3 является точкой минимума функции $\tilde{\varphi}$ на Ω: $\tilde{\varphi}(X)\geqslant 0\quad\forall X\in\Omega$.

Значит,

$$\max_{Y\in\widetilde{G}}\{F(X,Y)-F(X_0,Y)\}\geqslant 0 \quad \forall X\in\Omega,$$

откуда $\varphi(X)\geqslant \min_{i\in 0:p} F(X_0,Y_i)$ $\forall X\in\Omega$, что равносильно (14). Теорема доказана.

4°. В этом пункте будет найдено минимальное дополнительное условие, гарантирующее существование полного альтернанса. Предварительно введем определение.

Определение 3. Точку $X^*\in\Omega$ назовем чебышевской точкой функции φ на Ω, если для любых $Y_0, Y_1, \ldots, Y_p$ из $R(X^*)$ и любых $\beta_1, \ldots, \beta_q$ из $Q(X^*)$ при $p+q\leqslant n-1$ существуют коэффициенты $\alpha_1, \ldots, \alpha_n$ такие, что

$$\begin{aligned}&\textstyle\sum_{i=1}^{n}\alpha_i F'_{x_i}(X^*,Y_j)>0 \quad \forall j\in 0:p,\\ &\textstyle\sum_{i=1}^{n}\alpha_i h'_{x_i}(X^*,\beta_j)\geqslant 0 \quad \forall j\in 1:q.\end{aligned} \tag{15}$$

Лемма 3. Для того чтобы точка $X^*\in\Omega$ была чебышевской точкой функции φ на Ω, необходимо и достаточно, чтобы для любых $Y_0, Y_1, \ldots, Y_p$ из $R(X^*)$, $\beta_1, \ldots, \beta_q$ из $Q(X^*)$ таких, что векторы

$$F'_X(X^*,Y_0), \ldots, F'_X(X^*,Y_p), h'_X(X^*,\beta_1), \ldots, h'_X(X^*,\beta_q) \tag{16}$$

p-независимы, и любых положительных $\alpha_0, \alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_q$ при $p+q\leqslant n-1$ было невозможно равенство

$$\textstyle\sum_{i=0}^{p}\alpha_i F'_X(X^*,Y_i)+\sum_{j=1}^{q}\lambda_j h'_X(X^*,\beta_j)=\mathbf{0}. \tag{17}$$

Доказательство. Необходимость. Допустим, что при некоторых положительных α_i, λ_j выполняется равенство (17). Тогда

$$V\overset{\text{def}}{=}\textstyle\sum_{i=0}^{p}\alpha_i F'_X(X^*,Y_i)=-\sum_{j=1}^{q}\lambda_j h'_X(X^*,\beta_j). \tag{18}$$

По условию разрешима система (15), т.е. существует вектор $A=(\alpha_1, \ldots, \alpha_n)$ такой, что

$$(A, F'_X(X^*,Y_j))>0 \quad \forall j\in 1:p, \tag{19}$$

$$(A, h'_X(X^*,\beta_j))\geqslant 0 \quad \forall j\in 1:q. \tag{20}$$

Учитывая определение V и (19), получаем $(A,V)>0$. Однако в силу (18) и (20) $(A,V)\leqslant 0$. Полученное противоречие завершает доказательство необходимости.

Достаточность. Возьмем произвольные $Y_0, Y_1, \ldots, Y_p$ из $R(X^*)$ и $\beta_1, \ldots, \beta_q$ из $Q(X^*)$, где $p+q\leqslant n-1$, и докажем разрешимость системы (15). Обозначим через L выпуклую оболочку множества $\{F'_X(X^*,Y_i) \mid i\in 0:p\}$, а через K – выпуклую коническую оболочку множества $\{-h'_X(X^*,\beta_j) \mid j\in 1:q\}$. Покажем, что $L\cap K=\varnothing$. Допустим противное. Тогда справедливо равенство (17) с неотрицательными α_i, λ_j, причем $\sum_{i=0}^{p}\alpha_i=1$. Уменьшая при необходимости количество слагаемых в (17) так, как это делалось при доказательстве теоремы 2, можно добиться того, что в (17) все α_i и λ_j будут положительными, а векторы (16) – p-независимыми. Но это противоречит условию леммы. Итак, $L\cap K=\varnothing$. По теореме о строгой отделимости найдется вектор A такой, что

$$(A,V)>0 \quad \forall V\in L,$$

$$(A,V)\leqslant 0 \quad \forall V\in K.$$

В частности, $A=(\alpha_1, \ldots, \alpha_n)$ является решением системы (15). Лемма доказана.

Теорема 5. Если точка локального минимума X^* функции φ на Ω является чебышевской, то она обладает полным альтернансом.

Доказательство. Прежде всего заметим, что если X^* – точка локального минимума функции φ на Ω, то выполняется (7), и, значит, справедливо заключение теоремы 2. При этом поскольку X^* – чебышевская точка, то согласно лемме 3 $p+q=n$. Последнее утверждение равносильно существованию полного альтернанса. Теорема доказана.

5°. Наличие полного альтернанса является достаточным условием для того, чтобы X^* была точкой строгого локального минимума функции φ на Ω. Для доказательства этого вспомним следующие формулы [39, с. 233-237]:

$$\frac{\partial\varphi(X^*)}{\partial g}\overset{\text{def}}{=}\lim_{\lambda\to+0}\frac{\varphi(X^*+\lambda g)-\varphi(X^*)}{\lambda}=\max_{Y\in R(X^*)}(F'_X(X^*,Y),g), \tag{21}$$

$$\varphi(X^*+\lambda g)=\varphi(X^*)+\lambda\frac{\partial\varphi(X^*)}{\partial g}+o(g;\lambda), \tag{22}$$

где g – любой единичный вектор и $o(g;\lambda)/\lambda \xrightarrow[\lambda\to+0]{} 0$ равномерно относительно g.

Обозначим через $B_r(X^*)$ замкнутый шар радиусом r с центром в точке X^*, а через $\Gamma S(X^*)$ – совокупность единичных векторов из замыкания конуса возможных направлений $\Gamma(X^*)$ множества Ω в точке X^*.

Справедливо следующее предложение.

Предложение 2. Пусть в точке $X^*\in\Omega$ выполняется неравенство $\min\limits_{g\in\Gamma S(X^*)} \partial\varphi(X^*)/\partial g > 0$. Тогда X^* является точкой строгого локального минимума функции φ на Ω. Более того, найдутся положительные r и η такие, что

$$\varphi(X) \geqslant \varphi(X^*) + \eta\|X - X^*\| \quad \forall X\in B_r(X^*)\cap\Omega. \tag{23}$$

Доказательство. Положим $\eta = \frac{1}{2}\min\limits_{g\in\Gamma S(X^*)} \partial\varphi(X^*)/\partial g$ и выберем такое $r>0$, что в разложении (22) для $\lambda\in(0, r]$ равномерно по $g\in\Gamma S(X^*)$ будет

$$|o(g;\lambda)| \leqslant \eta\lambda. \tag{24}$$

Возьмем $X\in B_r(X^*)\cap\Omega$ и покажем, что выполняется неравенство (23). Если $X = X^*$, то это очевидно. Пусть $X\neq X^*$. Тогда $X = X^*+\lambda g$, где $\|g\| = 1$ и $\lambda = \|X - X^*\|\in(0, r]$. Очевидно, что $g\in\Gamma S(X^*)$. На основании (22), (24) и определения λ получаем

$$\varphi(X) \geqslant \varphi(X^*) + 2\eta\lambda - |o(g;\lambda)| \geqslant \varphi(X^*) + \eta\|X - X^*\|.$$

Утверждение доказано.

Теорема 6. Если точка $X^*\in\Omega$ обладает полным альтернансом, то она является точкой строгого локального минимума функции φ на Ω. Более того, выполняется неравенство (23) при некоторых $r>0$ и $\eta>0$.

Доказательство. По условию существуют точки $Y_0, Y_1, \ldots, Y_p$ из $R(X^*)$, $\beta_1, \ldots, \beta_q$ из $Q(X^*)$, где $p\geqslant 0$, $q\geqslant 0$, $p+q = n$, и положительные числа $\alpha_0, \alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_q$ такие, что

$$\sum_{i=0}^{p}\alpha_i F'_X(X^*, Y_i) + \sum_{j=1}^{q}\lambda_j h'_X(X^*, \beta_j) = \mathbf{0}, \tag{25}$$

причем векторы

$$F'_X(X^*, Y_0), \ldots, F'_X(X^*, Y_p), h'_X(X^*, \beta_1), \ldots, h'_X(X^*, \beta_q) \tag{26}$$

p-независимы. Поскольку в рассматриваемом случае

$$\Gamma(X^*) = \{V\in E_n \mid (V, h'_X(X^*, \beta))\leqslant 0 \quad \forall\beta\in Q(X^*)\} \tag{27}$$

(см., например, [39, с. 177, 246]), то для любого $g\in\Gamma S(X^*)$ с учетом (25) получаем

$$\sum_{i=0}^{p}\alpha_i(F'_X(X^*, Y_i), g) \geqslant 0.$$

Отсюда

$$\max_{i\in 0:p}(F'_X(X^*, Y_i), g) \geqslant 0. \tag{28}$$

Покажем, что в (28) равенство нулю невозможно. Действительно, в противном случае $(F'_X(X^*, Y_i), g) = 0$ $\forall i\in 0:p$, и в силу (25) и (27) $(h'_X(X^*, \beta_j), g) = 0$ $\forall j\in 1:q$. Так как система векторов (26) p-независима, то $g = \mathbf{0}$. Получили противоречие, ибо $\|g\| = 1$. Значит,

$$\max_{i\in 0:p}(F'_X(X^*, Y_i), g) > 0 \quad \forall g\in\Gamma S(X^*).$$

Отсюда и из (21) следует, что

$$\frac{\partial\varphi(X^*)}{\partial g} > 0 \quad \forall g\in\Gamma S(X^*).$$

Остается сослаться на предложение 2. Теорема доказана.

Замечание 1. При отсутствии ограничений эта теорема была установлена в [76], в общем случае – в [30].

Замечание 2. В условиях теоремы 5 точка X^* автоматически является точкой строгого локального минимума функции φ на Ω.

6°. Аналогичные результаты могут быть получены и для нелинейных задач аппроксимации:

$$\varphi_1(X) \stackrel{\text{def}}{=} \max_{t\in D}|f(X,t)| \to \min_{X\in\Omega}.$$

Здесь множество Ω то же, что и в (1), функция $f(X,t)$ непрерывна вместе с $f'_X(X,t)$ на $\Omega'\times D$, где Ω' – открытое множество, содержащее Ω, а D – непустой компакт. Кроме того, предполагается, что $\inf\limits_{X\in\Omega}\varphi_1(X) > 0$.

Не будем на этом подробно останавливаться. Заметим лишь, что

$$\varphi_1(X) = \max_{Y \in G} F(X,Y),$$

где $G = \{-1,1\} \times D$, $Y = (\xi, t)$, $F(X,Y) = \xi f(X,t)$. Покажем, как в данном случае выглядит определение полного альтернанса, и приведем без доказательства одну теорему, которая будет использована в дальнейшем.

Пусть $R_1(X) = \{ t \in D \mid |f(X,t)| = \varphi_1(X) \}$.

Определение 4. Говорят, что точка $X^* \in \Omega$ обладает полным альтернансом, если существуют точки $t_0, t_1, \ldots, t_p$ из $R_1(X^*)$ и $\beta_1, \ldots, \beta_q$ из $Q(X^*)$, где $p \geqslant 0$, $q \geqslant 0$, $p+q = n$, такие, что матрица $\mathcal{F}'$, составленная из столбцов

$$f'_X(X^*, t_0), \ldots, f'_X(X^*, t_p), h'_X(X^*, \beta_1), \ldots, h'_X(X^*, \beta_q),$$

удовлетворяет следующему условию: определители Δ_i подматриц, получаемых из $\mathcal{F}'$ исключением i-го столбца, отличны от нуля, а величины $\xi_i \Delta_i$, где

$$\xi_i = \begin{cases} \operatorname{sign} f(X^*, t_i), & \text{если } i \in 0:p, \\ 1, & \text{если } i \in p+1:n, \end{cases}$$

имеют чередующиеся знаки:

$$\operatorname{sign}\{\xi_{i-1}\Delta_{i-1}\} = -\operatorname{sign}\{\xi_i \Delta_i\} \quad \forall i \in 1:n.$$

Теорема 7. Если точка $X^* \in \Omega$ обладает полным альтернансом, то она является точкой строгого локального минимума функции φ_1 на Ω. Более того, найдутся $r > 0$ и $\eta > 0$ такие, что

$$\varphi_1(X) \geqslant \varphi_1(X^*) + \eta \|X - X^*\| \quad \forall X \in B_r(X^*) \cap \Omega.$$

В заключение отметим, что существуют другие варианты альтернансной теории нелинейных задач аппроксимации в случае отсутствия ограничений ([162, 172-174, 178], [125, 159, 160]). Обзор результатов по линейным задачам аппроксимации приведен в [153].

§ 1.2. Сходимость алгоритма Ремеза в нелинейном случае

Весьма эффективным средством решения линейных задач аппроксимации является метод последовательных чебышевских интерполяций, или, как его часто называют, второй алгоритм Ремеза (см. [39, 66, 99, 100, 130]). Ральстон [171] применил этот метод к решению задач дробно-рациональной аппроксимации и установил его сходимость с достаточно хорошего начального приближения. Баррар и Лоуб [126] обобщили результат Ральстона на случай аппроксимации элементами нелинейных семейств, для которых справедлива альтернансная теория Майнардуса - Шведта [159, 160]. В данном параграфе доказывается локальная сходимость метода последовательных чебышевских интерполяций в существенно более общей ситуации, чем в [126].

1°. Рассмотрим задачу

$$\varphi(X) \stackrel{\text{def}}{=} \max_{t \in [\alpha, \beta]} |f(X,t)| \to \min_{X \in U},$$

где $U \subset E_n$ - открытое множество и $f(X,t)$ - функция, непрерывная вместе с $f'_X(X,t)$ на $U \times [\alpha, \beta]$.

Допустим, что в точке $X^* \in U$ выполнены следующие условия:

I. $\varphi^* \stackrel{\text{def}}{=} \varphi(X^*) > 0$.

II. Множество $R(X^*) = \{ t \in [\alpha, \beta] \mid |f(X^*, t)| = \varphi(X^*) \}$ содержит ровно $n+1$ точек $t_0^* < t_1^* < \ldots < t_n^*$ и

$$\operatorname{sign} f(X^*, t_{i-1}^*) = -\operatorname{sign} f(X^*, t_i^*) \quad \forall i \in 1:n.$$

III. Все определители Δ_i^*, составленные из столбцов

$$f'_X(X^*, t_0^*), \ldots, f'_X(X^*, t_{i-1}^*), f'_X(X^*, t_{i+1}^*), \ldots, f'_X(X^*, t_n^*),$$

отличны от нуля и одного знака.

Наша цель - показать, что перечисленных условий достаточно для локальной сходимости алгоритма Ремеза.

2°. Прежде всего заметим, что в силу I-III точка X^* обладает полным альтернансом. На основании теоремы 7 из § 1.1 X^* является точкой строгого локального минимума функции φ и, более того,

$$\varphi(X) \geqslant \varphi(X^*) + \eta_0 \|X - X^*\| \quad \forall x \in B_{r_0}(X^*) \tag{1}$$

при некоторых $r_0 > 0$ и $\eta_0 > 0$ (напомним, что $B_{r_0}(X^*) = \{ X \mid \|X - X^*\| \leqslant r_0 \}$). В частности,

$$\min_{X\in B_{r_0}(X^*)} \varphi(X) = \varphi^* > 0 . \qquad (2)$$

Обозначим через T множество базисов $\tau = (t_0, t_1, \ldots, t_n)$, где $t_i \in [\alpha, \beta]$ и $t_0 < t_1 < \ldots < t_n$ (в частности, $\tau^* = (t_0^*, t_1^*, \ldots, t_n^*)$), через $T_\delta(\tau^*)$ - совокупность базисов τ, для которых $|t_i - t_i^*| \leqslant \delta$ $\forall i \in 0:n$ (при этом предполагается, что отрезки $[t_i^* - \delta, t_i^* + \delta]$ не имеют общих точек), а через CT_ρ - совокупность тех базисов τ, у которых $|t_{i-1} - t_i| \geqslant \rho$ $\forall i \in 1:n$. Очевидно, что $T_\delta(\tau^*)$ и CT_ρ - компакты.

Введем обозначение $\gamma(X, \tau) = \min\limits_{i\in 0:n} |f(X, t_i)|$. Пару $\{X, \tau\}$ назовем базисной парой, если $\gamma(X, \tau) > 0$ и

$$\operatorname{sign} f(X, t_{i-1}) = -\operatorname{sign} f(X, t_i) \quad \forall i \in 1:n .$$

Экстремальную базисную пару определим как базисную пару, у которой $\gamma(X, \tau) = \varphi(X)$ (очевидно, что в данном случае $|f(X, t_i)| = \varphi(X)$ $\forall i \in 0:n$). В частности, $\{X^*, \tau^*\}$ - экстремальная базисная пара. Если $\{X, \tau\}$ - экстремальная базисная пара, то τ назовем экстремальным базисом, соответствующим X. Заметим, что по условию II существует единственный экстремальный базис τ^*, соответствующий X^*.

Лемма 1. По выбранному $r_0 > 0$ найдется $\rho > 0$ такое, что если для базисной пары $\{X, \tau\}$, где $X \in B_{r_0}(X^*)$, будет $\gamma(X, \tau) \geqslant \varphi^*/2$, то $\tau \in CT_\rho$.

Доказательство. Допустим противное. Тогда найдется последовательность базисных пар $\{X_k, \tau_k\}$ со свойствами $t_i^{(k)} \to \bar{t}_i$ $\forall i \in 0:n$, причем $\bar{t}_{i'-1} = \bar{t}_{i'}$ при некотором $i' \in 0:n$; $X_k \to \bar{X}$, где $\bar{X} \in B_{r_0}(X^*)$. Поскольку $|f(X_k, t_{i'-1}^{(k)}) - f(X_k, t_{i'}^{(k)})| \geqslant \varphi^*$, то в пределе получим $0 \geqslant \varphi^*$, что невозможно. Лемма доказана.

Следующий результат является нестандартным обобщением известной в теории аппроксимации теоремы Валле-Пуссена.

Лемма 2. Найдутся $r_1 \in (0, r_0)$, $\delta_1 > 0$ и $\eta_1 > 0$ такие, что как только $X \in B_{r_1}(X^*)$, $\tau \in T_{\delta_1}(\tau^*)$, тотчас

$$\max_{i\in 0:n} |f(Z, t_i)| \geqslant \min_{i\in 0:n} |f(X, t_i)| + \eta_1 \|Z - X\| \quad \forall Z \in B_{r_1}(X^*).$$

В частности, $$\varphi(X^*) \geqslant \gamma(X, \tau) + \eta_1 \|X - X^*\|. \qquad (3)$$

Доказательство. Введем обозначение $S = \{g \in E_n \mid \|g\| = 1\}$. В силу I-III

$$\min_{g\in S} \max_{i\in 0:n} \xi_i^* f_X'(X^*, t_i^*), g) \stackrel{\text{def}}{=} 3\eta_1 > 0 ,$$

где $\xi_i^* = \operatorname{sign} f(X^*, t_i^*)$ (см. доказательство теоремы 6 из § I.1). По непрерывности f и f_X' найдутся $r_1' \in (0, r_0)$ и $\delta_1 > 0$ такие, что из $X \in B_{r_1'}(X^*)$, $\tau \in T_{\delta_1}(\tau^*)$ следует

$$\operatorname{sign} f(X, t_i) = \xi_i^* \quad \forall i \in 0:n ,$$

$$\min_{g\in S} \max_{i\in 0:n} \xi_i^* f_X'(X, t_i), g) \geqslant 2\eta_1 . \qquad (4)$$

Возьмем $X \in B_{r_1'}(X^*)$. Тогда для любых $t_i \in [\alpha, \beta]$ и $\xi_i = \pm 1$ будет

$$\left| \max_{i\in 0:n} \xi_i [f(Z, t_i) - f(X, t_i)] - \max_{i\in 0:n} \xi_i f_X'(X, t_i), Z - X) \right| \leqslant$$

$$\leqslant \|Z - X\| \max_{t\in[\alpha,\beta]} \|f_X'(X + \theta(Z - X), t) - f_X'(X, t)\| , \qquad (5)$$

где $0 < \theta < 1$. При этом можно добиться, чтобы правая часть (5) не превосходила $\eta_1 \|Z - X\|$ равномерно по $X \in B_{r_1'}(X^*)$ как только $\|Z - X\| \leqslant r_1''$, где r_1'' - положительное число меньше r_1':

Положим $r_1 = 1/2\, r_1''$, зафиксируем $X \in B_{r_1}(X^*)$, $\tau \in T_{\delta_1}(\tau^*)$ и проверим, что для любого $Z \in B_{r_1}(X^*)$ будет выполняться требуемое неравенство. Если $Z = X$, то это очевидно. Пусть $Z \neq X$. Поскольку $X \in B_{r_1'}(X^*)$, $\|Z - X\| \leqslant r_1''$ и $\tau \in T_{\delta_1}(\tau^*)$, то, положив $\xi_i = \operatorname{sign} f(X, t_i)$, на основании (5) и (4) получим

$$\max_{i\in 0:n} \xi_i [f(Z, t_i) - f(X, t_i)] \geqslant$$

$$\geqslant \max_{i\in 0:n} \xi_i^* \left(f_X'(X, t_i), \frac{Z - X}{\|Z - X\|} \right) \|Z - X\| - \eta_1 \|Z - X\| \geqslant \eta_1 \|Z - X\| .$$

Отсюда

$$\max_{i\in 0:n} |f(Z, t_i)| \geqslant \min_{i\in 0:n} |f(X, t_i)| + \eta_1 \|Z - X\| ,$$

что и требовалось доказать.

3°. Согласно условию III все определители Δ_i^* отличны от ну-

ля и одного знака. Будем считать для определенности, что все они положительны. Уменьшив при необходимости r_1 и δ_1, указанные в лемме 2, можно добиться того, что как только $X_k \in B_{r_1}(X^*)$ $(k \in 0:n)$, $\tau \in T_{\delta_1}(\tau^*)$, тотчас все определители Δ_i, составленные из столбцов

$$f'_x(X_0,t_0),\dots,f'_x(X_{i-1},t_{i-1}),f'_x(X_{i+1},t_{i+1}),\dots,f'_x(X_n,t_n),$$

будут положительными. Более того,

$$\min_{\substack{i\in 0:n,\ \tau\in T_{\delta_1}(\tau^*),\\ X_0\in B_{r_1}(X^*),\dots,X_n\in B_{r_1}(X^*)}} \Delta_i \stackrel{def}{=} \alpha > 0 . \qquad (6)$$

Это следует из свойств определителей Δ_i^* и непрерывности вектор-функции f'_x.

Лемма 3. Существует такое $r_2 \in (0, r_1)$, что если $\{X,\tau\}$ - экстремальная базисная пара и $X \in B_{r_2}(X^*)$, то $X = X^*$, $\tau = \tau^*$.

Доказательство. Достаточно установить, что $X = X^*$, ибо существует единственный экстремальный базис τ^*, соответствующий X^*. Допустим противное. Тогда найдется последовательность экстремальных базисных пар $\{X_k,\tau_k\}$, у которой $X_k \to X^*$, причем $X_k \neq X^*$ и $X_k \in B_{r_0}(X^*)$ $\forall k = 1,2,\dots$ Согласно (2) $\gamma(X_k,\tau_k) = \varphi(X_k) \geqslant \varphi^*$, поэтому на основании леммы 1 $\tau_k \in C T_\rho$. Перейдя, если нужно, к подпоследовательности, добьемся того, что $\tau_k \to \bar\tau$, где $\bar\tau \in C T_\rho$. По непрерывности $\gamma(X_k,\tau_k) \to \gamma(X^*,\bar\tau)$, $\varphi(X_k) \to \varphi(X^*)$. Значит, $\gamma(X^*,\bar\tau) = \varphi(X^*)$, откуда следует $\bar\tau = \tau^*$.

Теперь имеем $\{X_k,\tau_k\} \to \{X^*,\tau^*\}$. Поскольку при больших k $X_k \in B_{r_1}(X^*)$, $\tau_k \in T_{\delta_1}(\tau^*)$, то на основании (3) получаем $\varphi(X^*) \geqslant \gamma(X_k,\tau_k) = \varphi(X_k)$, что противоречит (1). Лемма доказана.

Лемма 4. Пусть последовательность базисных пар $\{X_k,\tau_k\}$, у которой $X_k \in B_{r_2}(X^*)$ $\forall k = 1,2,\dots$, такова, что

$$\varphi(X_k) - \gamma(X_k,\tau_k) \to 0 . \qquad (7)$$

Тогда $X_k \to X^*$.

Доказательство. Возьмем произвольную сходящуюся последовательность $\{X_{k_j}\}$ и покажем, что ее предел $\bar X$ равен X^*. Учитывая (2), (7) и лемму 1, можно считать, что $\tau_{k_j} \to \bar\tau$, где $\bar\tau \in C T_\rho$. По непрерывности $\gamma(X_{k_j},\tau_{k_j}) \to \gamma(\bar X,\bar\tau)$, причем в силу (7) $\gamma(\bar X,\bar\tau) = \varphi(\bar X)$. Значит, $\{\bar X,\bar\tau\}$ - экстремальная базисная пара и $\bar X \in B_{r_2}(X^*)$. На основании леммы 3 $\bar X = X^*$, что и требовалось доказать.

4°. Обратим теперь внимание на то, что нелинейная система $f(X,t_i^*) - (-1)^i h = 0$, $i \in 0:n$, имеет решение (X^*,h^*), где $|h^*| = \varphi^*$, причем якобиан этой системы в решении

$$\begin{vmatrix} f'_{x_1}(X^*,t_0^*) & \dots & f'_{x_n}(X^*,t_0^*) & -1 \\ f'_{x_1}(X^*,t_1^*) & \dots & f'_{x_n}(X^*,t_1^*) & 1 \\ \dots & \dots & \dots & \dots \\ f'_{x_1}(X^*,t_n^*) & \dots & f'_{x_n}(X^*,t_n^*) & (-1)^{n+1} \end{vmatrix} = (-1)^{n+1} \sum_{i=0}^n \Delta_i^*$$

отличен от нуля. По теореме о неявном отображении (см., например, [115, с. 68-72]) найдутся $r_3 \in (0,r_2)$, $\delta_2 \in (0,\delta_1)$ и единственное непрерывное отображение $\tau \to (X(\tau),h(\tau))$ множества $T_{\delta_2}(\tau^*)$ в E_{n+1} со свойствами:

(i) $X(\tau^*) = X^*$, $h(\tau^*) = h^*$;

(ii) $\|X(\tau) - X^*\|^2 + |h(\tau) - h^*|^2 \leq r_3^2$ $\forall \tau \in T_{\delta_2}(\tau^*)$,

причем можно считать, что $|h(\tau)| \geqslant \varphi^*/2$ $\forall \tau \in T_{\delta_2}(\tau^*)$;

(iii) для всех $\tau \in T_{\delta_2}(\tau^*)$

$$f(X(\tau),t_i) - (-1)^i h(\tau) = 0, \quad i \in 0:n.$$

Нетрудно понять, что пара $\{X(\tau),\tau\}$ при любом $\tau \in T_{\delta_2}(\tau^*)$ базисная и $\gamma(X(\tau),\tau) = |h(\tau)|$.

Лемма 5. Если $|h(\tau_k)| \to \varphi^*$, то $\{X(\tau_k),\tau_k\} \to \{X^*,\tau^*\}$.

Доказательство. В силу непрерывности отображения $\tau \to X(\tau)$ достаточно установить, что $\tau_k \to \tau^*$. Допустим противное. Поскольку все τ_k принадлежат $T_{\delta_2}(\tau^*)$, то найдется подпоследовательность $\{\tau_{k_j}\}$, сходящаяся к базису $\bar\tau \neq \tau^*$, причем $\bar\tau \in T_{\delta_2}(\tau^*)$. В силу непрерывности отображения $\tau \to h(\tau)$ имеем $h(\tau_{k_j}) \to h(\bar\tau)$, откуда с учетом условия леммы получаем $|h(\bar\tau)| = \varphi^*$. Теперь очевидно, что $X(\bar\tau) \neq X^*$, ибо единственным экстремальным базисом, соответствующим X^*, является τ^*. На основании (3) $\varphi(X^*) \geqslant \gamma(X(\bar\tau),\bar\tau) + \eta_1 \|X(\bar\tau) - X^*\|$, что невозможно, так как $\gamma(X(\bar\tau),\bar\tau) = |h(\bar\tau)| = \varphi(X^*)$. Лемма доказана.

Следующий важный результат принадлежит Баррару и Лоубу [126].

Лемма 6. Пусть базисная пара $\{\bar{X},\bar{\tau}\}$ такова, что $\bar{X}\in B_{r_3}(X^*)$, $\bar{\tau}\in T_{\delta_2}(\tau^*)$. Тогда

$$|h(\bar{\tau})| = \sum_{i=0}^{n}\theta_i |f(\bar{X},\bar{t}_i)|,$$

где $\sum_{i=0}^{n}\theta_i = 1$ и $\theta_i \geqslant c > 0$ $\forall i\in 0:n$ равномерно по всем $\{\bar{X},\bar{\tau}\}$ с указанными свойствами.

Доказательство. Имеем

$$f(\bar{X},\bar{t}_i)-(-1)^i h(\bar{\tau}) = f(\bar{X},\bar{t}_i)-f(X(\bar{\tau}),\bar{t}_i) =$$

$$= \sum_{j=1}^{n} f'_{x_j}(X_i,\bar{t}_i)\left(\bar{x}_j - x_j(\bar{\tau})\right), \quad i\in 0:n, \qquad (8)$$

где X_i принадлежат отрезку, соединяющему $\bar{X}$ и $X(\bar{\tau})$. В частности, $X_i\in B_{r_3}(X^*)$ $\forall i\in 0:n$. Рассматривая (8) как линейную систему относительно $h(\bar{\tau})$ и $\bar{x}_j - x_j(\bar{\tau})$, по формуле Крамера получаем

$$h(\bar{\tau}) = \frac{\begin{vmatrix} f(\bar{X},\bar{t}_0) & f'_{x_1}(X_0,\bar{t}_0) & \dots & f'_{x_n}(X_0,\bar{t}_0) \\ f(\bar{X},\bar{t}_1) & f'_{x_1}(X_1,\bar{t}_1) & \dots & f'_{x_n}(X_1,\bar{t}_1) \\ \dots & \dots & \dots & \dots \\ f(\bar{X},\bar{t}_n) & f'_{x_1}(X_n,\bar{t}_n) & \dots & f'_{x_n}(X_n,\bar{t}_n) \end{vmatrix}}{\begin{vmatrix} 1 & f'_{x_1}(X_0,\bar{t}_0) & \dots & f'_{x_n}(X_0,\bar{t}_0) \\ -1 & f'_{x_1}(X_1,\bar{t}_1) & \dots & f'_{x_n}(X_1,\bar{t}_1) \\ \dots & \dots & \dots & \dots \\ (-1)^n & f'_{x_1}(X_n,\bar{t}_n) & \dots & f'_{x_n}(X_n,\bar{t}_n) \end{vmatrix}}$$

Отсюда

$$h(\bar{\tau}) = \frac{\sum_{i=0}^{n}(-1)^i f(\bar{X},\bar{t}_i)\Delta_i}{\sum_{i=0}^{n}\Delta_i} = \sum_{i=0}^{n}(-1)^i f(\bar{X},\bar{t}_i)\,\frac{\Delta_i}{\sum_{i=0}^{n}\Delta_i}.$$

Положим $\theta_i = \Delta_i \,/\, \sum_{i=0}^{n}\Delta_i$. Теперь требуемое следует из (6) и определения базисной пары. Лемма доказана.

5°. Лемма 7. По выбранным r_3 и δ_2 найдется $\varepsilon_0>0$ такое, что как только для базисной пары $\{X,\tau\}$ окажется $X\in B_{r_3}(X^*)$, $\varphi(X)-\gamma(X,\tau)\leqslant\varepsilon_0$, тотчас $\tau\in T_{\delta_2}(\tau^*)$.

Доказательство. Допустим противное. Тогда найдется последовательность базисных пар $\{X_k,\tau_k\}$ со свойствами $X_k\in B_{r_3}(X^*)$, $\varphi(X_k)-\gamma(X_k,\tau_k)\to 0$, $\tau_k\overline{\in} T_{\delta_2}(\tau^*)$. Согласно лемме 4

$X_k\to X^*$, а на основании леммы 1 можно считать, что $\tau_k\to\bar{\tau}$, причем $\bar{\tau}\neq\tau^*$. В пределе получаем $\varphi(X^*)=\gamma(X^*,\bar{\tau})$, что невозможно, ибо единственным экстремальным базисом, соответствующим X^*, является τ^*. Лемма доказана.

Лемма 8. По выбранному ε_0 найдется $\varepsilon_1>0$ такое, что как только $\varphi^*-|h(\tau)|\leqslant\varepsilon_1$, тотчас $\varphi(X(\tau))-|h(\tau)|\leqslant\varepsilon_0$.

Доказательство. Допустим противное. Тогда найдется последовательность базисов $\{\tau_k\}$ из $T_{\delta_2}(\tau^*)$, для которой $|h(\tau_k)|\to\varphi^*$, но

$$\varphi(X(\tau_k))-|h(\tau_k)|>\varepsilon_0. \qquad (9)$$

По лемме 5 $X(\tau_k)\to X^*$, что противоречит (9). Лемма доказана.

6°. Опишем метод последовательных чебышевских интерполяций. Берем начальный базис $\tau_0\in T_{\delta_2}(\tau^*)$. По лемме 7 для этого достаточно, чтобы нашлась базисная пара $\{X_{-1},\tau_0\}$, у которой $X_{-1}\in B_{r_3}(X^*)$ и $\varphi(X_{-1})-\gamma(X_{-1},\tau_0)\leqslant\varepsilon_0$. Система $f(X,t_i^{(0)})=(-1)^i h$, $i\in 0:n$, имеет единственное решение $X_0=X(\tau_0)$, $h_0=h(\tau_0)$ такое, что $\|X_0-X^*\|^2+|h_0-h^*|^2\leqslant r_3^2$. При этом на основании (3) $\varphi^*\geqslant|h_0|$. Дополнительно предполагаем, что $\varphi^*-|h_0|\leqslant\varepsilon_1$. Это условие является естественным в силу непрерывности отображения $\tau\to h(\tau)$ и равенства $\varphi^*=|h(\tau^*)|$.

Допустим, что имеется k-е приближение $\tau_k\in T_{\delta_2}(\tau^*)$, $X_k=X(\tau_k)$, $h_k=h(\tau_k)$, $\varphi_k=\varphi(X_k)$ и выполняются неравенства

$$0\leqslant\varphi^*-|h_k|\leqslant\varepsilon_1. \qquad (10)$$

Если $\varphi_k=|h_k|$ $\left(=\gamma(X_k,\tau_k)\right)$, то на основании леммы 3 $X_k=X^*$, $\tau_k=\tau^*$. В этом случае вычисления заканчиваются. Пусть $\varphi_k>|h_k|$. Тогда выбираем новый базис τ_{k+1} такой, что $\{X_k,\tau_{k+1}\}$ – базисная пара и

$$m_k \overset{\text{def}}{=} \gamma(X_k,\tau_{k+1})\geqslant|h_k|, \qquad (11)$$

$$M_k \overset{\text{def}}{=} \max_{i\in 0:n}|f(X_k,t_i^{(k+1)})|>|h_k|. \qquad (12)$$

Прежде всего проверим включение $\tau_{k+1}\in T_{\delta_2}(\tau^*)$. Согласно (10) и лемме 8 $\varphi(X_k)-|h_k|\leqslant\varepsilon_0$, так что, учитывая (11), получаем $\varphi(X_k)-\gamma(X_k,\tau_{k+1})\leqslant\varepsilon_0$. На основании леммы 7 $\tau_{k+1}\in T_{\delta_2}(\tau^*)$, что и требовалось доказать.

Осуществляя чебышевскую интерполяцию на новом базисе τ_{k+1}

(т.е. решая систему $f(X,t_i^{(k+1)})=(-1)^i h$, $i\in 0:n$), находим $X_{k+1}=X(\tau_{k+1})$, $h_{k+1}=h(\tau_{k+1})$, $\varphi_{k+1}=\varphi(X_{k+1})$. По лемме 2 $\varphi^*\geqslant|h_{k+1}|$ и $|h_{k+1}|\geqslant\min\limits_{i\in 0:n}|f(X_k,t_i^{(k+1)})|=m_k\geqslant|h_k|$. Отсюда, в частности, следует, что $\varphi^*-|h_{k+1}|\leqslant\varepsilon_1$.

Теперь итерации можно продолжить. В результате получаем последовательности $\{\tau_k\}$, $\{X_k\}$, $\{h_k\}$, $\{\varphi_k\}$, $\{m_k\}$ и $\{M_k\}$, причем $\tau_k\in T_{\delta_2}$, $X_k\in B_{r_3}(X^*)$, $|h_{k+1}|\geqslant|h_k|$ и выполняются неравенства (11) и (12).

7°. Если последовательность базисных пар $\{X_k,\tau_k\}$ конечна, то последняя пара по построению совпадает с $\{X^*,\tau^*\}$. Допустим, что последовательность $\{X_k,\tau_k\}$ бесконечна. Тогда справедлива следующая теорема.

Теорема. Если

$$\varliminf_{k\to\infty}\frac{M_k-|h_k|}{\varphi_k-|h_k|}>0, \tag{13}$$

то $\{X_k,\tau_k\}\to\{X^*,\tau^*\}$. Более того, для всех $k=0,1,2,\ldots$

$$\|X_k-X^*\|\leqslant c\theta^k, \tag{14}$$

где $\theta\in(0,1)$ и константа $c>0$ не зависит от k.

Доказательство проводится по той же схеме, что и в линейном случае (ср., например, [39, с. 56-59]). Прежде всего заметим, что в силу (12) и (13) существует такая константа $l\in(0,1]$, что при всех $k=0,1,2,\ldots$

$$M_k-|h_k|\geqslant l(\varphi_k-|h_k|). \tag{15}$$

Применим лемму 6 к базисной паре $\{X_k,\tau_{k+1}\}$. Это даст

$$|h_{k+1}|=\sum_{i=0}^n\theta_i|f(X_k,t_i^{(k+1)})|.$$

Если обозначить через i' индекс, на котором $|f(X_k,t_{i'}^{(k+1)})|=M_k$, то в силу (11)

$$|h_{k+1}|=\sum_{\substack{i=0\\ i\neq i'}}^n\theta_i|f(X_k,t_i^{(k+1)})|+\theta_{i'}M_k\geqslant$$

$$\geqslant(1-\theta_{i'})|h_k|+\theta_{i'}M_k=|h_k|+\theta_{i'}(M_k-|h_k|).$$

Учитывая (15), (1) и лемму 6, получаем

$$|h_{k+1}|-|h_k|\geqslant cl(\varphi_k-|h_k|)\geqslant cl(\varphi^*-|h_k|). \tag{16}$$

Введем обозначение $\theta=1-cl$. Очевидно, что $\theta\in(0,1)$. Теперь неравенство (16) можно переписать в виде

$$\varphi^*-|h_{k+1}|\leqslant\theta(\varphi^*-|h_k|),$$

откуда $\varphi^*-|h_{k+1}|\leqslant\theta^{k+1}(\varphi^*-|h_0|)\overset{def}{=}c_1\theta^{k+1}$ (формально последнее неравенство справедливо и при $k=-1$). Осталось положить $C=c_1/\eta_1$ и сослаться на (3). Теорема доказана.

З а м е ч а н и е. Условие (13) заведомо выполняется, когда $M_k=\varphi_k$ при всех $k=0,1,2,\ldots$, т.е. когда в очередной базис τ_{k+1} включается точка из $R(X_k)$.

§ 1.3. Выравнивание максимумов

На каждом шаге алгоритма Ремеза приходится осуществлять чебышевскую интерполяцию, т.е. решать нелинейную систему вида $f(X,t_i)=(-1)^i h$, $i\in 0:n$. Эта иррациональная процедура затрудняет непосредственное использование алгоритма Ремеза для решения нелинейных задач аппроксимации. Метод выравнивания максимумов, к рассмотрению которого мы сейчас переходим, лишен указанного недостатка. В полном виде он был предложен в [74] (см. также [38]). В случае $f(X,t)=\sum_{i=1}^n x_i t^{i-1}-u_0(t)$ этот метод является одним из вариантов метода последовательных чебышевских интерполяций [147].

1°. Рассмотрим задачу

$$\max_{t\in[\alpha,\beta]}|f(X,t)|\to\min_{X\in E_n}.$$

Как следует из § 1.1, при наличии полного альтернанса и достаточно хорошего начального приближения решение этой задачи можно свести к решению нелинейной системы

$$\begin{aligned} &f'_t(X,t_i)=0, \quad i\in 1:r,\\ &f(X,t_i)-\xi_i h=0, \quad i\in 1:n+1. \end{aligned} \tag{1}$$

Предполагается, что числа $t_{r+1},\ldots,t_{n+1}$ и $\xi_i=\pm 1$ уже известны и фиксированы. Неизвестными являются векторы $U=(t_1,\ldots,t_r)$ и $V=(x_1,\ldots,x_n,h)$.

Перепишем систему (1) в векторном виде:

$$F(U,V)=\mathbf{0}, \qquad G(U,V)=\mathbf{0}. \tag{2}$$

Метод выравнивания максимумов заключается в следующем. Пусть имеется k-е приближение $\{U_k, V_k\}$. Фиксируем V_k, уточняем U с помощью одного шага по методу Ньютона в системе $F(U,V_k)=\mathbf{0}$:

$$U_{k+1}=U_k-F_U^{-1}(U_k,V_k)F(U_k,V_k).$$

Это сводится к пересчету по формулам

$$t_i^{(k+1)}=t_i^{(k)}-\frac{f_t'(X_k,t_i^{(k)})}{f_{tt}''(X_k,t_i^{(k)})}, \qquad i\in 1:r.$$

Далее, фиксируя U_{k+1}, уточняем V с помощью одного шага по методу Ньютона в системе $G(U_{k+1},V)=\mathbf{0}$:

$$V_{k+1}=V_k-G_V^{-1}(U_{k+1},V_k)G(U_{k+1},V_k).$$

Затем процесс повторяется. В [74] доказана следующая теорема.

Теорема 1. Пусть

(i) система (2) имеет решение $Z^*=\{U^*,V^*\}$ и отображения F и G непрерывно дифференцируемы в некоторой окрестности точки Z^*;

(ii) существуют $F_U^{-1}(Z^*)$ и $G_V^{-1}(Z^*)$.

Если $Z_0=\{U_0,V_0\}$ – достаточно близкое к Z^* начальное приближение, то последовательность $Z_k=\{U_k,V_k\}$ сходится к Z^* со сверхлинейной скоростью:

$$\|Z_k-Z^*\|\le \text{const}\, q_1q_2\dots q_k,$$

где $q_i\to 0$ при $i\to\infty$.

Ключевым при доказательстве этой теоремы является легко проверяемое соотношение $G_U(Z^*)=0$.

Рассмотрим теперь произвольную систему (2), обладающую тем свойством, что $\|G_U(U,V)\|$ мала в окрестности начального приближения $\{U_0,V_0\}$. Далее доказывается сходимость метода выравнивания максимумов без предположения о существовании решения. Предполагается, что $\|U_1-U_0\|$ и $\|V_1-V_0\|$ достаточно малы (как в теореме Л.В.Канторовича о методе Ньютона [55, 56]). Положим

$$Z=\{U,V\},\ \|Z\|=\max\{\|U\|,\|V\|\},\ \xi_k=U_{k+1}-U_k,\ \eta_k=V_{k+1}-V_k.$$

Теорема 2. Пусть в шаре $\Omega=\{Z\mid \|Z-Z_0\|\le r\}$ вектор-функции

$F(Z)$ и $G(Z)$ непрерывно дифференцируемы, существуют обратные матрицы $F_U^{-1}(Z)$, $G_V^{-1}(Z)$ и

$$\|G_U(Z)\|\le\varepsilon\le 1, \quad \|F_V(Z)\|\le C,$$

$$\|F_U^{-1}(Z)\|\le A, \quad \|G_V^{-1}(Z)\|\le A,$$

$$\|F_U(Z_1)-F_U(Z_2)\|\le 2L\|Z_1-Z_2\|, \quad \|G_V(Z_1)-G_V(Z_2)\|\le 2L\|Z_1-Z_2\|$$

для любых $Z_1,Z_2\in\Omega$. Пусть также

$$\|\xi_0\|\le\delta, \quad \|\eta_0\|\le\varepsilon\delta, \tag{3}$$

$$q\overset{\text{def}}{=}\max\{A(L\delta+C\varepsilon),\ A^2(L\delta+C\varepsilon)+AL\delta\}<1, \quad \delta/(1-q)\le r.$$

Тогда в Ω существует решение системы (2) и последовательность $Z_k=\{U_k,V_k\}$ сходится к некоторому решению $Z^*=\{U^*,V^*\}$, причем

$$\|U_k-U^*\|\le\frac{\delta}{1-q}q^k, \quad \|V_k-V^*\|\le\frac{\varepsilon\delta}{1-q}q^k, \quad k=0,1,2,\dots \tag{4}$$

Если при этом $G_U(Z^*)=0$, то скорость сходимости сверхлинейная.

Доказательство. Покажем по индукции, что

$$\|\xi_k\|\le\delta q^k, \quad \|\eta_k\|\le\varepsilon\delta q^k, \quad k=0,1,2,\dots$$

Отсюда следуют и сходимость и неравенства (4). Для $k=0$ (5) выполнено в силу (3), и очевидно $\{U_1,V_0\}$ и $\{U_1,V_1\}$ принадлежат Ω. Сделаем индукционный переход от k к $k+1$. Имеем

$$\|\xi_{k+1}\|\le A\|F(U_{k+1},V_{k+1})\|,$$

$$F(U_{k+1},V_{k+1})=F(U_{k+1},V_{k+1})-F(U_k,V_k)-F_U(U_k,V_k)\xi_k=$$

$$=\int_0^1\{[F_U(Y(t))-F_U(U_k,V_k)]\xi_k+F_V(Y(t))\eta_k\}\,dt,$$

где $Y(t)=(U_k+t\xi_k,V_k+t\eta_k)$. Отсюда

$$\|\xi_{k+1}\|\le A(L\delta\|\xi_k\|+C\|\eta_k\|)\le A(L\delta+C\varepsilon)\delta q^k\le\delta q^{k+1}.$$

Кроме того, $\{U_{k+2},V_{k+1}\}\in\Omega$. Действительно,

$$\|U_{k+2}-U_0\|\le\|\xi_0\|+\dots+\|\xi_{k+1}\|\le\delta+\dots+\delta q^{k+1}<\frac{\delta}{1-q}\le r.$$

Далее, $\|\eta_{k+1}\|\le A\|G(U_{k+2},V_{k+1})\|$,

$$G(U_{k+2},V_{k+1})=\int_0^1\{G_U(Z(t))\xi_{k+1}+[G_V(Z(t))-G_V(U_{k+1},V_k)]\eta_k\}\,dt,$$

где $z(t)=(U_{k+1}+t\xi_{k+1}, V_k+t\eta_k)$. Отсюда

$$\|\eta_{k+1}\| \le A(\varepsilon\|\xi_{k+1}\|+L\delta\|\eta_k\|) \le A(\varepsilon A(L\delta+C\varepsilon)+\varepsilon L\delta)\delta q^k \le \varepsilon\delta q^{k+1},$$

что и требовалось доказать. При этом $\{U_{k+2}, V_{k+2}\}\in\Omega$.

Из (5) следует, что $\{z_k\}$ сходится к некоторому решению z^* системы (2) и $\|U_k-U^*\|\le r_k$, $\|V_k-V^*\|\le\varepsilon r_k$, $k=0,1,2,\ldots$, где $r_k=\delta q^k/(1-q)$. Покажем, что в случае $G_U(z^*)=0$ скорость сходимости сверхлинейная. Положим

$$\varepsilon_k=\max_{\|z-z^*\|\le r_k}\|G_U(z)\|, \quad \lambda_k=\max\{Lr_k, A\varepsilon_k\},$$

$$q_k=\lambda_{k-1}+(AC+1)\lambda_{k-2}, \quad k=1,2,\ldots \quad (\lambda_{-1}=\varepsilon).$$

Так как $\lambda_k\to 0$, то $q_k\to 0$ при $k\to\infty$. Для $\xi_k^*=U_k-U^*$, $\eta_k^*=V_k-V^*$ докажем неравенства

$$\|\xi_k^*\|\le r_0 q_1\ldots q_k, \quad \|\eta_k^*\|\le\lambda_{k-1}r_0 q_1\ldots q_k.$$

При $k=0$ они выполнены. Сделаем индукционный переход. Так как

$$\xi_{k+1}^*=\xi_k^*-F_U^{-1}(z_k)[F(z_k)-F(z^*)]=$$

$$=-F_U^{-1}(z_k)\int_0^1\{[F_U(Y^*(t))-F_U(z_k)]\xi_k^*+F_V(Y^*(t))\eta_k^*\}\,dt,$$

где $Y^*(t)=(U^*+t\xi_k^*, V^*+t\eta_k^*)$, то

$$\|\xi_{k+1}^*\|\le Lr_k\|\xi_k^*\|+AC\|\eta_k^*\|\le(\lambda_k+AC\lambda_{k-1})r_0 q_1\ldots q_k.$$

Далее имеем

$$\eta_{k+1}^*=-G_V^{-1}(U_{k+1},V_k)\int_0^1\{G_U(z^*(t))\xi_{k+1}^*+[G_V(z^*(t))-G_V(U_{k+1},V_k)]\eta_k^*\}dt,$$

где $z^*(t)=(U^*+t\xi_{k+1}^*, V^*+t\eta_k^*)$. Отсюда

$$\|\eta_{k+1}^*\|\le\lambda_k\|\xi_{k+1}^*\|+\lambda_k\|\eta_k^*\|\le$$

$$\le\lambda_k(\lambda_k+AC\lambda_{k-1}+\lambda_{k-1})r_0 q_1\ldots q_k=\lambda_k r_0 q_1\ldots q_k q_{k+1}.$$

Итак, скорость сходимости сверхлинейная. Теорема доказана полностью.

Для установления единственности решения следует наложить более жесткое условие на q.

Теорема 3. Пусть в условиях теоремы 2 число $q<1/2$. Тогда решение системы (2) единственно в области

$$\|U-U_0\|\le\frac{\delta}{1-q}, \qquad \|V-V_0\|\le\frac{\varepsilon\delta}{1-q}. \tag{6}$$

Доказательство. Система (2) имеет в области (6) решение, что следует из теоремы 2. Пусть $\tilde z=\{\tilde U,\tilde V\}$ – любое решение системы (2) такое, что $\|\tilde U-U_0\|\le\delta_1$, $\|\tilde V-V_0\|\le\varepsilon\delta_1$, где $\delta_1=\delta/(1-q)$. Тогда можно получить оценки, аналогичные (5): $\|\tilde U-U_k\|\le\delta_1 q_1^k$, $\|\tilde V-V_k\|\le\varepsilon\delta_1 q_1^k$, $k=0,1,2,\ldots$, где $q_1=\max\{A(L\delta_1+C\varepsilon), A^2(L\delta_1+C\varepsilon)+AL\delta_1\}$. Для этого q_1 имеем $q_1\le q/(1-q)<1$. Значит, $\tilde U=\lim U_k=U^*$, $\tilde V=\lim V_k=V^*$. Единственность доказана.

2°. Рассмотрим модифицированный метод:

$$U_{k+1}=U_k-F_U^{-1}(z_0)F(U_k,V_k),$$

$$V_{k+1}=V_k-G_V^{-1}(z_0)G(U_{k+1},V_k).$$

Теорема 4. Допустим, что

1) в шаре $\Omega=\{z\mid\|z-z_0\|\le r\}$ вектор-функции $F(z)$, $G(z)$ непрерывно дифференцируемы и матрицы $F_U(z)$, $G_V(z)$ удовлетворяют условию Липшица с константой L;

2) существуют обратные матрицы $F_U^{-1}(z_0)$, $G_V^{-1}(z_0)$, причем $\|F_U^{-1}(z_0)\|\le A$, $\|G_V^{-1}(z_0)\|\le A$, $A\ge 1$;

3) $\|G_U(z)\|\le\varepsilon\le 1$, $\|F_V(z)\|\le C$ $\forall z\in\Omega$;

4) $\|U_1-U_0\|\le\delta$, $\|V_1-V_0\|\le\varepsilon\delta$;

5) $\gamma\overset{\text{def}}{=}A^2(L\delta+C\varepsilon)+AL\delta\le 1/4$;

6) справедливо неравенство $\delta/(1-q)\le r$, где $q=(\beta-\sqrt{\beta^2-4\gamma})/2$, $\beta=1+A^2C\varepsilon$.

Тогда в Ω существует решение $z^*=\{U^*,V^*\}$ системы (2), к которому сходятся последовательные приближения $z_k=\{U_k,V_k\}$, причем

$$\|U_k-U^*\|\le\frac{\delta}{1-q}q^k, \qquad \|V_k-V^*\|\le\frac{\varepsilon\delta}{1-q}q^k, \quad k=0,1,2,\ldots$$

Доказательство данной теоремы аналогично доказательству теоремы 2. Необходимо лишь учесть, что

$$q=\frac{\beta}{2}\left[1-\sqrt{1-\frac{4\gamma}{\beta^2}}\right]\le\frac{\beta}{2}\left[1-\left(1-\frac{4\gamma}{\beta^2}\right)\right]=\frac{2\gamma}{\beta}<\frac{1}{2}$$

и что q является меньшим корнем уравнения

$$A^2\left(L\frac{\delta}{1-t}+C\varepsilon\right)+AL\frac{\delta}{1-t}=t,$$

или $t^2-\beta t+\gamma=0$. При этом $A[L\delta/(1-q)+C\varepsilon]\le q$.

З а м е ч а н и е. Метод выравнивания максимумов и его модификация применимы и в случае, когда число точек альтернанса меньше $n+1$, т.е. применимы к решению системы

$$f'_t(X,t_i)=0,\quad i\in 1:r,$$

$$f(X,t_i)-\xi_i h=0,\quad i\in 1:m,$$

где $r\le m<n+1$, $X\in E_n$. Матрица $G_V(Z)$ имеет размеры $m\times(n+1)$, поэтому вместо обычной обратной можно использовать псевдообратную матрицу $B(Z)=G_V^*(Z)[G_V(Z)G_V^*(Z)]^{-1}$, где $*$ означает транспонирование. Теоремы 2 и 4 переносятся на этот случай, только $G_V^{-1}(Z)$ необходимо всюду заменить на $B(Z)$.

3°. Если несколько усложнить метод, то можно получить квадратичную скорость сходимости.

Рассмотрим систему (2), в которой $G_U(U^*,V^*)=0$.

Пусть имеется k-е приближение $Z_k=\{U_k,V_k\}$, $(k+1)$-е приближение находим из системы

$$G(U_k,V_k)+G_V(U_k,V_k)(V_{k+1}-V_k)=\mathbf{0},$$

$$F(U_k,V_k)+F_U(U_k,V_k)(U_{k+1}-U_k)+F_V(U_k,V_k)(V_{k+1}-V_k)=\mathbf{0}. \qquad (7)$$

Из первого уравнения определяем V_{k+1}, а затем из второго U_{k+1}. Этот метод называется ускоренным выравниванием максимумов.

У.А.Рыскелдиевым установлен следующий результат.

Теорема 5. Допустим, что система (2) имеет решение $Z^*=\{U^*,V^*\}$, причем:

(i) вектор-функции F, G дважды непрерывно дифференцируемы в некоторой окрестности точки Z^* и $G_U(Z^*)=0$;

(ii) существуют $F_U^{-1}(Z^*)$ и $G_V^{-1}(Z^*)$.

Если Z_0 – достаточно близкое к Z^* начальное приближение, то последовательность $\{Z_k\}$ сходится к Z^* с квадратичной скоростью: $\|Z_{k+1}-Z^*\|\le \mathrm{const}\,\|Z_k-Z^*\|^2$, $k=0,1,2,\ldots$

Более полным является следующее утверждение, доказанное А.В.Соломенниковым.

Теорема 6. Пусть в шаре $\Omega=\{Z\mid\|Z-Z_0\|\le r\}$ выполнены условия:

1) вектор-функции $F(Z),G(Z)$ непрерывно дифференцируемы и существуют обратные матрицы $F_U^{-1}(Z)$, $G_V^{-1}(Z)$;

2) $\|G_U(Z)\|\le\varepsilon\le 1$, $\|F_V(Z)\|\le C$, $\|F_U^{-1}(Z)\|\le A$, $\|G_V^{-1}(Z)\|\le A$;

3) матрицы $F_U(Z)$, $F_V(Z)$ и $G_V(Z)$ удовлетворяют условию Липшица с константой $2L$.

Пусть, кроме того:

4) $\|U_1-U_0\|\le\delta$, $\|V_1-V_0\|\le\delta$;

5) $q\overset{\text{def}}{=}\max\{A(\varepsilon+L\delta),A^2C(\varepsilon+L\delta)+2AL\delta\}<1$, $\delta/(1-q)\le r$.

Тогда в Ω существует решение системы (2) и последовательность $Z_k=\{U_k,V_k\}$, построенная согласно (7), сходится к некоторому решению $Z^*=\{U^*,V^*\}$. При этом $\|Z_k-Z^*\|\le\frac{\delta}{1-q}q^k$, $k=0,1,2,\ldots$

Если дополнительно $G_U(Z^*)=0$ и

$$\|G_U(Z_1)-G_U(Z_2)\|\le 2M\|Z_1-Z_2\|\quad\forall Z_1,Z_2\in\Omega,$$

то скорость сходимости квадратичная:

$$\|Z_{k+1}-Z^*\|\le\mathrm{const}\,\|Z_k-Z^*\|^2,\quad k=0,1,2,\ldots$$

Если в условии 5) $q<\min\left\{\frac{1}{2-A\varepsilon},\frac{1}{2-A^2C\varepsilon}\right\}$, то в области $\{Z\mid\|Z-Z_0\|\le\delta/(1-q)\}$ решение единственно.

4°. Применим ускоренное выравнивание максимумов к решению задачи наилучшей полиномиальной аппроксимации. Положим $P(X,t)=\sum_{i=1}^{n}x_it^{i-1}$, $f(X,t)=P(X,t)-u_0(t)$. Точки альтернанса будем считать упорядоченными по возрастанию: $\alpha\le t_1<t_2<\ldots<t_{n+1}\le\beta$. Через $I\subset 1:n+1$ обозначим индексы неизвестных точек альтернанса. Тогда система (1) примет вид

$$P'_t(X,t_i)-u'_0(t_i)=0,\quad i\in I,$$

$$(-1)^{i-1}h+P(X,t_i)=u_0(t_i),\quad i\in 1:n+1.$$

Нетрудно проверить, что расчетные формулы (7) в данном случае следующие:

$$(-1)^{i-1}h^{(k+1)}+P(X_{k+1},t_i^{(k)})=u_0(t_i^{(k)}),\quad i\in 1:n+1,$$

$$t_i^{(k+1)}=t_i^{(k)}-\frac{P'_t(X_{k+1},t_i^{(k)})-u'_0(t_i^{(k)})}{P''_{tt}(X_k,t_i^{(k)})-u''_0(t_i^{(k)})},\quad i\in I. \qquad (8)$$

Таким образом, вычисление $h^{(k+1)}$ и X_{k+1} сводится к чебышев-

ской интерполяции, а $t_i^{(k+1)}$ явно выражается через $t_i^{(k)}$, X_k и X_{k+1}.

Метод (8) является одним из вариантов метода последовательных чебышевских интерполяций. Согласно теоремам 5 и 6 он обладает квадратичной скоростью сходимости.

§ 1.4. Выравнивание максимумов. Теорема о квадратичной скорости сходимости

1°. Рассмотрим безмодульную минимаксную задачу

$$\varphi(X) \stackrel{\text{def}}{=} \max_{t \in E_p} f(X,t) \to \min_{X \in E_n} .$$

В тех случаях, когда точка локального минимума X^* функции φ имеет ровно $n+1$ точек альтернанса $t_1^*, t_2^*, \ldots, t_{n+1}^*$, систему, связывающую X^* и $\{t_i^*\}$, можно записать в виде

$$f'_t(X,t_i) = 0, \quad i \in 1:n+1, \tag{1}$$

$$f(X,t_i) - f(X,t_{n+1}) = 0, \quad i \in 1:n. \tag{2}$$

Систему (1), (2) будем решать методом выравнивания максимумов, выполняя итерации поочередно по X из (2) и по $\{t_i\}$ из (1).

В предположениях, которые формулируются далее, доказана квадратичная скорость сходимости этого метода.

2°. Для большей общности рассмотрим систему

$$\begin{aligned} f'_t(X,t_i) &= 0, \quad i \in 1:m, \\ f(X,t_i) - f(X,t_m) &= 0, \quad i \in 1:m-1, \end{aligned} \tag{3}$$

где $2 \leq m \leq n+1$.

Введем обозначения $U = (t_1, t_2, \ldots, t_m)$, $Z = \{U, X\}$,

$F_i(U,X) = f'_t(X,t_i)$, $G_i(U,X) = f(X,t_i) - f(X,t_m)$,

$F = (F_1, F_2, \ldots, F_m)$, $G = (G_1, G_2, \ldots, G_{m-1})$.

Систему (3) перепишем в векторном виде:

$$\begin{aligned} F(U,X) &= \mathbf{0}, \\ G(U,X) &= \mathbf{0}. \end{aligned} \tag{4}$$

Заметим, что в силу специфики системы (4) из того, что $F(U,X)=0$, следует $G_U(U,X)=0$ и из $F(U,X) + F_U(U,X)\Delta U = 0$ вытекает $G_U(U,X) + G_{UU}(U,X)\Delta U = 0$.

Предполагаем существование производных $G_z(z)$, $G_{UU}(z)$, $F_z(z)$ в области Ω, которая определяется ниже, и положительного числа L такого, что

$$\| G_z(z) - G_z(z') \| \leq L \| z - z' \|,$$

$$\| G_{UU}(U,X) - G_{UU}(U',X) \| \leq L \| U - U' \|,$$

$$\| F(U,X) - F(U,X') \| \leq L \| X - X' \|,$$

$$\| F_U(U,X) - F_U(U',X) \| \leq L \| U - U' \|,$$

$$\| F_X(U,X) - F_X(U,X') \| \leq L \| X - X' \|$$

для любых точек z, z', $\{U, X'\}$, $\{U', X\}$, принадлежащих Ω.

При доказательстве теоремы о сходимости часто будет использоваться следующий хорошо известный факт (см., например, [93]). Пусть $\mathcal{X}$ и $\mathcal{Y}$ – банаховы пространства и F – отображение из $\mathcal{X}$ в $\mathcal{Y}$, имеющее производную F_X, удовлетворяющую условию Липшица с константой L, т.е.

$$\| F_X(X) - F_X(X') \| \leq L \| X - X' \| \quad \forall X, X' \in \mathcal{X}.$$

Тогда

$$\| F(X') - F(X) - F_X(X)(X' - X) \| \leq \frac{L}{2} \| X' - X \|^2. \tag{5}$$

Введем обозначения $B(z) = G_X^*(z) [G_X(z) G_X^*(z)]^{-1}$, $z_{ik} = \{U_k, X_i\}$. Заметим, что если $m = n+1$, то $B(z) = G_X^{-1}(z)$.

Пусть z_{00} – начальное приближение. Рассмотрим следующий вариант метода выравнивания максимумов:

$$\begin{aligned} X_{i+1} &= X_i - B(z_{ii}) G(z_{ii}), \\ U_{i+1} &= U_i - F_U^{-1}(z_{i+1,i}) F(z_{i+1,i}). \end{aligned} \tag{6}$$

Теорема. Пусть

1) в шаре $\Omega = \{z \mid \| z - z_{00} \| \leq \delta\}$ существуют операторы $B(z)$, $F_U^{-1}(z)$ и $\max \{ \| B(z) \|, \| F_U^{-1}(z) \| \} \leq A$;

2) $\max\{\|F(Z_{00})\|,\ \|G_U(Z_{00})\|\} \leqslant \frac{\mu L}{2}\gamma_0$, где $\mu = \frac{\sqrt{17}-3}{2}$, $\gamma_0 = A\|G(Z_{00})\|$;

3) $q = \lambda c(1+c)^2\gamma_0 < 1$, где $\lambda = \frac{4+\mu}{4(2+\mu)}$, $c = \frac{AL}{2}(2+\mu)$;

4) $\delta \geqslant (1+c)\gamma_0 \sum_{i=0}^{\infty} q^{2^i-1}$.

Тогда в Ω существует решение системы (4), последовательные приближения (6) сходятся к решению этой системы, т.е.

$$Z_{kk} \underset{k\to\infty}{\to} Z^*, \quad F(Z^*) = 0, \quad G(Z^*) = 0,$$

и справедлива оценка скорости сходимости

$$\|Z_{kk} - Z^*\| \leqslant \frac{q^{2^k-1}}{1-q^{2^k}}(1+c)\gamma_0, \quad k \geqslant 0.$$

З а м е ч а н и е. Условие 2) нестандартно, но вполне естественно для выравнивания максимумов. Оно, в частности, ставит в зависимость норму левой части первого уравнения системы (4) от нормы левой части второго уравнения этой системы в начальной точке.

Д о к а з а т е л ь с т в о. Введем последовательности

$$\beta_i = c\gamma_i, \quad \alpha_i = \gamma_i + \beta_i, \quad \gamma_{i+1} = \lambda c \alpha_i^2, \quad i \geqslant 0,$$

где γ_0 - из условий теоремы.

Легко проверить, что

$$\sum_{i=0}^{\infty} \alpha_i = (1+c)\gamma_0 \sum_{i=0}^{\infty} q^{2^i-1} \tag{7}$$

и

$$\mu\lambda = 1/4. \tag{8}$$

Учитывая (8), получаем

$$\frac{\mu\gamma_{i+1}}{\beta_i^2} = \mu\lambda c\left(1+\frac{\gamma_i}{\beta_i}\right)^2 = \frac{1}{4}\left(\sqrt{c}+\frac{1}{\sqrt{c}}\right)^2, \quad i \geqslant 0,$$

откуда

$$\beta_i^2 \leqslant \mu\gamma_{i+1}, \quad i \geqslant 0. \tag{9}$$

Покажем теперь, что последовательные приближения могут быть построены, т.е. все они принадлежат Ω.

Введем обозначения $\eta_i = X_{i+1} - X_i$, $\xi_i = U_{i+1} - U_i$, $i \geqslant 0$. Очевидно, что $\|\eta_0\| \leqslant \gamma_0$ и $Z_{10} \in \Omega$. Убедимся в том, что $\|\xi_0\| \leqslant \beta_0$, и, следовательно, $Z_{11} \in \Omega$. Так как $\xi_0 = -F_U^{-1}(Z_{10})F(Z_{10})$, то достаточно оценить $\|F(Z_{10})\|$. Имеем

$$\|F(Z_{10})\| \leqslant \|F(Z_{00})\| + \|F(Z_{10}) - F(Z_{00})\| \leqslant \|F(Z_{00})\| + L\|\eta_0\|,$$

откуда, учитывая предположение 2) теоремы, получаем

$$\|\xi_0\| \leqslant A\left(\frac{\mu L}{2}\gamma_0 + L\gamma_0\right) = (2+\mu)\frac{AL}{2}\gamma_0 = \beta_0.$$

Значит, $Z_{11} \in \Omega$.

Для того чтобы иметь базу индукции, необходимо установить дополнительно, что $\|\eta_1\| \leqslant \gamma_1$, $\|\xi_1\| \leqslant \beta_1$. Применяя неравенство (5) к функции G в точках Z_{11} и Z_{00}, получаем

$$\|G(Z_{11}) - G(Z_{00}) - G_X(Z_{00})\eta_0 - G_U(Z_{00})\xi_0\| \leqslant \frac{L}{2}\left(\|\eta_0\| + \|\xi_0\|\right)^2,$$

откуда следует

$$\|G(Z_{11})\| \leqslant \|G(Z_{00}) + G_X(Z_{00})\eta_0 + G_U(Z_{00})\xi_0\| + \frac{L}{2}\left(\|\eta_0\| + \|\xi_0\|\right)^2.$$

Однако (см. (6))

$$G(Z_{ii}) + G_X(Z_{ii})\eta_i = 0, \quad i \geqslant 0. \tag{10}$$

Поэтому согласно предположению 2) теоремы

$$\|G(Z_{11})\| \leqslant \|G_U(Z_{00})\|\|\xi_0\| + \frac{L}{2}\alpha_0^2 \leqslant \frac{L}{2}(\alpha_0^2 + \mu\gamma_0\beta_0).$$

Наконец, так как для произвольных положительных чисел γ и β должно быть

$$\gamma\beta \leqslant \frac{1}{4}(\gamma+\beta)^2, \tag{11}$$

то окончательно получаем

$$\|G(Z_{11})\| \leqslant \frac{L}{2}\left(1+\frac{\mu}{4}\right)\alpha_0^2, \quad \|\eta_1\| \leqslant \gamma_1 \text{ и } Z_{21} \in \Omega.$$

Применим теперь неравенство (5) к функции F в точках Z_{11} и Z_{10}. Имеем

$$\|F(Z_{11}) - F(Z_{10}) - F_U(Z_{10})\xi_0\| \leqslant \frac{L}{2}\|\xi_0\|^2.$$

Далее,

$$\|F(z_{21})\| \le \|F(z_{11})\| + \|F(z_{21}) - F(z_{11})\| \le$$

$$\le \|F(z_{10}) + F_U(z_{10})\xi_0\| + \frac{L}{2}\|\xi_0\|^2 + L\|\eta_1\| .$$

На основании (6)

$$F(z_{i+1,i}) + F_U(z_{i+1,i})\xi_i = 0 , \quad i \ge 0 . \tag{12}$$

Поэтому, учитывая (9), получаем

$$\|F(z_{21})\| \le \frac{L}{2}(2+\mu)\gamma_1 , \quad \|\xi_1\| \le \beta_1 \quad \text{и} \quad z_{22} \in \Omega .$$

Предположим теперь, что точки z_{ii}, $i \in 0:k$, $z_{i+1,i}$, $i \in 0:k-1$, $(k \ge 2)$ принадлежат области Ω и

$$\|\eta_i\| \le \gamma_i , \quad \|\xi_i\| \le \beta_i , \quad i \in 0:k-1 .$$

Покажем, что $z_{k+1,k}$ и $z_{k+1,k+1}$ также принадлежат Ω и

$$\|\eta_k\| \le \gamma_k , \quad \|\xi_k\| \le \beta_k .$$

Применим (5) к функции G в точках z_{kk} и $z_{k-1,k-1}$:

$$\|G(z_{kk})\| \le \|G(z_{k-1,k-1}) + G_X(z_{k-1,k-1})\eta_{k-1} + G_U(z_{k-1,k-1})\xi_{k-1}\| + \frac{L}{2}(\|\eta_{k-1}\| + \|\xi_{k-1}\|)^2 .$$

Согласно (10)

$$\|G(z_{kk})\| \le \|G_U z_{k-1,k-1}\|\beta_{k-1} + \frac{L}{2}\alpha_{k-1}^2 .$$

Заметим теперь, что из (12) следует

$$G_U(z_{i+1,i}) + G_{UU}(z_{i+1,i})\xi_i = 0 , \quad i \ge 0 .$$

Поэтому, используя опять же (5), имеем

$$\|G_U(z_{k-1,k-1})\| \le \|G_U(z_{k-1,k-2}) + G_{UU}(z_{k-1,k-2})\xi_{k-2}\| + \frac{L}{2}\|\xi_{k-2}\|^2 \le \frac{L}{2}\beta_{k-2}^2 ,$$

откуда $\|G(z_{kk})\| \le \frac{L}{2}(\beta_{k-1}\beta_{k-2}^2 + \alpha_{k-1}^2)$. Учитывая (9), (11), получаем

$$\|G(z_{kk})\| \le \frac{L}{2}(\mu\gamma_{k-1}\beta_{k-1} + \alpha_{k-1}^2) \le \frac{L}{2}\left(1 + \frac{\mu}{4}\right)\alpha_{k-1}^2$$

и $\|\eta_k\| \le A\|G(z_{kk})\| \le \frac{AL}{2}\left(1 + \frac{\mu}{4}\right)\alpha_{k-1}^2 = \gamma_k$.

Значит, $\|z_{k+1,k} - z_{00}\| \le \sum_{i=0}^{k-1}\beta_i + \sum_{i=0}^{k}\gamma_i \le \sum_{i=0}^{\infty}\alpha_i \le \delta$.
Таким образом, точка $z_{k+1,k}$ принадлежит Ω, и можно построить точку $z_{k+1,k+1}$, причем $\|\xi_k\| \le A\|F(z_{k+1,k})\|$.

Оценивая $\|F(z_{k+1,k})\|$ так же, как и $\|F(z_{21})\|$, получаем $\|\xi_k\| \le \beta_k$, и, следовательно, $z_{k+1,k+1} \in \Omega$. Итак, последовательные приближения могут быть построены. Для $l > k$ имеем

$$\|z_{ll} - z_{kk}\| \le \sum_{i=k}^{l-1}\alpha_i . \tag{13}$$

Полагая в (13) $l = k+\nu$ и устремляя k к бесконечности, убеждаемся в том, что последовательность $\{z_{kk}\}_{k=0}^{\infty}$ фундаментальная, и, следовательно, существует предел $\lim\limits_{k\to\infty} z_{kk}$, который обозначим через z^*.

Ясно, что z^* есть решение системы (4). Устремляя в (13) l к бесконечности, получаем оценку скорости сходимости:

$$\|z^* - z_{kk}\| \le \sum_{i=k}^{\infty}\alpha_i = \frac{1}{\lambda c(1+c)}\sum_{i=k}^{\infty}q^{2^i} \le$$

$$\le \frac{q^{2^k}}{\lambda c(1+c)}\sum_{i=0}^{\infty}(q^{2^k})^i = \frac{q^{2^k-1}}{1-q^{2^k}}\alpha_0 , \quad k \ge 0 .$$

Теорема доказана.

3°. Необходимость решения системы (4), обладающей указаннными выше свойствами, может возникнуть не только при решении задачи из п.1°. Рассмотрим, например, минимаксную задачу со связанными ограничениями: найти $\min\limits_{X\in\Omega}\varphi(X)$, где

$$\varphi(X) = \max_{t\in G(X)} f(X,t),$$

$$\Omega = \{X \in E_n \mid g_j(X) \le 0, \ j \in 1:l_1\},$$

$$G(X) = \{t \in E_p \mid h_j(X,t) \le 0, \ j \in 1:l_2\}.$$

Введем множества

$$Q_1(X)=\{\, j\in 1:l_1 \mid\ g_j(X)=0\,\},$$

$$Q_2(X,t)=\{\, j\in 1:l_2 \mid\ h_j(X,t)=0\,\}$$

и функцию Лагранжа

$$\mathcal{L}(X,t,\lambda,Q)=f(X,t)-\Sigma_{j\in Q}\lambda_j h_j(X,t),\quad Q\subset 1:l_2\,.$$

Пусть в точке $X^*\in\Omega$ функция $\varphi(X)$ достигает минимума на множестве Ω, $R(X^*)=\{t_1^*, t_2^*, \ldots, t_m^*\}$, а число индексов в $Q_1(X^*)$ равно m_1, причем $m+m_1=n+1$. Так как точки t_i^*, $i\in 1:m$, принадлежат $R(X^*)$, то в силу необходимых условий максимума найдутся такие множители Лагранжа $\lambda_i^*\geqslant 0$, $i\in 1:m$, что

$$\mathcal{L}'_t(X^*, t_i^*, \lambda_i^*, Q_2(X^*, t_i^*))=0,\quad i\in 1:m\,.$$

Поэтому точки X^*, t_i^*, $i\in 1:m$, и множители Лагранжа λ_i^*, $i\in 1:m$, будут удовлетворять следующей системе уравнений:

$$\mathcal{L}'_t(X,t_i,\lambda_i,Q_2(X^*,t_i^*))=0,$$
$$h_j(X,t_i)=0,\quad j\in Q_2(X^*,t_i^*),\qquad i\in 1:m, \tag{14}$$

$$\mathcal{L}(X,t_i,\lambda_i,Q_2(X^*,t_i^*))-\mathcal{L}(X,t_m,\lambda_m,Q_2(X^*,t_m^*))=0,$$
$$g_j(X)=0,\quad j\in Q_1(X^*),\qquad i\in 1:m-1. \tag{15}$$

Допустим, что индексные множества $Q_1(X^*)$, $Q_2(X^*, t_i^*)$ известны. Тогда для решения системы (14), (15) можно применить метод выравнивания максимумов. Итерации следует вести поочередно: сначала по X из (15), а затем по t_i, λ_i из (14).

§ 1.5. Метод обобщенного градиента с полной релаксацией

Настоящий параграф посвящен решению задачи минимизации произвольной выпуклой функции $f(X)$ на всем пространстве E_n и на выпуклом замкнутом множестве $\Omega\subset E_n$. Изложенные алгоритмы могут быть использованы и для минимизации выпуклой функции максимума.

1°. Пусть на E_n задана выпуклая функция $f(X)$. Требуется

найти точку $X^*\in E_n$ такую, что

$$f(X^*)=\min_{X\in E_n} f(X). \tag{1}$$

Как известно (см., например, [34, 97, 101]), функция $f(X)$ является непрерывной в каждой точке E_n и дифференцируемой по направлениям, причем

$$\frac{\partial f(X)}{\partial g}\equiv\lim_{\alpha\to+0}\frac{1}{\alpha}[\,f(X+\alpha g)-f(X)]=\max_{V\in M(X)}(V,g), \tag{2}$$

где $$M(X)=\{\, V\in E_n \mid\ f(Z)-f(X)\geqslant(V,Z-X)\quad\forall Z\in E_n\,\}.$$

Множество $M(X)$ называется множеством субградиентов (или обобщенных градиентов). Оно замкнуто, выпукло и ограничено в каждой точке $X\in E_n$. Нетрудно показать, что это множество является ограниченным равномерно по X из любого компактного множества. Точечно-множественное отображение $M(X)$ полунепрерывно сверху, т.е. из $X_k\to X_0$, $Z_k\to Z_0$ и $Z_k\in M(X_k)$ следует $Z_0\in M(X_0)$. Справедливо следующее утверждение (см. [39, 97]):

для того чтобы функция $f(X)$ достигала минимального на E_n значения в точке X^*, необходимо и достаточно, чтобы

$$0\in M(X^*). \tag{3}$$

Пусть в точке X_0 условие (3) не выполнено. Найдем точку $V(X_0)\in M(X_0)$ такую, что

$$\rho(X_0)=\|V(X_0)\|=\min_{V\in M(X_0)}\|V\|. \tag{4}$$

Очевидно $\|V(X_0)\|>0$. Тогда направление $g(X_0)=-V(X_0)/\|V(X_0)\|$ является направлением наискорейшего спуска функции $f(X)$ в точке X_0 (и оно единственно).

Методы последовательных приближений для минимизации функции максимума, основанные на использовании направлений наискорейшего спуска или близких к ним направлений, содержатся в [38]. Для минимизации произвольных выпуклых функций в [48, 116] разработаны методы, использующие не только направления наискорейшего спуска, но и произвольные обобщенные градиенты. По методу обобщенного градиента, например, последовательные приближения строятся по формуле

(см. [116, 117])

$$x_{k+1}=x_k+\alpha_k g_k,\quad g_k=-v_k/\|v_k\|,\quad v_k\in M(x_k),\qquad \alpha_k\to+0,\ \sum_{k=0}^{\infty}\alpha_k=\infty. \tag{5}$$

Этот метод представляет собой значительный вклад в теорию минимизации выпуклых функций. Существует и ряд обобщений этого метода (например, метод с растяжением пространства [117-119] и т.д.). В [90, 91] метод обобщенного градиента распространен на случай ограничений (см. также [127]). Заметим, что метод обобщенного градиента не является релаксационным (метод последовательных приближений называется релаксационным, если $f(x_{k+1})\le f(x_k)$), и, кроме того, условие (5) не всегда удобно для вычислений.

Метод последовательных приближений называется методом с полной релаксацией, если $x_{k+1}=x_k+\alpha_k g_k$ и $f(x_k+\alpha_k g_k)=\min_{\alpha\ge 0} f(x_k+\alpha g_k)$. Функция $f(x)$ называется строго выпуклой, если

$$f(\alpha x_1+(1-\alpha)x_2)<\alpha f(x_1)+(1-\alpha)f(x_2)$$

для всех $\alpha\in(0,1)$ и всех $x_1\ne x_2$.

Лемма 1. Если $f(x)$ - строго выпуклая функция, точка $x_0\in E_n$ такова, что множество

$$D(x_0)=\{x \mid f(x)\le f(x_0)\} \tag{6}$$

ограничено, то всякий метод с полной релаксацией, начинающийся в точке x_0, удовлетворяет условию

$$\|x_{k+1}-x_k\|\to 0\ (k\to\infty). \tag{7}$$

Доказательство. Допустим противное. Пусть существуют число $h>0$ и подпоследовательность $\{x_{k_s}\}$ такие, что $\|x_{k_s+1}-x_{k_s}\|\ge h$. В силу ограниченности множества $D(x_0)$ без ограничения общности можно считать, что $x_{k_s}\to\bar{x}$, $x_{k_s+1}\to\bar{\bar{x}}$. Очевидно также, что $f(\bar{x})=f(\bar{\bar{x}})$ и для всех k

$$f(x_k)\ge f(\bar{x}). \tag{8}$$

Но вследствие строгой выпуклости функции f окажется $f(\frac{1}{2}\bar{x}+\frac{1}{2}\bar{\bar{x}})=\bar{f}<f(\bar{x})$ и при достаточно больших k_s (поскольку метод с полной релаксацией)

$$f(x_{k_s+1})=\min_{\alpha\ge 0} f(x_{k_s}+\alpha(x_{k_s+1}-x_{k_s}))\le f\left(\tfrac{1}{2}x_{k_s}+\tfrac{1}{2}x_{k_s+1}\right)<f(\bar{x}),$$

что противоречит (8).

Пусть $\varepsilon>0$. Точку x^* назовем ε-стационарной точкой функции $f(x)$, если $\rho(x^*)\le\varepsilon$. Предположим, что $f(x)$ - строго выпуклая функция. Опишем метод с полной релаксацией для нахождения ε_0-стационарных точек. Пусть числа $\varepsilon_0>0$, $\mu>0$, $h>0$ фиксированы. Выберем $x_0\in E_n$ и предположим, что множество $D(x_0)$ (см. (6)) ограничено. Положим $b=\sup_{x\in D(x_0)}\sup_{v\in M(x)}\|v\|$. Это число конечно. Можно считать, что $\varepsilon_0<b$.

Возьмем целое число m такое, что

$$m>\frac{\ln\frac{\varepsilon_0}{b}}{\ln q},\qquad q=\frac{1}{\sqrt{1+\left(\frac{\varepsilon_0}{b}\right)^2}}<1.$$

Пусть уже найдено x_k. Положим $x_{k0}=x_k$. Возьмем любое $v_{k0}\in M(x_{k0})$. Если $\|v_{k0}\|\le\varepsilon_0$, то точка x_k ε_0-стационарна, и процесс прекращается. В противном случае положим $x_{k0\alpha}=x_{k0}-\alpha v_{k0}$ и найдем $x_{k1}=x_{k0\alpha_{k0}}$ такое, что

$$f(x_{k0\alpha_{k0}})=\min_{\alpha\ge 0} f(x_{k0\alpha}). \tag{9}$$

Если $\|x_{k1}-x_{k0}\|=\|\alpha_{k0}v_{k0}\|\ge h$ или $f(x_{k0})-f(x_{k1})\ge\mu$, то полагаем $x_{k+1}=x_{k1}$ (т.е. переходим к следующему шагу, если значительно удалились от точки x_{k0} либо значительно уменьшилось значение функции).

В противном случае ищем $\bar{v}_{k0}\in M(x_{k1})$ такое, что

$$(\bar{v}_{k0},v_{k0})\le 0 \tag{10}$$

(указанное $\bar{v}_{k0}$ найдется, ибо согласно (9) в точке x_{k1} должно быть $\partial f(x_{k1})/\partial g_{k0}\ge 0$, где $g_{k0}=-v_{k0}$. Тогда из (2) и вытекает существование $\bar{v}_{k0}$, удовлетворяющего (10)).

Если $\|v_{k0}\|\le\varepsilon_0$, то полагаем $x_{k+1}=x_{k1}$, и процесс прекращается (ибо точка x_{k+1} уже ε_0-стационарна). В противном случае возьмем такой вектор:

$$v_{k1}=\beta_1 v_{k0}+(1-\beta_1)\bar{v}_{k0},\qquad \|v_{k1}\|=\min_{\beta\in[0,1]}\|\beta v_{k0}+(1-\beta)\bar{v}_{k0}\|.$$

Теперь рассмотрим луч $x_{k1\alpha}=x_{k1}-\alpha v_{k1}$ и найдем такое $\alpha_{k1}>0$, что $f(x_{k1\alpha_{k1}})=\min_{\alpha\ge 0} f(x_{k1\alpha})$. Полагаем $x_{k2}=x_{k1\alpha_{k1}}$.

Если $\|X_{k2}-X_{k1}\|=\|\alpha_{k1}V_{k1}\|\geqslant h$ или $f(X_{k1})-f(X_{k2})\geqslant\mu$, то берем $X_{k+1}=X_{k2}$. В противном случае находим $\bar{V}_{k1}\in M(X_{k2})$ такое, что

$$(\bar{V}_{k1},V_{k1})\leqslant 0\,.$$

Если $\|\bar{V}_{k1}\|\leqslant\varepsilon_0$, то полагаем $X_{k+1}=X_{k2}$, и процесс прекращается (точка X_{k+1} ε_0-стационарна). Если же $\|\bar{V}_{k1}\|>\varepsilon_0$, продолжаем аналогично до тех пор, пока не отыщется такое r_k, что либо $\|X_{kr_k}-X_{k,r_k-1}\|=\|\alpha_{kr_k}V_{kr_k}\|\geqslant h$, либо $f(X_{k,r_k-1})-f(X_{kr_k})\geqslant\mu$, либо $\|\bar{V}_{kr_k}\|\leqslant\varepsilon_0$, либо $r_k=m$. Тогда полагаем $X_{k+1}=X_{kr_k}$. Заметим, что

$$(\bar{V}_{ki},V_{ki})\leqslant 0\quad\forall i\in 0:r_k\,. \tag{11}$$

В результате строим последовательность точек $\{X_k\}$. Если она конечна, то для последней полученной точки X_k оказывается $\rho(X_k)\leqslant\varepsilon_0$ (см. (4)). Если же последовательность $\{X_k\}$ содержит бесконечное число элементов, то справедлива следующая теорема.

Теорема 1. Всякая предельная точка X^* последовательности $\{X_k\}$ является ε_0-стационарной точкой функции $f(X)$, т.е.

$$\rho(X^*)\leqslant\varepsilon_0\,. \tag{12}$$

Доказательство. Пусть $X_{k_s}\to X^*$. Допустим, что (12) не выполнено. Вследствие полунепрерывности отображения $M(X)$ найдется такое $\delta>0$, что

$$\min_{X\in B_\delta(X^*)}\ \min_{V\in M(X)}\|V\|>\varepsilon_0,\ \text{где}\ B_\delta(X^*)=\{X\mid\|X-X^*\|\leqslant\delta\}\,,$$

т.е. при X, достаточно близких к X^*, все множества $M(X)$ будут содержаться в некоторой окрестности множества $M(X^*)$ (на рис. 1 эта окрестность показана пунктиром), расстояние которой до начала координат больше ε_0.

По лемме 1 при достаточно больших k_s окажется $r_{k_s}=m$, $X_{k_si}\in B_\delta(X^*)$ $\forall i\in 0:r_{k_s}$, $\|\bar{V}_{k_si}\|>\varepsilon_0$, а также

$$\|V_{k_si}\|>\varepsilon_0\,. \tag{13}$$

Но так как $\|V_{k_si}\|\leqslant b$, то из (11) и (13) нетрудно установить (поскольку на рис. 1 угол $\alpha\geqslant\pi/2$), что

$$\|V_{k_s,i+1}\|\leqslant\|V_{k_si}\|\frac{1}{\sqrt{1+(\varepsilon_0/b)^2}}=\|V_{k_si}\|q\,,$$

где $q<1$. Отсюда $\|V_{k_si}\|\leqslant\|V_{k_s0}\|q^i\leqslant bq^i$, значит, $\|V_{k_s+1}\|=\|V_{k_sm}\|\leqslant bq^m$. Но поскольку $m>\ln(\varepsilon_0/b)/\ln q$, то $\|V_{k_s+1}\|<\varepsilon_0$, что противоречит (13). Теорема доказана.

Замечание 1. Предположение о строгой выпуклости $f(X)$ не является ограничительным, ибо всякую выпуклую функцию $f(X)$ можно аппроксимировать с любой заданной точностью строго выпуклой функцией, например $f_1(X)=f(X)+\varepsilon X^2,\varepsilon>0$.

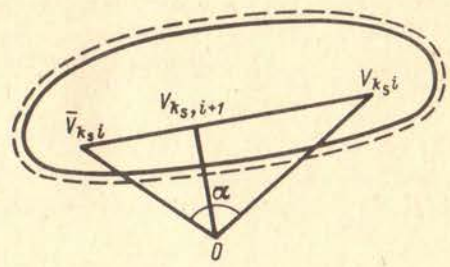


Рис. 1.

Замечание 2. Выше была изложена "принципиальная" схема алгоритма. Можно построить методы, в которых полная релаксация не будет производиться (тогда и условие (11) заменится условием $(\bar{V}_{ki},V_{ki})\leqslant\gamma$, $\gamma>0$). Один такой метод с неполной релаксацией, близкий по идее к описанному выше, изложен в [149-151, 181-183].

Замечание 3. Проверка условий $\|X_{ki}-X_{k,i-1}\|\geqslant h$ и $f(X_{k,i-1})-f(X_{ki})\geqslant\mu$ производится во избежание излишних вычислений в начале процесса.

Замечание 4. В случае непрерывно дифференцируемых функций изложенный метод является одним из методов сопряженных направлений (см. [92, 134, 138]).

Замечание 5. Обозначим $f^*=\min_{X\in E_n}f(X)$, а через D - диаметр множества $D(X_0)$. Пусть X^* - точка, удовлетворяющая (12). По определению множества $M(X^*)$ имеем

$$f(Z)-f(X^*)\geqslant(V,Z-X^*)\quad\forall Z\in E_n,\ \forall V\in M(X^*)\,.$$

Отсюда

$$f(Z)-f(X^*)\geqslant\max_{V\in M(X^*)}(V,Z-X^*)\,.$$

Так как $\min_{X\in E_n}f(X)=\min_{X\in D(X_0)}f(X)$, то

$$f^*-f(X^*)\geqslant\min_{Z\in D(X_0)}\ \max_{V\in M(X^*)}(V,Z-X^*)\geqslant D\min_{\|g\|\leqslant 1}\ \max_{V\in M(X^*)}(V,g)\,.$$

Как и в [39], можно показать, что $\min\limits_{\|g\|\le 1} \max\limits_{V\in M(X^*)} (V, g) = -\rho(X^*)$.

Поэтому $f(X^*)\le f^*+\varepsilon_0 D$. Отсюда ясно также, что для любого фиксированного $\varepsilon > \varepsilon_0$ с помощью описанного выше метода за конечное число шагов будет найдена такая точка X_k, что $f(X_k)\le f^*+\varepsilon D$.

2°. Рассмотрим задачу минимизации выпуклой функции $f(X)$ на выпуклом замкнутом множестве Ω. Положим

$$\gamma(X)=\left\{ V=\lambda(Z-X) \;\middle|\; \lambda>0,\ Z\in\Omega \right\}.$$

Замыкание множества $\gamma(X)$ обозначим через $\Gamma(X)$. Конус $\Gamma(X)$ называется конусом возможных направлений множества Ω в точке X. Положим

$$\Gamma^+(X)=\left\{ W\in E_n \;\middle|\; (W,V)\ge 0 \quad \forall V\in\Gamma(X) \right\}. \tag{14}$$

Конус $\Gamma^+(X)$ называется сопряженным конусу $\Gamma(X)$. Нетрудно установить, что точечно-множественное отображение $\Gamma^+(X)$ полунепрерывно сверху. Поскольку $\Gamma^{++}=\Gamma$ (см. [57]), то для любого $V\in\gamma(X)$ найдется такое $g\in\Gamma^+(X)$, что

$$(V, g)\le 0. \tag{15}$$

В [97] установлено следующее условие. Для того чтобы функция $f(X)$ достигала минимального на Ω значения в точке $X^*\in\Omega$, необходимо и достаточно, чтобы

$$M(X^*)\cap\Gamma^+(X^*)\ne\emptyset. \tag{16}$$

Лемма 2. Если множество Ω имеет внутренние точки, то условие (16) эквивалентно условию

$$0\in L(X^*), \tag{17}$$

где $L(X)=\operatorname{co}\left\{M(X)\cup T_\eta(X)\right\}$, $T_\eta(X)=\left\{ V \;\middle|\; -V\in\Gamma^+(X),\ \|V\|=\eta \right\}$, $\eta>0$ – произвольное фиксированное число.

Доказательство. Вначале установим, что из (16) следует (17). Действительно, если выполнено (16), то существуют $V\in M(X^*)$ и $W\in\Gamma^+(X^*)$ такие, что $V=W$. Если $W=0$, то и $V=0$, т.е. (17) выполнено. Если же $W\ne 0$, то $g=-\eta W/\|W\|\in T_\eta(X^*)$ и $\eta V/\|W\|+g=0$. Значит, $V_\alpha=\alpha V+(1-\alpha)g=0$, где $\alpha=\frac{\eta}{\|W\|}\Big/\left(1+\frac{\eta}{\|W\|}\right)$. Но $V_\alpha\in L(X^*)$, поэтому (17) выполнено.

Покажем теперь, что из (17) вытекает (16). Из (17) следует существование таких $V_i\in M(X^*)$, $i\in 1:m_1$ и $g_j\in T_\eta(X^*)$, $j\in 1:m_2$, что $\sum_{i=1}^{m_1}\alpha_i V_i+\sum_{j=1}^{m_2}\beta_j g_j=0$, где $\alpha_i\ge 0$, $\beta_j\ge 0$, $\sum\alpha_i+\sum\beta_j=1$.

Если среди коэффициентов α_i есть хотя бы один положительный, то $V=W$, где

$$V=\sum_{i=1}^{m_1}\alpha_i' V_i,\quad W=-\sum_{j=1}^{m_2}\beta_j' g_j,\quad \alpha_i'=\frac{\alpha_i}{\alpha},\quad \beta_j'=\frac{\beta_j}{\alpha},\quad \alpha=\sum_{i=1}^{m_1}\alpha_i.$$

Так как $\sum_{i=1}^{m_1}\alpha_i'=1$, то $V=\sum\alpha_i' V_i\in M(X^*)$, $W=-\sum\beta_j' g_j\in\Gamma^+(X^*)$, и поскольку $V=W$, то (16) выполнено. Осталось показать, что среди α_i есть хотя бы один ненулевой коэффициент. Допустим противное. Тогда

$$W=-\sum_{j=1}^{m_2}\beta_j g_j=0, \tag{18}$$

причем $\beta_j\ge 0$, $\sum_{j=1}^{m_2}\beta_j=1$, $W_j=-g_j\in\Gamma^+(X^*)$, $\|W_j\|=\eta>0$. Пусть $\bar X$ – внутренняя точка множества Ω, существование которой предполагается. Тогда найдется $\delta>0$ такое, что $B_\delta(\bar X)=\{X \mid \|X-\bar X\|\le\delta\}\subset\Omega$. Так как $W_j\in\Gamma^+(X^*)$, то для всех $V\in\Gamma(X^*)$ окажется

$$(V, W_j)\ge 0, \tag{19}$$

в частности $(X_j-X^*, W_j)\ge 0$, $(\bar X-X^*, W_j)\ge 0$, где $X_j=\bar X+r_j$, $r_j=\delta W_j/\eta$, ибо $X_j\in B_\delta(\bar X)$, $\|r_j\|=\delta$. Отсюда

$$(X_j-X^*, W_j)\ge\delta\eta>0. \tag{20}$$

Для $i\ne j$ из (19) следует

$$(X_i-X^*, W_j)\ge 0. \tag{21}$$

Умножаем обе части неравенства (20) на β_j, а i-го неравенства (21) – на β_i. Складывая эти неравенства, получаем

$$\left(\sum_{i=1}^{m_2}\beta_i X_i-X^*, W_j\right)\ge\beta_j\delta\eta. \tag{22}$$

Полагаем $\tilde X=\sum_{i=1}^{m_2}\beta_i X_i$. Ясно, что $\tilde X\in\Omega$. Неравенство (22) справедливо для всех $j\in 1:m_2$. Умножая j-е неравенство на β_j и складывая все полученные неравенства, имеем

$$\left(\tilde X-X^*, \sum_{j=1}^{m_2}\beta_j W_j\right)\ge\delta\eta\sum_{j=1}^{m_2}\beta_j^2>0,$$

что противоречит (18). Лемма доказана.

Положим $\varphi(X)=\min\limits_{\substack{g\in\Gamma(X)\\ \|g\|\leqslant 1}}\max\limits_{V\in M(X)}(V,g)$, $\quad d(X)=\min\limits_{\substack{V\in M(X)\\ W\in\Gamma^+(X)}}\|V-W\|$.

В [39, гл. IV] доказано, что

$$\varphi(X)=-d(X). \tag{23}$$

Пусть $f^*=\min\limits_{X\in\Omega} f(X)$, множество Ω ограничено и D – его диаметр. Тогда, как и в замечании 5,

$$f(X)\leqslant f^*-D\varphi(X)=f^*+Dd(X). \tag{24}$$

3°. В лемме 2 была установлена эквивалентность условий (16) и (17) (если множество Ω содержит внутренние точки).

Положим

$$\rho(X)=\min_{V\in L(X)}\|V\|, \tag{25}$$

где множество $L(X)=L_\eta(X)$ определено в лемме 2. Установим связь между $\rho(X)$ и $d(X)$. Пусть

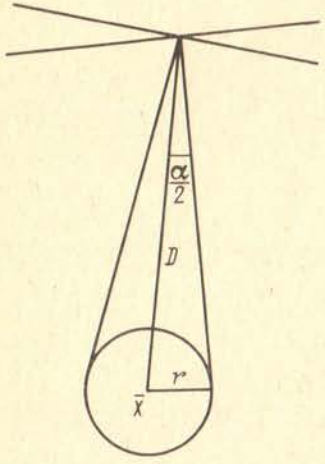


Рис. 2.

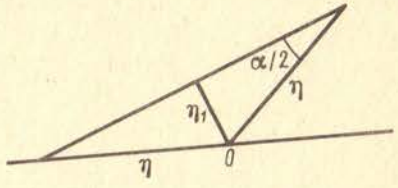


Рис. 3.

$$C=\sup_{X\in\Omega}\sup_{V\in M(X)}\|V\|,$$

D – диаметр множества Ω. Пусть $\bar{X}$ – внутренняя точка множества Ω, $r>0$ – наибольшее число такое, что

$$B_r(\bar{X})=\{X\mid\|X-\bar{X}\|\leqslant r\}\subset\Omega.$$

Тогда, рассуждая как при доказательстве леммы 2 (рис. 2), можно установить, что для любой точки $X\in\Omega$ сопряженный конус $\Gamma^+(X)$ содержится в телесном угле, не превосходящем $2\pi-\alpha$, где $\sin(\alpha/2)=r/D$. Отсюда (рис. 3)

$$\min_{V\in\operatorname{co}T_\eta(X)}\|V\|\geqslant\eta_1=\eta\sin\frac{\alpha}{2}=\frac{r\eta}{D}.$$

А.Б.Певный доказал следующую лемму.

Лемма 3. Для любого $X\in\Omega$ справедливо неравенство

$$d(X)\leqslant\left(1+\frac{2CD}{\eta r}\right)\rho(X). \tag{26}$$

Доказательство. Пусть λ – произвольное число из $(0,1)$. Если $\rho(X)>\lambda\eta_1$, то

$$d(X)\leqslant C\leqslant\frac{C}{\lambda\eta_1}\rho(X). \tag{27}$$

Пусть теперь $\rho(X)\leqslant\lambda\eta_1$. Тогда $\rho(X)=\|V(X)\|$, причем $V(X)$ можно представить в виде $V(X)=\alpha V_1+(1-\alpha)V_2$, где $V_1\in M(X)$, $V_2\in\operatorname{co}T_\eta(X)$, $\alpha\in[0,1]$. Так как $\|V_1\|\leqslant C$, $\|V_2\|\leqslant\eta_1$, то

$$\lambda\eta_1\geqslant\rho(X)=\|\alpha V_1+(1-\alpha)V_2\|\geqslant(1-\alpha)\|V_2\|-\alpha\|V_1\|\geqslant(1-\alpha)\eta_1-\alpha C.$$

Отсюда $$\alpha\geqslant\frac{(1-\lambda)\eta_1}{\eta_1+C}>0.$$

Берем $W_2=-V_2\in\Gamma^+(X)$. Тогда имеем

$$\rho(X)=\|\alpha V_1-(1-\alpha)W_2\|=\alpha\left\|V_1-\frac{1-\alpha}{\alpha}W_2\right\|\geqslant\alpha d(X),$$

т.е. $d(X)\leqslant\frac{1}{\alpha}\rho(X)\leqslant\frac{\eta_1+C}{(1-\lambda)\eta_1}\rho(X)$. Отсюда и из (27) $d(X)\leqslant$ $\leqslant K\rho(X)$, где

$$K = \min_{\lambda\in(0,1)} \max\left\{\frac{C}{\lambda\eta_1},\ \frac{\eta_1+C}{(1-\lambda)\eta_1}\right\}.$$

Минимум достигается, когда

$$\frac{C}{\lambda\eta_1} = \frac{\eta_1+C}{(1-\lambda)\eta_1}, \quad \text{т.е.} \quad \frac{\lambda}{C} = \frac{1-\lambda}{\eta_1+C}, \quad \lambda = \frac{C}{\eta_1+2C}.$$

При этом

$$K = \frac{\eta_1+2C}{C}\,\frac{C}{\eta_1} = 1+\frac{2C}{\eta_1} = 1+\frac{2CD}{\eta r},$$

что и требовалось доказать.

С л е д с т в и е. Из (26) и (24) заключаем, что для любого $X\in\Omega$ справедлива оценка

$$0\leqslant f(X)-f^*\leqslant D\,d(X)\leqslant D\left(1+\frac{2CD}{\eta r}\right)\rho(X). \tag{28}$$

З а м е ч а н и е 6. Так как $\rho(X)=\rho_\eta(X)\leqslant d(X)$ при любом $\eta>0$, то из (26) имеем

$$\rho_\eta(X)\leqslant d(X)\leqslant\left(1+\frac{2CD}{\eta r}\right)\rho_\eta(X).$$

Отсюда $\rho_\eta(X)\to d(X)$ при $\eta\to\infty$.

4°. Опишем метод последовательных приближений для минимизации строго выпуклой функции $f(X)$ на выпуклом замкнутом множестве Ω. Возьмем $X_0\in\Omega$. Предположим, что множество $D(X_0)=\{X\in\Omega \mid f(X)\leqslant f(X_0)\}$ ограничено. Пусть m и b – числа, определенные в п. 1° (можно считать, что $b\geqslant\eta$). Зафиксируем $\mu>0$, $h>0$, $\varepsilon_0>0$. В качестве начального приближения выберем точку $X_0\in\Omega$. Пусть уже найдено $X_k\in\Omega$. Положим $X_{k0}=X_k$ и возьмем произвольное $V_{k0}\in L(X_{k0})$. Если $\|V_{k0}\|\leqslant\varepsilon_0$, то процесс прекращается. Если же $\|V_{k0}\|>\varepsilon_0$, то рассмотрим луч $X_{k0\alpha}=X_{k0}-\alpha V_{k0}$ и найдем

$$\min_{\substack{\alpha\geqslant 0\\ X_{k0\alpha}\in\Omega}} f(X_{k0\alpha}) = f(X_{k0\alpha_0}). \tag{29}$$

Положим теперь $X_{k1}=X_{k0\alpha_0}$. Если $\|X_{k1}-X_{k0}\|\geqslant h>0$ или $f(X_{k0})-f(X_{k1})\geqslant\mu>0$, то возьмем $X_{k+1}=X_{k1}$. В противном случае согласно (29) в точке X_{k1} должен найтись $\bar V_{k0}\in L(X_{k1})$ такой, что

$$(V_{k0},\ \bar V_{k0})\leqslant 0. \tag{30}$$

Действительно, если в точке X_{k1} достигается минимум в (29), то это может произойти по двум причинам:

а) либо $\dfrac{\partial f(X_{k1})}{\partial g_{k0}}\geqslant 0$ (где $g_{k0}=-V_{k0}$). Тогда найдется вектор $\bar V_{k0}\in M(X_{k1})$ такой, что имеет место (30);

б) либо $X_{k1}-\beta V_{k0}\in\Omega$ при любом $\beta>0$, т.е. $-V_{k0}\in\tau(X_{k1})$. Тогда (см. (15)) найдется вектор $\bar V_{k0}\in-\Gamma^+(X_{k1})$ такой, что имеет место (30). Без ограничения общности можем считать, что $\|\bar V_{k0}\|=\eta$. Ищем теперь $\min_{\beta\in[0,1]}\|V_{k0\beta}\|=\|V_{k0\beta_0}\|$, где $V_{k0\beta}=\beta V_{k0}+(1-\beta)\bar V_{k0}$, и положим $V_{k1}=V_{k0\beta_0}$. Далее продолжаем аналогично, пока не находим X_{kr_k} такое, что либо $r_k=m$, либо $\|X_{kr_k}-X_{k,r_k-1}\|\geqslant h$, либо $\|\bar V_{kr_k}\|\leqslant\varepsilon_0$. В первых двух случаях полагаем $X_{k+1}=X_{kr_k}$, а в третьем процесс прекращается.

В результате строим последовательность $\{X_k\}$. Если она конечна, то для последней полученной точки X_k оказывается $\rho(X_k)\leqslant\varepsilon_0$ (см. определение (25)). В противном случае (если последовательность $\{X_k\}$ содержит бесконечное число элементов) справедлива следующая теорема.

Теорема 2. Для всякой предельной точки X^* последовательности $\{X_k\}$ выполнено соотношение

$$\rho(X^*)\leqslant\varepsilon_0. \tag{31}$$

Доказательство этой теоремы аналогично доказательству теоремы 1.

З а м е ч а н и е 7. Из (28) имеем

$$0\leqslant f(X^*)-f^*\leqslant D\left(1+\frac{2CD}{\eta r}\right)\varepsilon_0. \tag{32}$$

5°. Выше рассматривалась задача минимизации произвольной строго выпуклой функции на произвольном выпуклом множестве. При дополнительной информации о функции f и множестве Ω можно ослабить некоторые условия, при которых сходится алгоритм.

Пусть

$$\varphi(X)=\max_{Y\in G} f(X,Y),\quad \Omega=\{X\in E_n \mid h(X,Z)\leqslant 0 \quad \forall Z\in\omega\},$$

где $G\subset E_m$, $\omega\subset E_p$ - замкнутые ограниченные множества в соответствующих пространствах, функции $f(X,Y)$ и $h(X,Z)$ непрерывны вместе с f'_X и h'_X соответственно на $\Omega\times G$ и $E_n\times\omega$. Функция $h(X,Z)$ предполагается выпуклой по X при каждом фиксированном $Z\in\omega$, кроме того, предполагается выполненным условие Слейтера.

Функция $\varphi(X)$ не обязательно выпукла (не говоря уже о строгой выпуклости), однако она является дифференцируемой по направлениям, причем (см. [39, 97])

$$\frac{\partial\varphi(X)}{\partial g}=\max_{Y\in R(X)}(f'_X(X,Y),g)=\max_{Z\in M(X)}(Z,g), \tag{33}$$

где $R(X)=\{Y\in G \mid f(X,Y)=\varphi(X)\}$, $M(X)=\operatorname{co}H(X)$, $H(X)=\{Z=f'_X(X,Y) \mid Y\in R(X)\}$.

Формула (33) дает возможность применить описанный в п. 4° метод. Поскольку для $X\in\Omega$

$$\Gamma^+(X)=\{W=-\textstyle\sum_{i=1}^{r}\alpha_i h'_X(X,Z_i) \mid \alpha_i\geqslant 0,\ Z_i\in Q(X),$$

r - произвольное натуральное число$\}$,

где $Q(X)=\{Z\in\omega \mid h(X,Z)=0\}$, то при нахождении вектора $\bar V_{k0}$, удовлетворяющего (30), вместо множества $L(X_{k1})$ достаточно взять множество

$$H(X_{k1})\cup\{-h'_X(X_k,Z) \mid Z\in Q(X_{k1})\}.$$

Задача нахождения вектора $\bar V_{k0}$ особенно упрощается, если множества $R(X_{k1})$ и $Q(X_{k1})$ состоят из конечного числа точек.

§ 1.6. Метод экстремального базиса

1°. Пусть

$$\varphi(X)=\max_{Y\in G} f(X,Y), \tag{1}$$

где функция $f(X,Y)$ непрерывна вместе с $f'_X(X,Y)$ по совокупности переменных на $S\times G$ и сильно выпукла по X на S при каждом фиксированном $Y\in G$, G - компактное множество, $S\subset E_n$ - ограниченное замкнутое множество. Предполагается, что множество S имеет простую структуру, например задается неравенствами $|x_i|\leqslant C_i$, $i\in 1:n$ (здесь $X=(x_1,\dots,x_n)$).

Положив $$h_i(X)=\begin{cases} x_i-C_i, & i\in 1:n,\\ -x_{i-n}-C_i, & i\in n+1:2n,\end{cases}$$

можно задать S так:

$$S=\{X\in E_n \mid h_i(X)\leqslant 0 \quad \forall i\in 1:2n\}. \tag{2}$$

Требуется найти

$$\min_{X\in\Omega}\varphi(X) \tag{3}$$

где $\Omega=\Omega'\cap S$,

$$\Omega'=\{X\in E_n \mid h(X,Z)\leqslant 0 \quad \forall Z\in\omega\}, \tag{4}$$

ω - компактное множество, функция $h(X,Z)$ непрерывна вместе с $h'_X(X,Z)$ по совокупности переменных на $S\times\omega$ и выпукла по X при каждом фиксированном $Z\in\omega$. Кроме того, предполагается выполненным условие Слейтера: существует точка $\bar X\in\Omega$ такая, что

$$\max_{Z\in\omega} h(\bar X,Z)=-\mu<0. \tag{5}$$

Многие методы последовательных приближений, разработанные для решения задачи (3), требуют многократного вычисления функции (1), что может оказаться трудоемким, особенно если множество G содержит много точек.

Цель настоящего параграфа - свести решение задачи (3) к последовательности задач минимизации функций более простого, чем (1), вида. Исходной точкой для дальнейшего послужила идея алгоритмов Валле-Пуссена и Ремеза для задачи чебышевской аппроксимации [99, 100].

Пусть $X\in\Omega$. Введем множества

$$R(X)=\{Y\in G \mid f(X,Y)=\varphi(X)\},$$

$$H(X)=\{f'_X(X,Y)\mid Y\in R(X)\},\quad L(X)=\operatorname{co}H(X).$$

Пусть $\gamma(X)=\{V=\lambda(Z-X)\mid \lambda>0,\ Z\in\Omega\}$. Замыкание $\gamma(X)$ обозначим через $\Gamma(X)$. Это конус возможных направлений множества Ω в точке X. Через $\Gamma^+(X)$ обозначим конус, сопряженный конусу $\Gamma(X)$, т.е.

$$\Gamma^+(X)=\{W\in E_n\mid (W,V)\geqslant 0\ \ \forall V\in\Gamma(X)\}.$$

Если $\Omega=\Omega'\cap S$, где Ω' задано соотношениями (4), а S - соотношениями (2), то

$$\Gamma^+(X)=\Gamma^+_{\Omega'}(X)+\Gamma^+_S(X),$$

где $\Gamma^+_\Omega(X)$ и $\Gamma^+_S(X)$ - конусы, сопряженные конусам возможных направлений множеств Ω' и S соответственно, т.е.

$$\Gamma^+_{\Omega'}(X)=\{W=-\textstyle\sum_{z_i\in Q'(X)}\beta_i h'_X(X,z_i)\mid \beta_i\geqslant 0\},$$

$$Q'(X)=\{z\in\omega\mid h(X,z)=0\},$$

$$\Gamma^+_S(X)=\{W=-\textstyle\sum_{i\in Q_S(X)}\gamma_i h'_i(X)\mid \gamma_i\geqslant 0\},$$

$$Q_S(X)=\{i\in 1{:}2n\mid h_i(X)=0\}.$$

Если $Q'(X)=\emptyset$, то $\Gamma^+_{\Omega'}(X)=\{0\}$, а если $Q_S(X)=\emptyset$, то $\Gamma^+_S(X)=\{0\}$.

Напомним необходимые и достаточные условия минимума для задачи 3. Для того чтобы точка $X^*\in\Omega$ доставляла минимальное значение функции $\varphi(X)$ вида (1), необходимо и достаточно, чтобы

$$\Gamma^+(X^*)\cap L(X^*)\neq\emptyset. \tag{6}$$

Условие (6) эквивалентно условию

$$0\in\tilde L(X^*), \tag{7}$$

где $$\tilde L(X)=\operatorname{co}\{L(X)\cup T(X)\},$$

$$T(X)=\{h'_X(X,z)\mid z\in Q'(X)\}\cup\{h'_i(X)\mid i\in Q_S(X)\}.$$

Изложенный далее метод основан на том, что по теореме Каратеодори (см., например, [57, с. 783] или [39, с. 306]) при $z\in\tilde L(X)$ существуют такие множества

$$\sigma'_*=\{Y_0,\dots,Y_q\}\quad (Y_i\in R(X)),\quad \sigma''_*=\{z_0,\dots,z_t\}\quad (z_j\in\omega),$$

$$\sigma'''_*=\{l_1,\dots,l_p\}\quad (l_k\in Q_S(X))$$

(причем $q\geqslant 0$, $q+t+p\leqslant n-1$), что

$$z=\textstyle\sum_{i=0}^q\alpha_i f'_X(X,Y_i)+\sum_{j=0}^t\beta_j h'_X(X,z_j)+\sum_{k=1}^p\gamma_k h'_k(X),$$

где $\alpha_i\geqslant 0$, $\beta_j\geqslant 0$, $\gamma_k\geqslant 0$, $\sum_{i=0}^q\alpha_i+\sum_{j=0}^t\beta_j+\sum_{k=1}^p\gamma_k=1$, $\sum_{i=0}^q\alpha_i>0$.

Отсюда, в частности, следует, что точка X^* является также точкой минимума функции $\varphi^*(X)=\max\limits_{Y\in\sigma'_*} f(X,Y)$ на множестве $\Omega^*=\{X\mid h(X,z)\leqslant 0\ \ \forall z\in\sigma''_*\}\cap S$.

Таким образом, если бы множества σ'_* и σ''_* были известны заранее, то достаточно было бы найти $\min\limits_{X\in\Omega^*}\varphi^*(X)$. Изложенный ниже метод состоит в последовательном уточнении указанных множеств. При решении задачи полиномиальной чебышевской аппроксимации методом Ремеза на каждом шаге удерживалось ровно $n+1$ точек (множества σ'_* и σ''_* содержат оба не более $n+1$ точек). Этого достаточно вследствие полного альтернанса в задаче аппроксимации. Общая минимаксная задача может и не обладать таким свойством, поэтому ниже приходится вводить дополнительные точки, гарантирующие сходимость.

Определение. Пусть $\varepsilon>0$, $\mu\geqslant 0$. Точку $X^*\in\Omega$ будем называть (ε,μ)-стационарной точкой функции $\varphi(X)$ на множестве Ω, если

$$0\in\tilde L_{\varepsilon\mu}(X^*), \tag{8}$$

где $\tilde L_{\varepsilon\mu}(X)=\operatorname{co}\{L_\varepsilon(X)\cup T_\mu(X)\}$, $L_\varepsilon(X)=\operatorname{co}H_\varepsilon(X)$,

$$H_\varepsilon(X)=\{f'_X(X,Y)\mid Y\in R_\varepsilon(X)\},\quad R_\varepsilon(X)=\{Y\in G\mid \varphi(X)-f(X,Y)\leqslant\varepsilon\},$$

$$T_\mu(X)=\{h'_X(X,z)\mid z\in Q'_\mu(X)\}\cup\{h'_i(X)\mid i\in Q_\varepsilon(X)\},$$

$$Q'_\mu(X)=\{z\in\omega\mid -\mu\leqslant h(X,z)\leqslant 0\}.$$

Базисом будем называть набор точек $\sigma=\sigma'\cup\sigma''$, где

$$\sigma' = \{Y_0, \ldots, Y_q\} \quad (Y_i \in G), \qquad \sigma'' = \{Z_0, \ldots, Z_t\} \quad (Z_j \in \omega),$$

причем $q \geq 0$, $s+t=n+1$ (так что базис содержит $n+3$ точек, причем множество σ'' может быть пустым).

Выберем в качестве начального произвольный базис $\sigma_0 = \sigma_0' \cup \sigma_0''$, где

$$\sigma_0' = \{Y_{00}, \ldots, Y_{0q_0}\} \quad (Y_{0i} \in G), \qquad \sigma_0'' = \{Z_{00}, \ldots, Z_{0t_0}\} \quad (Z_{0j} \in \omega),$$

причем $q_0 \geq 0$, $q_0 + t_0 = n+1$. Пусть уже построен базис $\sigma_k = \sigma_k' \cup \sigma_k''$, где

$$\sigma_k' = \{Y_{k0}, \ldots, Y_{kq_k}\} \quad (Y_{ki} \in G), \qquad \sigma_k'' = \{Z_{k0}, \ldots, Z_{kt_k}\} \quad (Z_{kj} \in \omega),$$

причем $q_k \geq 0$, $q_k + t_k = n+1$. Положим

$$\varphi_k(X) = \max_{Y \in \sigma_k'} f(X, Y), \qquad \Omega_k = \{X \in E_n \mid h(X, Z) \leq 0 \ \ \forall Z \in \sigma_k''\} \cap S$$

и найдем точку $X_k \in \Omega_k$ такую, что $\varphi_k(X_k) = \min_{X \in \Omega_k} \varphi_k(X)$. В точке X_k выполнено условие

$$0 \in L_k, \tag{9}$$

где $L_k = \operatorname{co} H_k$, $H_k = H_k' \cup H_k'' \cup H_k'''$, $H_k' = \{f_X'(X_k, Y) \mid Y \in R_k\}$,
$H_k'' = \{h_X'(X_k, Z) \mid Z \in Q_k\}$, $H_k''' = \{h_i'(X_k) \mid h_i(X_k) = 0\}$,
$R_k = \{Y \in \sigma_k' \mid f(X_k, Y) = \varphi_k(X_k)\}$, $Q_k = \{Z \in \sigma_k'' \mid h(X_k, Z) = 0\}$.
Существует представление начала координат (9), содержащее не более $n+1$ вектор из H_k. Из условия Слейтера вытекает, что в это представление входит хотя бы один вектор из H_k'. Таким образом, в базисе σ_k есть по крайней мере две точки V_{k1}, V_{k2}, которым соответствуют точки из H_k, не входящие в представление начала координат (9).

Положим $\bar{\sigma}_k = \bar{\sigma}_k' \cup \bar{\sigma}_k'' = \sigma_k \setminus \{V_{k1}, V_{k2}\}$,

$$\bar{\varphi}_k(X) = \max_{Y \in \bar{\sigma}_k'} f(X, Y), \qquad \bar{\Omega}_k = \{X \mid h(X, Z) \leq 0 \ \ \forall Z \in \bar{\sigma}_k''\} \cap S.$$

Ясно, что тогда

$$\bar{\varphi}_k(X) = \min_{X \in \bar{\Omega}_k} \varphi_k(X). \tag{10}$$

Найдем теперь точки $Y_k' \in G$ и $Z_k' \in \omega$ такие, что

$$f(X_k, Y_k') = \varphi(X_k), \tag{11}$$

$$h(X_k, Z_k') = \max_{Z \in \omega} h(X_k, Z) = h_k, \tag{12}$$

и построим базис σ_{k+1}, в котором точки V_{k1} и V_{k2} заменены точками Y_k' и Z_k', т.е. $\sigma_{k+1} = \sigma_{k+1}' \cup \sigma_{k+1}''$, где $\sigma_{k+1}' = \bar{\sigma}_k' \cup \{Y_k'\}$, $\sigma_{k+1}'' = \bar{\sigma}_k'' \cup \{Z_k'\}$.

Если $\varphi(X_k) = \varphi_k(X_k)$, то точку V_{k1} не меняем, а если $h_k = 0$, то не меняем точку V_{k2}. Если $\varphi(X_k) = \varphi_k(X_k)$ и $h_k = 0$ одновременно, то X_k – точка минимума функции $\varphi(X)$ на множестве Ω. Если последовательность точек $\{X_k\}$ конечна, то последняя полученная точка является точкой минимума функции $\varphi(X)$ на множестве Ω. Если же последовательность $\{X_k\}$ содержит бесконечное число точек, то справедлива следующая теорема.

Теорема 1. Последовательность точек $\{X_k\}$ сходится к точке минимума функции $\varphi(X)$ на множестве Ω.

Доказательство. Функцию $f(X, Y)$ предполагаем сильно выпуклой по X на S при всяком фиксированном $Y \in G$, т.е. существует такое $m > 0$, что

$$f(X+\Delta, Y) \geq f(X, Y) + (f_X'(X, Y), \Delta) + \frac{1}{2} m \|\Delta\|^2 \tag{13}$$

для всех $X, X+\Delta \in S$, $Y \in G$.

Последовательность точек $\{X_k\}$ ограниченная (ибо все $X_k \in S$). Положим $\Delta_k = X_{k+1} - X_k$. Из (10) следует

$$0 \in \bar{L}_k, \tag{14}$$

где $\bar{L}_k = \operatorname{co} \bar{H}_k$, $\bar{H}_k = \bar{H}_k' \cup \bar{H}_k'' \cup H_k'''$,
$\bar{H}_k' = \{f_X'(X_k, Y) \mid Y \in R_k \cap \bar{\sigma}_k'\}$, $\bar{H}_k'' = \{h_X'(X_k, Z) \mid Z \in Q_k \cap \bar{\sigma}_k''\}$.
Поскольку $X_{k+1} \in \Omega_{k+1} \subset \bar{\Omega}_k$, то $(\Delta_k, V) \leq 0$ для всех $V \in \bar{H}_k'' \cup H_k'''$. Поэтому из (14) заключаем, что найдется такое $Y_k \in \bar{\sigma}_k'$, что

$$(f_X'(X_k, Y_k), \Delta_k) \geq 0. \tag{15}$$

Далее из (13) имеем

$$f(X_{k+1}, Y_k) \geq f(X_k, Y_k) + (f_X'(X_k, Y_k), \Delta_k) + \frac{1}{2} m \|\Delta_k\|^2. \tag{16}$$

Так как $f(X_k, Y_k) = \varphi_k(X_k)$, то

$$\varphi_{k+1}(X_{k+1}) = \max_{Y \in \sigma'_{k+1}} f(X_{k+1}, Y) \geq \max_{Y \in \bar{\sigma}'_k} f(X_{k+1}, Y) \geq$$

$$\geq f(X_{k+1}, Y_k) \geq \varphi_k(X_k) + \frac{1}{2} m \|\Delta_k\|^2 . \qquad (17)$$

Отсюда

$$\varphi_{k+1}(X_{k+1}) \geq \varphi_k(X_k) . \qquad (18)$$

Покажем, что $\Delta_k \to 0$. Действительно, допустив существование подпоследовательности $\{k_s\}$ такой, что $\|\Delta_{k_s}\| \geq \alpha > 0$, из (17) и (18) получаем $\varphi_k(X_k) \to +\infty$, что невозможно, ибо $\varphi(X_k) \geq \varphi_k(X_k)$, а функция $\varphi(X)$ ограничена на множестве S. Итак,

$$\Delta_k \to 0 . \qquad (19)$$

Отсюда и из (12) заключаем, что

$$h_k \to 0 , \qquad (20)$$

ибо $h(X_k, Z'_k) = h_k \geq 0$, а $h(X_{k+1}, Z'_k) \leq 0$. Значит,

$$0 \geq h(X_{k+1}, Z'_k) = h_k + (h'_X(X'_k, Z'_k), \Delta_k), \text{ где } X'_k \in (X_k, X_{k+1}) . \qquad (21)$$

Второе слагаемое правой части (21) стремится к нулю, ибо $\|h'_X(X, Z)\|$ ограничена, а $\Delta_k \to 0$. Поэтому и $h_k \to 0$. Из (19) также следует

$$\varphi(X_k) - \varphi_k(X_k) \to 0 . \qquad (22)$$

Из (9), (20), (22) заключаем, что точка X_k является (ε_k, μ_k)-стационарной точкой функции $\varphi(X)$ на множестве

$$\bar{\Omega}_k = \{ X \mid h(X, Z) - h_k \leq 0 \quad \forall Z \in \omega \} \cap S ,$$

где $\varepsilon_k = \varphi(X_k) - \varphi_k(X_k)$, $\mu_k = h_k$, т.е.

$$0 \in \tilde{L}_{\varepsilon_k \mu_k}(X_k) . \qquad (23)$$

Теперь, пользуясь (23), нетрудно показать, что если $X_{k_q} \to X^*$, а X_{k_q} есть $(\varepsilon_{k_q}, \mu_{k_q})$-стационарная точка, причем $\varepsilon_{k_q} \to 0$, $\mu_{k_q} \to 0$, то X^* - точка минимума функции $\varphi(X)$ на множестве Ω. Из сильной выпуклости функции $\varphi(X)$ следует единственность предельной точки. Теорема доказана.

З а м е ч а н и е 1. Требование сильной выпуклости функции $\varphi(X)$ существенно, так как при отсутствии этого условия может произойти зацикливание.

2°. Приведенный алгоритм практически неосуществим, ибо каждый раз требуется находить минимум функции $\varphi_k(X)$ на множестве Ω_k, что в большинстве случаев является задачей, не решаемой за конечное число шагов. Рассмотрим модификацию метода экстремального базиса (МЭБ), изложенного выше, которая не требует бесконечной процедуры на каждом шаге.

Выберем последовательности $\{\varepsilon_k\}$, $\{\mu_k\}$, $\{\rho_k\}$ такие, что $\varepsilon_k \to +0$, $\mu_k \to +0$, $\rho_k \to +0$,

$$\sum_{k=0}^{\infty} \varepsilon_k < \infty , \quad \sum_{k=0}^{\infty} \mu_k < \infty , \quad \sum_{k=0}^{\infty} \rho_k < \infty . \qquad (24)$$

В качестве начального выберем произвольный базис $\sigma_0 = \sigma'_0 \cup \sigma''_0$. Пусть уже найден базис $\sigma_k = \sigma'_k \cup \sigma''_k$. Строим, как и ранее, функцию $\varphi_k(X)$ и множество Ω_k. За конечное число шагов найдем (используя один из методов первого порядка, например из [39]) точку $X_k \in \Omega_k$ такую, что

$$\min_{Z \in L_{k\varepsilon_k\mu_k}} \|Z\| = \|Z_k\| \leq \rho_k ,$$

где $L_{k\varepsilon\mu} = \operatorname{co} H_{k\varepsilon\mu}$, $H_{k\varepsilon\mu} = H'_{k\varepsilon} \cup H''_{k\mu} \cup H'''_{k\varepsilon}$,

$$H'_{k\varepsilon} = \{ f'_X(X_k, Y) \mid Y \in R_{k\varepsilon} \}, \quad R_{k\varepsilon} = \{ Y \in \sigma'_k \mid \varphi_k(X_k) - f(X_k, Y) \leq \varepsilon \},$$

$$H''_{k\mu} = \{ h'_X(X_k, Z) \mid Z \in Q_{k\mu} \}, \quad Q_{k\mu} = \{ Z \in \sigma''_k \mid 0 \leq h(X_k, Z) \leq \mu \},$$

$$H'''_{k\varepsilon} = \{ h'_{iX}(X_k) \mid -\varepsilon \leq h_i(X_k) \leq 0 \} .$$

Как и в п. 1°, точка Z_k может быть представлена в виде выпуклой комбинации не более чем $n+1$ точек из $H_{k\varepsilon_k\mu_k}$. Поэтому в базисе σ_k есть две точки, которые можно заменить по правилу, приведенному в п. 1°. Но можно заменять точки и другим способом. Точки Y_k и Z'_k можно находить не из условий (11) и (12), а из соотношений

$$\varphi(X_k) - f(X_k, Y'_k) \leq \alpha_k , \qquad (25)$$

$$h_k - h(X_k, Z'_k) \leqslant \beta_k , \qquad (26)$$

где $h_k = \max\limits_{Z\in\Omega} h(X_k, Z)$, $\alpha_k \to +0$, $\beta_k \to +0$, т.е. задачи (II) и (I2) можно решать приближенно.

В результате строим последовательность точек, для которой также справедлива теорема I.

3°. Рассмотрим задачу минимизации произвольной выпуклой функции $f(X)$ на произвольном выпуклом множестве $\Omega \subset S \subset E_n$ (где S - ограниченное замкнутое множество сравнительно простой структуры), содержащем внутренние точки. Пусть $f(X)$ задана на S. Тогда с каждой точкой $X \in S$ связано множество $M(X)$ субградиентов функции $f(X)$, т.е.

$$f(Z) \geqslant f(X) + (A(X), Z - X) \qquad (27)$$

для всех $A(X) \in M(X)$ и $Z \in S$. Легко убедиться, что

$$f(X) = \max_{Y\in S} [(A(Y), X - Y) + f(Y)] , \qquad (28)$$

где $A(Y)$ - произвольный субградиент функции f в точке Y. Действительно, из (27)

$$f(X) \geqslant \max_{Y\in S} [f(Y) + (A(Y), X - Y)] . \qquad (29)$$

Неравенство же $\max\limits_{Y\in S} [(A(Y), X - Y) + f(Y)] \geqslant f(X)$ очевидно. Отсюда и из (29) следует (28).

Пусть $\Gamma^+(X)$ - конус, сопряженный конусу возможных направлений множества Ω в точке X. Через Γ_Ω обозначим множество граничных точек множества Ω. Нетрудно убедиться, что

$$\Omega = \{ X \mid (B(Z), X - Z) \geqslant 0 \quad \forall Z \in \Gamma_\Omega \} , \qquad (30)$$

где $B(Z) \in \Gamma^+(Z)$, $\|B(Z)\| = 1$.

Таким образом, задача минимизации выпуклой функции $f(X)$ на выпуклом множестве Ω эквивалентна задаче минимизации функции максимума линейных функций (28) на множестве Ω, заданном линейными неравенствами (30). Теперь можно применить изложенный в п. I° метод экстремального базиса.

Выберем произвольный базис $\sigma_0 = \sigma'_0 \cup \sigma''_0$. Пусть уже найден базис $\sigma_k = \sigma'_k \cup \sigma''_k$, где

$$\sigma'_k = \{ Y_{k0}, \ldots, Y_{kq_k} \} \ (Y_{ki} \in S), \quad \sigma''_k = \{ Z_{k0}, \ldots, Z_{kt_k} \} \ (Z_{kj} \in \Gamma_\Omega),$$

причем $q_k \geqslant 0$, $q_k + t_k = n+1$. Для каждой точки Y_{ki} найдем любой вектор $A(Y_{ki}) \in M(Y_{ki})$, а для каждой точки Z_{kj} - любой вектор $B(Z_{kj}) \in \Gamma^+(Z_{kj})$. Образуем функцию

$$\varphi_k(X) = \max_{Y\in\sigma'_k} [(A(Y), X - Y) + f(Y)] \qquad (31)$$

и множество

$$\Omega_k = \{ X \mid (B(Z), X - Z) \geqslant 0 \quad \forall Z \in \sigma''_k \} \cap S . \qquad (32)$$

Найдем $\min\limits_{X\in\Omega_k} \varphi_k(X) = \varphi_k(X_k)$. Это задача линейного программирования, ибо с учетом (2) ее можно переписать так: найти

$$\min v \qquad (33)$$

при ограничениях

$$\begin{aligned} &(A(Y), X - Y) + f(Y) \leqslant v \quad \forall Y \in \sigma'_k , \\ &(-B(Z), X - Z) \leqslant 0 \quad \forall Z \in \sigma''_k , \\ &x_i \leqslant c_i , \quad -x_i \leqslant c_i \quad \forall i \in 1:n \end{aligned} \qquad (34)$$

(общее число ограничений $3n+3$, но последние $2n$ из них тривиальные).

"Лишние" точки в базисе σ_k находим так же, как и в п. I° (с помощью теоремы Каратеодори). В качестве новых точек вводим точки $Y'_k = X_k$ и Z'_k, где

$$\|Z'_k - X_k\| = \min_{Z\in\Omega} \|Z - X_k\| . \qquad (35)$$

Если $Z'_k = X_k$ (т.е. $X_k \in \Omega$), то в множество σ''_{k+1} новую точку не вводим. Вообще говоря, задача нахождения Z'_k (см. задачу (35)) может оказаться довольно сложной, поэтому если $X_k \bar{\in} \Omega$ и известна точка Слейтера $\bar{X}$, удовлетворяющая неравенству (5), то в качестве Z'_k можно взять точку $Z'_k = \alpha_k X_k + (1 - \alpha_k)\bar{X}$ ($\alpha_k \in (0,1)$), являющуюся граничной точкой множества Ω (на отрезке $[X_k, \bar{X}]$ существует единственная точка множества Γ_Ω).

Итак, задача минимизации выпуклой функции $f(X)$ на выпук-

лом множестве Ω свелась к последовательности задач линейного программирования. Функция в правой части (28), по которой берется максимум, не является сильно выпуклой, поэтому, как уже отмечалось в п. 1°, метод может "зациклиться". Избежать зацикливания можно так же, как это делается в линейном программировании. Кроме того, можно заменить линейную функцию в (33) функцией

$$v + \varepsilon [X^2 + v^2], \tag{36}$$

где $\varepsilon > 0$ - достаточно малое число. Тогда задача (33), (34) станет задачей квадратичного программирования, а функция превратится в сильно выпуклую, поэтому метод экстремального базиса будет сходиться. Правда, при этом на каждом шаге придется решать задачу квадратичного программирования (для этого можно воспользоваться одним из конечных методов). К решению задачи квадратичного программирования можно переходить лишь тогда, когда четко определилось зацикливание. Минимум функции (36) на множестве, заданном неравенствами (34), близок к минимуму исходной функции.

З а м е ч а н и е 2. Изложенный метод отличается от метода секущих плоскостей Келли [146] тем, что число линейных ограничений не возрастает с каждым шагом, а остается постоянным. В работе [111] предложена модификация метода Келли, в которой каждый раз удерживается ровно $n+1$ точек (а не $n+3$, как в данной работе). Однако сходимость такого метода гарантируется лишь при выполнении некоторых дополнительных условий.

4°*. Пусть $f(X)$ - сильно выпуклая функция, определенная на S, т.е. существует такое $\mu > 0$, что

$$f(X) \geq f(Y) + (V, X-Y) + \mu(X-Y)^2 \quad \forall X, Y \in S, \; V \in M(Y). \tag{37}$$

Тогда, как легко видеть,

$$f(X) = \max_{Y \in S} [(V(Y), X-Y) + f(Y) + \mu(X-Y)^2], \tag{38}$$

где $V(Y)$ - произвольный вектор из $M(Y)$.

Как и в п. 3°, задача минимизации сильно выпуклой функции $f(X)$ на множестве $\Omega \subset S$ сводится к минимизации функции (38) на множестве Ω, заданном неравенствами (30).

Применим теперь метод экстремального базиса.

---

* П. 4° написан В.Н.Тарасовым.

Поскольку функция, по которой берется максимум в (38), сильно выпуклая, то метод экстремального базиса сходится к точке минимума. При этом на каждом шаге необходимо решать такую задачу: найти минимум функции

$$\varphi(X) = \max_{i \in 1:l} f_i(X) \tag{39}$$

при ограничениях

$$(A_i, X) \leq b_i \quad \forall i \in l+1 : n+3, \tag{40}$$

где $f_i(X) = (X, X) + (a_i, X) + h_i$; a_i и A_i - заданные n-мерные векторы, h_i - заданные числа. По теореме Куна-Таккера найдутся такие коэффициенты $\alpha_i^* \geq 0$, $\sum_{i=1}^{l} \alpha_i^* = 1$ $(i \in 1:n+3)$, что

$$F(X^*, \alpha) \leq F(X^*, \alpha^*) \leq F(X, \alpha^*) \quad \forall X \in E_n, \; \alpha \in \Lambda, \tag{41}$$

где

$$\Lambda = \{\alpha = (\alpha_1, \ldots, \alpha_{n+3}) \mid \alpha_i \geq 0, \; \textstyle\sum_{i=1}^{l} \alpha_i = 1\},$$
$$F(X, \alpha) = \textstyle\sum_{i=1}^{l} \alpha_i f_i(X) + \sum_{i=l+1}^{n+3} \alpha_i [(A_i, X) - b_i], \tag{42}$$

т.е. $[X^*, \alpha^*]$ - седловая точка функции $F(X, \alpha)$ на $E_n \times \Lambda$. Кроме того,

$$\max_{\alpha \in \Lambda} \min_{X \in E_n} F(X, \alpha) = \min_{X \in E_n} \max_{\alpha \in \Lambda} F(X, \alpha) = F(X^*, \alpha^*). \tag{43}$$

Приравнивая нулю F'_X, из (41) имеем

$$2X + \textstyle\sum_{i=1}^{l} \alpha_i a_i + \sum_{i=l+1}^{n+3} \alpha_i A_i = 0.$$

Отсюда

$$X = -\frac{1}{2} \left[\textstyle\sum_{i=1}^{l} \alpha_i a_i + \sum_{i=l+1}^{n+3} \alpha_i A_i \right]. \tag{44}$$

Подставляя (44) в (42), из (43) получаем следующую задачу: найти $\max_{\alpha \in \Lambda} F_1(\alpha)$, причем функция $F_1(\alpha)$ квадратичная. Таким образом, задача (39), (40) свелась к задаче квадратичного программирования.

Итак, для сильно выпуклых функций метод экстремального базиса приводит к последовательности задач квадратичного программирования. При этом зацикливание исключается.

5°. Рассмотрим задачу отыскания седловых точек. Пусть заданы множества

$$\Omega = \{ X \in E_n \mid h_{1i}(X) \leqslant 0 \quad \forall i \in I \},$$

$$G = \{ Y \in E_m \mid h_{2j}(Y) \leqslant 0 \quad \forall j \in J \},$$

где $I = 1:N_1$, $J = N_1+1:N$, $h_{1i}(X)$ и $h_{2j}(X)$ – выпуклые функции в соответствующих пространствах, множества Ω и G удовлетворяют условию Слейтера. Очевидно, что Ω и G – выпуклые множества.

Пусть на $E_n \times E_m$ определена функция $f(X,Y)$, строго выпуклая по X и строго вогнутая по Y. Требуется найти седловую точку функции $f(X,Y)$ на множестве $\Omega \times G$, т.е. точку $[X^*, Y^*] \in \Omega \times G$, такую, что

$$f(X^*,Y) \leqslant f(X^*,Y^*) \leqslant f(X,Y^*) \quad \forall X \in \Omega, Y \in G. \qquad (45)$$

Задача нахождения седловых точек является минимаксной задачей, ибо $f(X^*,Y^*) = \min\limits_{X \in \Omega} \max\limits_{Y \in G} f(X,Y)$, поэтому для их отыскания можно применять методы, разработанные для решения минимаксных задач. Однако использование особенностей рассматриваемой задачи позволяет разработать специальные методы (см. [23,40,52,54]).

Если $N_1 \gg n$, $(N-N_1) \gg m$, то для нахождения точки, удовлетворяющей (45), можно применить метод экстремального базиса.

Выберем последовательность $\{\beta_k\}$, $\beta_k > 0$, $\beta_k \to 0$. Возьмем произвольный базис $\sigma_{10} = \{i_{00}, \ldots, i_{0,n+1}\}$ $(i_{0s} \in I)$. Пусть уже найден базис σ_{1k}. Поскольку

$$\min_{X \in \Omega_k} \max_{Y \in G} f(X,Y) = \max_{Y \in G} \min_{X \in \Omega_k} f(X,Y),$$

где $\Omega_k = \{ X \in E_n \mid h_{1i}(X) \leqslant 0 \quad \forall i \in \sigma_{1k} \}$, то, применяя метод экстремального базиса к задаче

$$\max_{Y \in G} F_k(Y), \quad \text{где} \quad F_k(Y) = \min_{X \in \Omega_k} f(X,Y),$$

за конечное число шагов найдем такой базис $\sigma_{2k} = \{ j_{k0}, \ldots, j_{k,m+1} \}$ $(j_{ks} \in J)$ и такую точку $[X_k, Y_k] \in \Omega_k \times G_k$, где $G_k = \{ Y \in E_m \mid h_{2j}(Y) \leqslant 0 \quad \forall j \in \sigma_{2k} \}$, что

$$f(X_k,Y_k) = \max_{X \in G_k} \min_{X \in \Omega_k} f(X,Y) = \min_{X \in \Omega_k} \max_{Y \in G_k} f(X,Y), \qquad (46)$$

при этом

$$\max_{j \in J} h_{2j}(Y_k) \leqslant \beta_k. \qquad (47)$$

Новый базис $\sigma_{1,k+1}$ строим, как обычно (вводим в базис $i_k \in I$, для которого $h_{1i_k}(X_k) = \max\limits_{i \in I} h_{1i}(X_k)$).

Здесь предполагается, что задача отыскания седловых точек на множествах, заданных небольшим числом ограничений (см. задачу (46)), может быть эффективно решена.

Теорема 2. Изложенный метод за конечное число шагов позволит найти седловую точку функции $f(X,Y)$ на множестве $\Omega \times G$.

Доказательство. Поскольку число множеств G_k и Ω_k конечно, то из условия (47) следует, что при достаточно больших k справедливо $Y_k \in G$, т.е.

$$f(X_k,Y_k) = \max_{Y \in G} f(X_k,Y) \stackrel{def}{=} \varphi(X_k).$$

Но так как мы применяем метод экстремального базиса для минимизации функции $\varphi(X)$, то за конечное число шагов получаем точку X^* – точку минимума функции $\varphi(X)$:

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

Но тогда $[X^*, Y^*]$ – седловая точка функции $f(X,Y)$ на $\Omega \times G$ (здесь $f(X^*,Y^*) = \max\limits_{Y \in G} f(X^*,Y) \stackrel{def}{=} \varphi(X^*)$).

Замечание 3. Как и в п. 2°, изложенный метод можно применять для нахождения седловых точек на множествах, заданных континуумом неравенств. Правда, в этом случае сходимости в конечное число шагов, вообще говоря, уже не будет.

Замечание 4. В описанном алгоритме задачу (46) можно решать приближенно (как в [37]), но метод перестает быть конечным.

§ 1.7. О непрерывных методах минимизации выпуклых функций

Рассмотрим задачу минимизации на E_n функции $f(X)$. Если функция $f(X)$ непрерывно дифференцируема, то хорошо известно, что при весьма естественных предположениях система дифференциальных уравнений $\dot{X} = -f'_X(X)$, $X(0) = X_0$, имеет решение $f(t, X_0)$, причем всякая предельная точка этого решения является стационарной точкой функции f на E_n, т.е. если $X(t_k, X_0) \to X^*$ при $t_k \to \infty$, то $f'_X(X^*) = 0$. Ряд численных методов минимизации функции $f(X)$

можно рассматривать с точки зрения разностных аналогов приведенной ранее системы дифференциальных уравнений. Так, метод наискорейшего спуска состоит в замене кривой $X(t, X_k)$ ее касательной $X_{k\alpha} = X_k - \alpha f'_X(X_k)$ и последующей минимизации функции $f(X)$ на луче $X_{k\alpha}$ $(\alpha \geq 0)$. Поэтому изучение подобных систем дифференциальных уравнений представляет не только теоретический интерес.

Если же функция $f(X)$ дифференцируема лишь по направлениям, то дифференциальное уравнение естественным образом заменяется дифференциальным включением. Далее будет показано существование решения соответствующего дифференциального включения для выпуклой функции и установлено, что всякое решение, определенное на бесконечном интервале, стремится к точке минимума функции $f(X)$ на E_n.

Пусть $f(X)$ - выпуклая функция, определенная на n-мерном евклидовом пространстве E_n. Тогда функция f непрерывна и в каждой точке $X \in E_n$ имеет субдифференциал $M(X)$ ($M(X)$ - множество субградиентов. См. § I.5). При этом точечно-множественное отображение $M(X)$ полунепрерывно сверху. В дальнейшем нам понадобится следующая лемма.

Лемма I. Для любого $Y \in E_n$ выполняется соотношение

$$M(Y) = \bigcap_{\delta>0} \overline{\text{co}} \bigcup_{X \in S_\delta(Y)} M(X),$$

где $S_\delta(X) = \{X \mid \|X - Y\| < \delta\}$, $\overline{\text{co}}\, A$ - замыкание выпуклой оболочки множества A.

Доказательство. Пусть $\partial f(X)/\partial u$ - производная по направлению u функции f в точке X. Известно, что $\partial f(X)/\partial u$ - опорная функция Минковского субдифференциала $M(X)$. Пусть q_δ - опорная функция множества $\overline{\text{co}} \bigcup_{X \in S_\delta(Y)} M(X)$. Тогда (см. например, [43])

$$q_\delta(u) = \sup_{X \in S_\delta(u)} \frac{\partial f(X)}{\partial u}.$$

При фиксированном $u \in E_n$ функция $\varphi(\delta) = q_\delta(u)$ убывает, и потому существует предел

$$q(u) = \lim_{\delta \to 0} q_\delta(u) = \inf_{\delta>0} q_\delta(u) = \inf_{\delta>0} \sup_{X \in S_\delta(Y)} \frac{\partial f(X)}{\partial u} = \overline{\lim_{X \to Y}} \frac{\partial f(X)}{\partial u}.$$

Непосредственно можно проверить, что сублинейный функционал q является опорной функцией множества $\bigcap_{\delta>0} \overline{\text{co}} \bigcup_{X \in S_\delta(Y)} M(X)$. Так как

отображение $M(X)$ полунепрерывно сверху, то при любом u функция $\omega(X) = \partial f(X)/\partial u = \max_{l \in M(X)} l(u)$ непрерывна сверху (см., например, [67]), т.е. $\partial f(X)/\partial u \geq \overline{\lim_{X \to Y}} \partial f(X)/\partial u = q(u)$. Отсюда следует

$$M(Y) \supset \bigcap_{\delta>0} \overline{\text{co}} \bigcup_{X \in S_\delta(Y)} M(X).$$

Обратное включение очевидно. Лемма доказана.

В дальнейшем будем предполагать, что лебеговы множества $D_c = \{X \mid f(X) \leq c\}$ функции f ограничены и, кроме того, ограничены множества $\bigcup_{X \in D_c} M(X)$.

Рассмотрим дифференциальное включение (уравнение в контингенциях)

$$\dot{X} \in -M(X), \quad X(0) = X_0, \tag{I}$$

предположив, что отображение $M(X)$ определено на множестве D_c. Покажем, что это включение имеет решение, т.е. существует абсолютно непрерывная вектор-функция $X(t)$, определенная на некотором интервале $(t_1, t_2) \ni 0$ и такая, что $\dot{X}(t) \in M(X(t))$ почти при всех $t \in (t_1, t_2)$. Кроме того, опишем начальные условия X_0, при которых решение определено на полупрямой $(t_1, +\infty)$.

Так как отображение $M(X)$ полунепрерывно сверху, то [6I] существует функция $Z(X)$ такая, что:

I) для всех $X \in D_c$ выполняется соотношение $Z(X) \in M(X)$;

2) функция Z - первого класса Бэра, т.е. является пределом некоторой последовательности непрерывных функций. Отсюда вытекает, в частности, что Z измерима. Отметим, кроме того, что Z - ограниченная функция.

Рассмотрим дифференциальное уравнение $\dot{X} = -Z(X)$ с начальным условием $X(0) = X_0$, где $X_0 \in D_c$. Из теоремы А.Ф.Филиппова [II0] следует, что это уравнение имеет решение в смысле Филиппова, т.е. существует абсолютно непрерывная функция $X(t)$, определенная на некотором интервале (t_1, t_2), содержащем нуль, и такая, что $X(0) = X_0$ и почти для всех $t \in (t_1, t_2)$

$$-\dot{X}(t) \in \bigcap_{\delta>0} \text{co} \bigcup_{u \in S_\delta(X(t))} Z(u) \subset \bigcap_{\delta>0} \text{co} \bigcup_{u \in S_\delta(X(t))} M(u).$$

Из леммы I следует $-\dot{X}(t) \in M(X(t))$ почти для всех $t \in (t_1, t_2)$. Из результатов, полученных Филипповым, вытекает также, что

решение определено либо на полуинтервале $(t_1, t_2]$ таком, что $X(t_2)$ лежит на границе множества D_c, либо на полупрямой $(t_1, +\infty)$. Покажем, что для достаточно широкого класса начальных условий решение определено на полупрямой.

Так как лебеговы множества D_c функции f ограничены, то они компактны, и, следовательно, f достигает минимума на E_n. Пусть D^* - множество точек минимума функции f и $\rho(X) = \min\limits_{Y \in D^*} \| Y - X \|$. Положим $b(X) = \min\limits_{\{Y \mid f(Y) = f(X)\}} \rho(Y)$. Можно показать (см., например, [38, с. 76]), что для $X^* \in D^*$, $X \in E_n$ и $Z \in M(X)$ выполняется неравенство

$$(X^* - X,\ Z/\|Z\|) \leqslant -b(X). \qquad (2)$$

Пусть $X(t)$ - решение дифференциального включения (1), определенное на промежутке (t_1, t_2). Тогда почти для всех t из промежутка, на котором определено решение, используя (2), имеем

$$(X^* - X(t),\ -\dot{X}(t)/\|\dot{X}(t)\|) \leqslant -b(X(t)). \qquad (3)$$

(Если на множестве полной меры, для которого выполняется включение $-\dot{X}(t) \in M(X(t))$, существует точка t такая, что $\dot{X}(t) = 0$, то $0 \in M(X(t))$, и, следовательно, f достигает в точке $X(t)$ минимума. Таким образом, в рассматриваемом случае решение дифференциального включения (1) приводит к минимуму. Поэтому будем считать в дальнейшем, что $\|\dot{X}(t)\| > 0$. Тогда неравенство (3) имеет смысл.)

Перепишем (3) в виде $(X(t) - X^*,\ \dot{X}(t)) \leqslant -b(X(t))\,\|\dot{X}(t)\|$. Обозначив $b(X(t))\,\|\dot{X}(t)\|$ через $v(t)$ и проинтегрировав в пределах от t_0 до t, получим

$$\int_{t_0}^{t} (X(\tau) - X^*,\ \dot{X}(\tau))\, d\tau \leqslant -\int_{t_0}^{t} v(\tau)\, d\tau. \qquad (4)$$

Поскольку

$$\int_{t_0}^{t} (X(\tau), \dot{X}(\tau))\, d\tau = \frac{1}{2}\int_{t_0}^{t} \frac{d}{dt}\|X(\tau)\|^2 d\tau = \frac{1}{2}\left(\|X(t)\|^2 - \|X(t_0)\|^2\right),$$

$$\int_{t_0}^{t} (X^*, \dot{X}(\tau))\, d\tau = (X^*, X(t) - X(t_0)),$$

то $\int_{t_0}^{t} (X(\tau) - X^*,\ \dot{X}(\tau))\, d\tau = \frac{1}{2}\left(\|X(t)\|^2 - \|X(t_0)\|^2\right) - (X^*,\ X(t) - X(t_0)) =$

$= \frac{1}{2}\left(\|X(t) - X^*\|^2 - \|X(t_0) - X^*\|^2\right)$, и потому, как вытекает из (4),

$$\|X(t) - X^*\|^2 - \|X(t_0) - X^*\|^2 \leqslant -2\int_{t_0}^{t} v(\tau)\, d\tau. \qquad (5)$$

Так как $v(t) \geqslant 0$ (и $v(t) = 0$ тогда и только тогда, когда $X(t) \in M^*$), то расстояние от точки, лежащей на траектории, до произвольной точки X^* минимума функции f убывает с возрастанием t. Переписывая (5) в виде

$$\|X(t) - X^*\|^2 \leqslant \|X(t_0) - X^*\|^2 - 2\int_{t_0}^{t} v(\tau)\, d\tau$$

и переходя к минимуму по множеству D^*, имеем

$$\rho^2(X(t)) \leqslant \rho^2(X(t_0)) - 2\int_{t_0}^{t} v(\tau)\, d\tau \quad (t_0 < t), \qquad (6)$$

откуда следует, что функция $\rho(X(t))$ убывает.

Пусть число $\varepsilon > 0$ таково, что ε-окрестность $D_\varepsilon^* = \{ Y \mid \rho(Y) < \varepsilon \}$ множества D^* целиком входит во внутренность множества D_c. Если $X_0 \in D_\varepsilon^*$, то из (6) следует $X(t) \in D_\varepsilon^*$ при тех t, для которых определено решение, и, следовательно, $X(t)$ не принадлежит границе множества D_c при любом t. Отсюда, как уже отмечалось, следует, что решение определено на полупрямой $(t_1, +\infty)$. Таким образом, доказана следующая лемма.

Лемма 2. Если окрестность D_ε^* множества D^* лежит во внутренности лебегова множества D_c функции f, то при $X_0 \in D_\varepsilon^*$ дифференциальное включение $\dot{X} \in -M(X)$ с начальным условием $X(0) = X_0$ имеет решение, определенное на полуоси $(t_1, +\infty)$, где $t_1 < 0$.

Справедлива следующая теорема.

Теорема 1. Пусть начальное условие выбрано так, как указано в лемме 2. Тогда для любой траектории $X(t)$, удовлетворяющей дифференциальному включению (1) с данным начальным условием, существует предел $\lim\limits_{t \to +\infty} X(t) = X(\infty)$, причем $X(\infty) \in D^*$.

Доказательство. Перепишем неравенство (6) в виде

$$\int_{t_0}^{t} v(\tau)\, d\tau \leqslant \frac{1}{2}\left[\rho^2(X(t_0)) - \rho^2(X(t))\right]. \qquad (7)$$

Из (7) вытекает, что интеграл $\int_0^\infty v(t)\, dt = \int_0^\infty b(X(t))\,\|\dot{X}(t)\|\, dt$ сходится. Возможны два случая:

1) $\operatorname{vrai}\min\|\dot{X}(t)\| = 0$. Тогда существует последовательность точек $\{t_k\}$ $(t_k \in (0,\infty))$ такая, что $\dot{X}(t_k) \in M(X(t_k))$ и $\dot{X}(t_k) \to 0$. Не умаляя общности, можем считать, что последовательности $\{t_k\}$ и $\{X(t_k)\}$ сходятся. Предположим сначала, что $t_k \to +\infty$. Пусть $Y = \lim X(t_k)$. Так как отображение $M(X)$ полунепрерывно сверху, то $0 \in M(Y)$, откуда следует, что $Y \in D^*$. Как отмечалось выше, расстояние от любой точки $X^* \in D^*$ до точки на траектории $X(t)$ убывает. Отсюда, в частности, вытекает, что $X(t) \underset{t\to\infty}{\to} Y$, т.е. существует предел $X(\infty) = \lim\limits_{t\to+\infty} X(t)$. Если $t_k \to \tau < +\infty$, то, рассуждая аналогично, получаем $\dot{X}(\tau) \in D^*$ и $X(t) = X(\tau)$ при $t > \tau$, т.е. $X(\tau) = X(\infty)$.

2) $\operatorname{vrai}\min\|\dot{X}(t)\| = c > 0$. В этом случае интеграл $\int_0^\infty b(X(t))\,dt$ сходится, и потому существует последовательность $\{t_k\}$ такая, что $t_k \to +\infty$, $b(X(t_k)) \to 0$. Отсюда легко получить (см., например, [38]), что и $\rho(X(t_k)) \to 0$, и, следовательно, предельная точка Y последовательности $\{X(t_k)\}$ входит в D^*. Так как функция $\|X(t) - Y\|$ убывает, то $Y = X(\infty) = \lim\limits_{t\to\infty} X(t)$. Теорема доказана.

Укажем в заключение подход, с помощью которого можно оценить рост функции $\rho(X(t))$. Предположим, что существует число $\theta > 0$ такое, что $b(X) \geqslant \theta\rho(X)$ для всех $X \in D_c$. Положим $z(X) = \min\limits_{l\in M(X)}\|l\|$. Если $X(t)$ – траектория включения (1), то, используя (7), для $t > t_0$ получаем

$$\rho^2(X(t_0)) - \rho^2(X(t)) \geqslant 2\theta\int_{t_0}^{t} \rho(X(\tau))\,z(X(\tau))\,d\tau .$$

Положим $\rho(X(t)) = \varphi(t)$. Отметим прежде всего, что поскольку эта функция убывает, то она почти всюду дифференцируема. Поэтому почти при всех t

$$-2\varphi(t)\dot{\varphi}(t) \geqslant 2\theta\varphi(t)\,z(X(t)),$$

откуда
$$-\dot{\varphi}(t) \geqslant \theta z(X(t)) . \qquad (8)$$

Предположим теперь, что существует функция γ такая, что $z(X) \geqslant \gamma(\rho(X))$ $(X \in D_c)$. Тогда соотношение (8) приводит к дифференциальному неравенству

$$-\dot{\varphi}(t) \geqslant \theta\gamma(\varphi(t)) .$$

Пусть $\gamma(u) = \lambda u$, где $\lambda > 0$ (если f дифференцируема и достигает минимума в единственной точке X^*, то это неравенство означает, что $\|\partial f(X)/\partial X\| \geqslant \gamma\|X - X^*\|$, или, что то же самое,

$$\left\| \frac{\partial f(X)}{\partial X} - \frac{\partial f(X^*)}{\partial X} \right\| \geqslant \gamma\|X - X^*\| \quad).$$

Тогда $\dot{\varphi} + \theta\lambda\varphi \leqslant 0$. Используя теоремы о дифференциальных неравенствах (см., например, [112]), получаем, что $\varphi(t)$ не превосходит решения задачи Коши $\dot{\varphi} + \theta\lambda\varphi = 0$, $\varphi(0) = \rho(X_0)$, т.е.

$$\rho(X(t)) \leqslant \rho(X_0)e^{-\theta\lambda t} .$$

Если $\gamma(u) = \lambda u^\alpha$, то, рассуждая таким же образом, имеем

$$\rho(X(t)) \leqslant \left(\rho^{1-\alpha}(X_0) - \theta\lambda(1-\alpha)t\right)^{\frac{1}{1-\alpha}} .$$

§ 1.8. Метод возможных направлений без управляющих последовательностей

В настоящем параграфе предлагается метод минимизации функции максимума при ограничениях. Метод целесообразно применять, когда число функций, образующих функцию максимума, и число ограничений значительно превышают размерность пространства, в котором производится минимизация.

Пусть $f_i(X)$, $i \in I = 1:N_1$, $h_j(X)$, $j \in J = N_1+1:N_2$, – непрерывно дифференцируемые на E_n функции. Образуем множество $\Omega = \{X \in E_n \mid h_j(X) \leqslant 0 \quad \forall j \in J\}$ и функцию $\varphi(X) = \max\limits_{i\in I} f_i(X)$. Предположим, что функции $h_j(X)$ выпуклы и выполнено условие Слейтера. Пусть $X \in \Omega$. Введем обозначения:

$$H(X) = \{f'_{iX}(X) \mid i \in R(X)\},$$

$$H'(X) = \{h'_{jX}(X) \mid i \in Q(X)\},$$

$$L(X) = \operatorname{co}\{H(X) \cup H'(X)\} .$$

Рассмотрим задачу минимизации функции $\varphi(X)$ на множестве Ω. Известно [39, с. 183], что для того чтобы точка X^* была точкой минимума функции $\varphi(X)$ на множестве Ω, необходимо, а

в случае выпуклости функции φ и достаточно, чтобы

$$0 \in L(X^*). \tag{1}$$

Пусть $\rho(X) = \min_{V \in L(X)} \|V\|$. Тогда (1) эквивалентно соотношению $\rho(X^*) = 0$.

Опишем следующий метод для отыскания стационарных точек.

В качестве начального приближения возьмем произвольную точку $X_0 \in \Omega$. Предположим, что множество $D(X_0) = \{X \in \Omega \mid \varphi(X) \le \varphi(X_0)\}$ ограничено. Пусть уже найдена точка $X_k \in \Omega$. Положим $X_{k0} = X_k$, $R_{k0} = R(X_{k0})$, $Q_{k0} = Q(X_{k0})$, $L_{k0} = L(X_{k0})$, $\rho_{k0} = \rho(X_{k0}) = \|V_{k0}\|$, $V_{k0} \in L_{k0}$. Если $\rho_{k0} = 0$, то точка X_k стационарная, и процесс прекращается. Если же $\rho_{k0} > 0$, то полагаем $g_{k0} = -V_{k0}/\rho_{k0}$, строим луч $X = X_{k0}(\alpha) \stackrel{def}{=} X_{k0} + \alpha g_{k0}$ $(\alpha \ge 0)$. Найдем

$$\min_{\substack{\alpha \ge 0,\ X_{k0}(\alpha) \in \Omega \\ R(X_{k0}(\beta)) \subset R_{k0}\ \forall \beta \in [0,\alpha)}} \varphi(X_{k0}(\alpha)) = \varphi(X_{k0}(\alpha_{k0}))$$

и положим $X_{k1} = X_{k0}(\alpha_{k0})$. Если

$$R(X_{k1}) \subset R_{k0}, \tag{2}$$

$$Q(X_{k1}) \subset Q_{k0}, \tag{3}$$

то выберем $X_{k+1} = X_{k1}$. Если хотя бы одно из включений (2) или (3) не выполнено, то положим

$$R_{k1} = R_{k0} \cup R(X_{k1}), \quad Q_{k1} = Q_{k0} \cup Q(X_{k1}),$$

$$H_{k1} = \{ f'_{iX}(X_{k1}) \quad i \in R_{k1} \},$$

$$H'_{k1} = \{ h'_{jX}(X_{k1}) \quad j \in Q_{k1} \},$$

$$L_{k1} = \operatorname{co}\{ H_{k1} \cup H'_{k1} \},$$

$$\rho_{k1} = \min_{V \in L_{k1}} \|V\| = \|V_{k1}\|, \quad V_{k1} \in L_{k1}.$$

Если $\rho_{k1} = 0$, то полагаем $X_{k+1} = X_{k1}$. Если же $\rho_{k1} > 0$, то строим луч $X = X_{k1}(\alpha) \stackrel{def}{=} X_{k1} + \alpha g_{k1}$ $(\alpha \ge 0)$, где $g_{k1} = -V_{k1}/\rho_{k1}$, и находим

$$\min_{\substack{\alpha \ge 0,\ X_{k1}(\alpha) \in \Omega \\ R(X_{k1}(\beta)) \subset R_{k1}\ \forall \beta \in [0,\alpha)}} \varphi(X_{k1}(\alpha)) = \varphi(X_{k1}(\alpha_{k1})).$$

Положим $X_{k2} = X_{k1}(\alpha_{k1})$. Если $R(X_{k2}) \subset R_{k1}$, $Q(X_{k2}) \subset Q_{k1}$, то берем $X_{k+1} = X_{k2}$. В противном случае полагаем

$$R_{k2} = R_{k1} \cup R(X_{k2}), \quad Q_{k2} = Q_{k1} \cup Q(X_{k2}),$$

$$H_{k2} = \{ f'_{iX}(X_{k2}) \mid i \in R_{k2} \}, \quad H'_{k2} = \{ h'_{jX}(X_{k2}) \mid j \in Q_{k2} \},$$

$$L_{k2} = \operatorname{co}\{ H_{k2} \cup H'_{k2} \},$$

$$\rho_{k2} = \min_{V \in L_{k2}} \|V\| = \|V_{k2}\|, \quad V_{k2} \in L_{k2},$$

и так продолжаем до тех пор, пока не находим t_k такое, что либо $\rho_{kt_k} = 0$ либо

$$R(X_{kt_k}) \subset R_{k,t_k-1}, \quad Q(X_{kt_k}) \subset Q_{k,t_k-1} \tag{4}$$

(число t_k конечно, так как конечно число индексов в множествах I и J).

Теперь положим $X_{k+1} = X_{kt_k}$. Заметим, что $X_{k+1} \in \Omega$, $\varphi(X_{k+1}) < \varphi(X_k)$. В результате построим последовательность $\{X_k\}$. Если эта последовательность состоит из конечного числа точек, то последняя полученная точка стационарна по построению. В противном случае справедлива следующая теорема.

Теорема. Всякая предельная точка последовательности $\{X_k\}$ является стационарной точкой функции $\varphi(X)$ на множестве Ω.

Доказательство. Существование предельных точек следует из ограниченности множества $D(X_0)$. Пусть $X_{k_s} \to X^*$. Очевидно, что $X^* \in \Omega$. Покажем, что $\rho(X^*) = 0$.

Допустим противное, т.е. что

$$\rho(X^*) = \rho^* > 0. \tag{5}$$

Тогда существует такое $\delta > 0$, что для всех $X \in S_\delta(X^*)$

$$R(X) \subset R(X^*), \quad Q(X) \subset Q(X^*),$$

$$\min_{V \in L^*(X)} \|V\| \ge \rho^*/2, \tag{6}$$

где $L^*(X) = \operatorname{co}\{ f'_{iX}(X), h'_{jX}(X) \mid i \in R(X^*), \ j \in Q(X^*) \}$. Без ограничения общности можем считать, что $X_{k_s} \in S_{\delta/2}(X^*)$. В этом случае из (6) следует $\rho(X_{k_s}) \ge \rho^*/2$. Переход от точки X_{k_s} к

точке x_{k_s+1} может произойти по одной из следующих причин:

$$\rho_{k_s t_{k_s}} = 0, \tag{7}$$

$$R(x_{k_s t_{k_s}}) \subset R_{k_s t_{k_s}-1}, \quad Q(x_{k_s t_{k_s}}) \subset Q_{k_s t_{k_s}-1}. \tag{8}$$

При этом: а) либо $x_{k_s l} \in S_\delta(x^*)$ для всех $l \in 0 : t_{k_s}$; б) либо существует такое число $l_{k_s} \in 0 : t_{k_s}$, что $x_{k_s l_{k_s}} \overline{\in} S_\delta(x^*)$.

Без ограничения общности будем считать, что либо для всех k_s выполнено условие а) либо для всех k_s выполнено условие б). В случае а) должно выполняться соотношение (8) (см. (6)). Тогда из (6) следует неравенство

$$\varphi\left(x_{k_s t_{k_s}-1}(\alpha)\right) \leqslant \varphi\left(x_{k_s t_{k_s}-1}\right) - \alpha\rho^*/2 + o_{k_s}(\alpha).$$

Так как $o_{k_s}(\alpha)/\alpha \underset{\alpha \to +0}{\to} 0$ равномерно по k_s, то существует $\beta > 0$ (одно и то же для всех k_s) такое, что

$$\varphi(x_{k_s+1}) \leqslant \varphi(x_{k_s}) - \beta. \tag{9}$$

Рассмотрим случай б). Выберем число $\tau_{k_s} \in 0 : t_{k_s}$ такое, что $x_{k_s l} \in S_\delta(x^*)$ для всех $l \in 0 : \tau_{k_s}$, $x_{k_s \tau_{k_s}+1} \overline{\in} S_\delta(x^*)$. Так как $x_{k_s} \in S_{\delta/2}(x^*)$, то

$$\left\|x_{k_s 1} - x_{k_s 0}\right\| + \left\|x_{k_s 2} - x_{k_s 1}\right\| + \ldots + \left\|x_{k_s \tau_{k_s}+1} - x_{k_s \tau_{k_s}}\right\| \geqslant \frac{\delta}{2},$$

причем при движении из точек $x_{k_s l}$, $l \in 0 : \tau_{k_s}$, имеем $\partial\varphi(x_{k_s l})/\partial g_{k_s l} \leqslant -\rho^*/2$ (в силу (6)). Поэтому снова найдется число $\beta > 0$ такое, что $\varphi(x_{k_s \tau_{k_s}+1}) \leqslant \varphi(x_{k_s}) - \beta$. Тем более выполнено (9). Отсюда вытекает, что $\varphi(x_k) \to -\infty$, что невозможно, ибо множество $D(x_0)$ ограничено. Теорема доказана полностью.

З а м е ч а н и е 1. В описанном методе переход к точке x_{k+1} происходил, если $\rho_{k t_k} = 0$ либо выполнялось условие (4). К точке x_{k+1} можно переходить, если $\rho_{k t_k} = 0$ или условие (4) выполняется n_0 раз подряд (n_0 - любое заранее фиксированное положительное число). При этом утверждение теоремы остается справедливым.

З а м е ч а н и е 2. Рассмотренный метод является одним из способов устранения зацикливания (см. [50]).

З а м е ч а н и е 3. В описанном методе множества R_{kt} и Q_{kt} можно формировать следующим образом.

Пусть уже найдены множества $R_{k,t-1}$, $Q_{k,t-1}$, $R(x_{kt})$ и $Q(x_{kt})$. Через $|A|$ обозначим число элементов в множестве A. Положим

$$l_{kt} = |R_{k,t-1}| + |R(x_{kt}) \setminus R_{k,t-1}|,$$

$$p_{kt} = |Q_{k,t-1}| + |Q(x_{kt}) \setminus Q_{k,t-1}|.$$

Расположим числа $f_i(x_{kt})$ в порядке убывания. В качестве R_{kt} возьмем множество индексов $i \in I$, которым соответствуют первые l_{kt} чисел в упорядоченном множестве $\{f_i(x_{kt})\}$. Множество Q_{kt} строится аналогично: упорядочиваем множество чисел $\{h_j(x_{kt})\}$, в качестве Q_{kt} берем множество индексов $j \in J$, которым соответствуют первые p_{kt} чисел в упорядоченном множестве $\{h_j(x_{kt})\}$.

Указанная модификация метода позволяет ускорить сходимость, ибо тогда из множеств R_{kt} и Q_{kt} "уходят" индексы, соответствующие "неперспективным" функциям f_i и ограничениям h_j. Доказательство сходимости такого модифицированного метода практически не изменяется.

З а м е ч а н и е 4. В описанном выше методе при построении множеств H_{kt} и H'_{kt} требуется находить градиенты функций f_i и h_j в точке x_{kt}. Вместо этого можно положить

$$H_{kt} = H_{k,t-1} \cup \left\{ f'_{ix}(x_{kt}) \mid i \in R(x_{kt}) \right\},$$

$$H'_{kt} = H'_{k,t-1} \cup \left\{ h'_{jx}(x_{kt}) \mid j \in Q(x_{kt}) \right\}.$$

Если применяется модификация метода, описанная в замечании 3, то при построении множеств H_{kt} и H'_{kt} поступаем следующим образом. Поскольку для всякого $i \in R_{k,t-1}$ найдется такое t_i, $0 \leqslant t_i \leqslant t-1$, что $f'_{ix}(x_{kt_i}) \in H_{k,t-1}$, а для всякого $j \in Q_{k,t-1}$ такое t_j, $0 \leqslant t_j \leqslant t-1$, что $h'_{jx}(x_{kt_j}) \in H'_{k,t-1}$, то положим

$$H_{kt} = \left\{ f'_{ix}(x_{kt}) \mid i \in R_{kt} \setminus R_{k,t-1} \right\} \cup \left\{ f'_{ix}(x_{kt_i}) \mid i \in R_{kt} \cap R_{k,t-1} \right\},$$

$$H'_{kt} = \left\{ h'_{jx}(x_{kt}) \mid j \in Q_{kt} \setminus Q_{k,t-1} \right\} \cup \left\{ h'_{jx}(x_{kt_j}) \mid j \in Q_{kt} \cap Q_{k,t-1} \right\}.$$

Иначе говоря, градиенты функций f_i и h_j вычисляем лишь для индексов i и j, которые не входили в множества $R_{k,t-1}$ и $Q_{k,t-1}$ соответственно. Для остальных i и j берем вычисленные в предыдущих точках градиенты. Этот прием целесообразно использовать, если вычисление градиентов трудоемко.

Доказательство теоремы в этом случае также изменяется незначительно.

§ 1.9. Ускорение сходимости при решении нелинейных минимаксных задач

В настоящем параграфе предлагается метод минимизации функции максимума на множестве, заданном неравенствами. Исходная задача сводится к последовательности задач минимизации функции Лагранжа при линейных ограничениях-равенствах. Предполагается, что множество "активных функций" в точке минимума уже известно. Особенностью метода является то, что можно не изменять множители Лагранжа. При этом метод сходится со скоростью геометрической прогрессии. Если множители Лагранжа уточняются на каждом шаге, то скорость сходимости сверхлинейная.

Пусть на открытом множестве E_n заданы дважды непрерывно дифференцируемые функции $f_i(X)$, $i \in I = 1:N$, и $h_j(X)$, $j \in J = 1:N_1$. Предположим, что функции $f_i(X)$ и $h_j(X)$ сильно выпуклые, т.е. для всех $j \in J$, $i \in I$ найдутся такие числа $m_i > 0$, $\bar{m}_j > 0$, $M_i > 0$, $\bar{M}_j > 0$, что для любых $X, g \in E_n$

$$m_i \|g\|^2 \leqslant \left(f''_{iXX}(X)g, g\right) \leqslant M_i \|g\|^2,$$
$$\bar{m}_j \|g\|^2 \leqslant \left(h''_{jXX}(X)g, g\right) \leqslant \bar{M}_j \|g\|^2.$$

Введем в рассмотрение множество

$$\Omega = \left\{ X \in E_n \mid h_j(X) \leqslant 0 \quad \forall j \in J \right\}.$$

Предположим, что Ω ограничено и удовлетворяет условию Слейтера. Образуем функцию $\varphi(X) = \max\limits_{i \in I} f_i(X)$. Рассмотрим задачу минимизации функции $\varphi(X)$ на множестве Ω.

Пусть $X \in \Omega$. Справедливо следующее утверждение (см. § 1.1). Для того чтобы точка X^* была точкой минимума функции $\varphi(X)$ на множестве Ω, необходимо (а в данном случае и достаточно), чтобы выполнялось соотношение

$$\Gamma^+(X^*) \cap L(X^*) \neq \emptyset. \qquad (1)$$

Условие (1) эквивалентно существованию $\alpha_i^* \geqslant 0$, $\lambda_j^* \geqslant 0$ таких, что $\sum_{i \in R(X^*)} \alpha_i^* = 1$ и

$$\sum_{i \in R(X^*)} \alpha_i^* f'_{iX}(X^*) + \sum_{j \in Q(X^*)} \lambda_j h'_{iX}(X^*) = \mathbf{0}. \qquad (2)$$

Представление нуля в (2) может оказаться не единственным. Выберем то из них, которое содержит наименьшее число положительных коэффициентов. Такое представление будем называть минимальным. Оно тоже может оказаться не единственным. Выберем любое из них. Пусть это представление содержит $p \geqslant 1$ положительных коэффициентов α_i^* и $p_1 \geqslant 1$ положительных коэффициентов λ_j^*. Без ограничения общности можем считать, что $R(X^*) = 1:p$, $Q(X^*) = 1:p_1$. Тогда (2) можно переписать в виде

$$\sum_{i=1}^{p} \alpha_i^* f'_{iX}(X^*) + \sum_{j=1}^{p_1} \lambda_j^* h'_{jX}(X^*) = 0,$$
$$\sum_{i=1}^{p} \alpha_i^* = 1, \quad \alpha_i^* > 0, \quad \lambda_j^* > 0. \qquad (3)$$

Лемма 1. Если в (3) $p = 1$, то векторы $\left\{ h'_{jX}(X^*),\ j \in 1:p_1 \right\}$ линейно-независимы, а если $p > 1$, то линейно-независимы и векторы $\left\{ f'_{iX}(X^*) - f'_{pX}(X^*),\ i \in 1:p-1,\ h'_{jX}(X^*),\ j \in 1:p_1 \right\}$.

Доказательство см. в [33].

Далее рассмотрим случай $p > 1$, $p_1 \geqslant 1$.

Пусть $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{p_1})$. Определим функцию

$$F(\alpha, \lambda; X) = \sum_{i=1}^{p} \alpha_i f_i(X) + \sum_{j=1}^{p_1} \lambda_j h_j(X) \qquad (\alpha_i \geqslant 0,\ \lambda_j \geqslant 0),$$

множество

$$B(y) = \left\{ X \in E_n \;\middle|\; \begin{array}{l} f_i(y) - f_p(y) + (f'_{iX}(y) - f'_{pX}(y),\ X - y) = 0,\ i \in 1:p-1 \\ h_j(y) + (h'_{jX}(y),\ X - y) = 0,\ j \in 1:p_1 \end{array} \right\}$$

и при фиксированных α и λ отображение

$$A_{\alpha,\lambda}(y) = \left\{ X \;\middle|\; F(\alpha, \lambda; X) = \min_{Z \in B(y)} F(\alpha, \lambda; Z) \right\}.$$

Отображение $A_{\alpha,\lambda}(y)$ есть отображение E_n в E_n и задает неявную функцию $X(\alpha, \lambda; y)$, определяемую системой нелинейных

уравнений

$$F'_x(\alpha,\lambda;x)+\sum_{i=1}^{p-1}\tau_i(f'_{ix}(y)-f'_{px}(y))+\sum_{j=1}^{p_1}\mu_j h'_{jx}(y)=0,$$

$$f_i(y)-f_p(y)+(f'_{ix}(y)-f'_{px}(y),\ x-y)=0,\quad i\in 1:p-1, \tag{4}$$

$$h_j(y)+(h'_{jx}(y),\ x-y)=0,\quad j\in 1:p_1.$$

Неизвестными в системе (4) являются x,τ_i,μ_j; $i\in 1:p-1$, $j\in 1:p_1$. Решение системы (4) обозначим через $x(\alpha,\lambda;y),\tau_i(\alpha,\lambda;y),\mu_j(\alpha,\lambda;y)$.

Лемма 2. Отображение $A_{\alpha,\lambda}(y)$ непрерывно и непрерывно дифференцируемо по параметрам α, λ и аргументу y в некоторой окрестности точки α^*,λ^*,x^*.

Доказательство следует из теоремы существования и непрерывной дифференцируемости неявной функции (см. [93]), так как якобиан системы (4) не равен нулю [33].

Лемма 3. Производная от отображения $A_{\alpha,\lambda}(y)$ по y при $\alpha=\alpha^*$, $\lambda=\lambda^*$ и $y=x^*$ равна нулевому оператору.

Доказательство. Заметим, что $\tau_i:=\tau_i(\alpha^*,\lambda^*;x^*)=0$ и $\mu_i^*:=\mu_i(\alpha^*,\lambda^*;x^*)=0$. Применяя правило дифференцирования неявной функции $x(\alpha,\lambda;y)$ по y_j, получаем следующую систему линейных уравнений:

$$F''_{xx}(\alpha^*,\lambda^*;x^*)x'_{y_j}(\alpha^*,\lambda^*;x^*)+\sum_{i=1}^{p-1}\tau^*_{iy_j}(f'_{ix}(x^*)-f'_{px}(x^*))+$$
$$+\sum_{k=1}^{p_1}\mu^{*\prime}_{ky_j}h'_{kx}(x^*)=0,$$

$$f'_{iy_j}(x^*)-f'_{py_j}(x^*)+(f'_{ix}(x^*)-f'_{px}(x^*),\ x'_{y_j}(\alpha^*,\lambda^*;x^*))-$$
$$-(f'_{iy_j}(x^*)-f'_{py_j}(x^*))=0,\quad i\in 1:p-1, \tag{5}$$

$$h'_{ly_j}(x^*)+(h'_{ly_j}(x^*),x'_{y_j}(\alpha^*,\lambda^*;x^*))-h'_{ly_j}(x^*)=0,\quad j\in 1:p_1,$$

где $\tau^{*\prime}_{iy_j}=\tau'_{iy_j}(\alpha^*,\lambda^*;x^*)$, $\mu^{*\prime}_{ky_j}=\mu'_{ky_j}(\alpha^*,\lambda^*;x^*)$.

Неизвестными в системе (5) являются $\tau^{*\prime}_{iy_j}$, $\mu^{*\prime}_{ky_j}$, $x'_{y_j}(\alpha^*,\lambda^*;x^*)$. В [33] показано, что якобиан системы (5) не равен нулю. Поэтому $x'_{y_j}(\alpha^*,\lambda^*;x^*)=0$.

Лемма 4. Для любых $\alpha_i\geq 0$, $\sum_{i=1}^{p}\alpha_i>0$, $\lambda_j>0$, $j\in 1:p_1$, точка x^* является неподвижной точкой оператора $A_{\alpha,\lambda}(x)$, т.е. $A_{\alpha,\lambda}(x^*)=x^*$.

Доказательство. Допустим противное, т.е. пусть существует точка $\bar{x}\in B(x^*)$ такая, что

$$F(\alpha,\lambda;\bar{x})<F(\alpha,\lambda;x^*). \tag{6}$$

Пусть $g=(\bar{x}-x^*)/\|\bar{x}-x^*\|$. Из определения множества $B(y)$ получаем

$$(f'_{ix}(x^*)-f'_{px}(x^*),\ g)=0,\quad i\in 1:p-1,$$
$$(h'_{jx}(x^*),\ g)=0,\quad j\in 1:p_1. \tag{7}$$

Из (6) и (7) следует, что существует число $\alpha>0$ такое, что

$$(f'_{ix}(x^*),g)=(f'_{px}(x^*),g)=-\alpha,\quad i\in 1:p-1,$$
$$(h'_{jx}(x^*),g)=0,\quad i\in 1:p_1. \tag{8}$$

Так как $\sum_{i=1}^{p}\alpha_i^*=1$, то из (8) получаем

$$\Big(\sum_{i=1}^{p}\alpha_i^*f'_{ix}(x^*),g\Big)=(F'_x(\alpha^*,\lambda^*;x^*),g)=-\alpha<0. \tag{9}$$

Но (9) противоречит (3). Лемма доказана.

Рассмотрим следующий итерационный процесс. Даны α^0,λ^0,x^0. Пусть получена точка x^k. Тогда

$$x^{k+1}=A_0(x^k)\overset{def}{=}A_{\alpha^0,\lambda^0}(x^k).$$

Теорема. Существует $\delta>0$ такое, что если $\|x^0-x^*\|\leq\delta$, $\|\alpha^0-\alpha^*\|\leq\delta$, $\|\lambda^0-\lambda^*\|\leq\delta$, то

$$\|x^{k+1}-x^*\|\leq q_0\|x^k-x^*\|, \tag{10}$$

где $q_0<1$.

Доказательство. Обозначим через ρ_0 спектральный радиус оператора $A'_{0x}(x^*)$. Известно, что (см. [93])

$$\rho_0\leq\|A'_{0x}(x^*)\|. \tag{11}$$

Из лемм 2, 3 и неравенства (11) вытекает, что для достаточно малого δ отображение $A_0(x)$ является оператором сжатия на шаре $\|x^0-x^*\|\leq\delta$ при $\|\alpha^0-\alpha^*\|\leq\delta$ и $\|\lambda^0-\lambda^*\|\leq\delta$. Тогда из принципа сжатых отображений и леммы 4 следует (10) (см. [93, с. 28-31]). Теорема доказана.

Заметим, что q_0 зависит только от начального приближения и параметров α^0 и λ^0. Для ускорения сходимости можно уточнять

α^{k+1} и λ^{k+1} из следующей системы уравнений:

$$\left(F'_x(\alpha^{k+1}, \lambda^{k+1}; x^{k+1}),\ f'_{ix}(x^{k+1}) - f'_{px}(x^{k+1})\right) = 0,$$

$$i \in 1:p-1,\quad \sum_{i=1}^{p} \alpha_i = 1, \tag{12}$$

$$\left(F'_x(\alpha^{k+1}, \lambda^{k+1}; x^{k+1}),\ h'_{jx}(x^{k+1})\right) = 0,\quad j \in 1:p_1.$$

Из леммы 1 следует, что при x^{k+1}, достаточно близких к x^*, система (12) имеет единственное решение. Нетрудно показать, что при $x^k \to x^*$ параметры α^k и λ^k, определяемые системой (12), стремятся к α^* и λ^*.

Теперь можно описать еще один алгоритм. Даны x^0, α^0, λ^0, достаточно близкие к x^*, α^*, λ^*. Пусть уже получены x^k, α^k, λ^k. Положим $x^{k+1} = A_k(x^k) \stackrel{def}{=} A_{\alpha^k \lambda^k}(x^k)$. Далее α^{k+1} и λ^{k+1} находим из системы (12). Можно показать, что скорость сходимости этого алгоритма сверхлинейная, т.е.

$$\|x^k - x^*\| \leqslant c \prod_{i=1}^{k} \tau_i,\quad c = \text{const} > 0,\quad \tau_i \to 0.$$

Замечание 1. Для использования предложенного метода нужно знать множества $\overline{R} \subset R(x^*)$ и $\overline{Q} \subset Q(x^*)$ индексов положительных коэффициентов в минимальном представлении (3) и иметь достаточно хорошее начальное приближение. При этом получение точки, достаточно близкой к точке минимума, не представляет принципиальной трудности, так как для этого можно применить любой метод первого порядка (см., например, [39, гл. 5]). Введем множества $R_\varepsilon(x) = \{i \in I \mid \varphi(x) - f_i(x) \leqslant \varepsilon\}$, $\varepsilon > 0$, $Q_\mu(x) = \{j \in J \mid -\mu \leqslant h_j(x) \leqslant 0\}$, $\mu \geqslant 0$, $x \in \Omega$. Легко видеть, что существуют числа $\varepsilon > 0, \mu > 0, \delta > 0$ такие, что при $\|x - x^*\| \leqslant \delta$ оказывается $R_\varepsilon(x) = R(x^*)$ и $Q_\mu(x) = Q(x^*)$. Таким образом, зная множества $R_\varepsilon(x)$ и $Q_\mu(x)$ для x, достаточно близких к точке x^*, при соответствующих ε и μ узнаем множества $R(x^*)$ и $Q(x^*)$. Тогда необходимую информацию о множествах можно получить при решении задачи $\min \|z - v\|$, $z \in L_\varepsilon(x)$, $v \in \Gamma^+_\mu(x)$, где $L_\varepsilon(x) = \operatorname{co} H_\varepsilon(x)$, $H_\varepsilon(x) = \{z = f'_{ix}(x) \mid i \in R_\varepsilon(x)\}$, $\Gamma^+_\mu(x) = \{v \mid v = -\sum_{j \in Q_\mu(x)} \lambda_j h'_{jx}(x),\ \lambda_j \geqslant 0\}$.

Замечание 2. Описанный алгоритм требует на каждом шаге решения задачи минимизации выпуклой функции при линейных ограничениях-равенствах. Так как градиенты ограничений линейно-независимы, то эту задачу можно свести к задаче безусловной оптимизации, пользуясь методом, предложенным в [98].

Замечание 3. Задача выпуклого программирования при наличии хорошего начального приближения может быть сведена к последовательности задач безусловной оптимизации. При этом сложность задач от шага к шагу не возрастает, чем предложенный метод выгодно отличается от метода штрафных функций. При реализации метода можно использовать такие эффективные методы безусловной оптимизации, как метод сопряженных градиентов, методы с переменной метрикой, методы двойственных направлений. Предложенный метод рекомендуется использовать, когда методы первого порядка сходятся медленно, а применение метода Ньютона вызывает значительные трудности из-за необходимости вычислять матрицу вторых производных функции Лагранжа.

Замечание 4. Полученный результат можно обобщить на решение задачи минимизации непрерывно дифференцируемой функции при нелинейных ограничениях-равенствах.

Замечание 5. В [175] предложен алгоритм решения задачи нелинейного программирования, близкий по идее к изложенному в настоящем параграфе алгоритму. Уточнение множителей Лагранжа в [175] производится на каждом шаге, что обеспечивает квадратичную сходимость.

§ 1.10. Совместное приближение

1°. Рассматривается задача

$$\max_{k \in 1:m} \max_{t_k \in D_k} w_k(t_k) \left| \sum_{j=0}^{n} x_j u_{kj}(t_k) \right| \to \min,$$

$$\sum_{j=0}^{n} c_{ij} x_j \leqslant b_i,\quad i \in 1:l, \tag{1}$$

где $w_k, u_{k0}, u_{k1}, \ldots, u_{kn}$ – непрерывные на компакте D_k функции, причем $w_k(t_k) > 0$ $\forall t_k \in D_k$. К задаче (1) сводятся задача совместного приближения функции и ее производных [143, 144, 157, 158], задача совместного приближения нескольких функций одним полиномом (см., например, [22, с. 246-251]) и задача совместного приближения вещественной части функции комплексного переменного вещественной частью полинома на одной дуге и мнимой части функции мнимой частью того же полинома на другой.

В этом параграфе доказана разрешимость задачи (1) в предполо-

жении непротиворечивости ограничений, установлена сходимость сеточного метода и получен критерий оптимальности.

2°. Обозначим через C матрицу ограничений задачи (1), а через E – единичную матрицу. Положим $A=[C,-C,E]$. Нетрудно проверить, что (1) равносильно аналогичной задаче с ограничениями в канонической форме записи:

$$\max_{k\in 1:m}\ \max_{t_k\in D_k} w_k(t_k)\left|\Sigma_{j=0}^{s} x_j u_{kj}(t_k)\right| \to \min, \qquad (2)$$

$$AX=B,\quad X\geq \mathbf{0},$$

где $s=2n+2+l$, $X=(x_0,x_1,\dots,x_s)$, $B=(b_1,\dots,b_l)$, $u_{k,n+1+j}=-u_{kj}$ $\forall j\in 0:n+1$, $u_{k,2n+2+j}=0$ $\forall j\in 1:l$.

Введем обозначение $D=D_1\times D_2\times\dots\times D_m$ и определим на D векторный полином $P(X,t)=\Sigma_{j=0}^{s} x_j U_j(t)$, где $U_j(t)=(u_{1j}(t_1),\dots,u_{mj}(t_m))$. Пусть

$$\|P(X,\cdot)\|_D=\max_{k\in 1:m}\ \max_{t_k\in D_k} w_k(t_k)\left|\Sigma_{j=0}^{s} x_j u_{kj}(t_k)\right|.$$

Тогда (2) можно переписать в виде

$$\|P(X,\cdot)\|_D\to\min, \qquad (3)$$

$$AX=B,\quad X\geq\mathbf{0}.$$

Множество векторов X, удовлетворяющих ограничениям задачи (3), обозначим через Ω. Полином $P(X,t)$ назовем допустимым, если $X\in\Omega$, и наименее уклоняющимся от нуля, если он является решением задачи (3). Через $N_+(X)=\{j\in 0:s \mid x_j>0\}$ обозначим базис индексов вектора X, а через A_j – j-й столбец матрицы A.

Определение. Допустимый полином $P(X,t)$ будем называть базисным, если система

$$\Sigma_{j\in N_+(X)} z_j U_j(t)=\mathbf{0}\quad \forall t\in D, \qquad (4)$$

$$\Sigma_{j\in N_+(X)} z_j A_j=\mathbf{0}$$

имеет только нулевое решение. В случае $\mathbf{0}\in\Omega$ полином $P(\mathbf{0},t)$ по определению считается базисным.

Лемма. Для любого допустимого полинома существует тождественно равный ему базисный полином.

Доказательство. Возьмем допустимый полином $P(X,t)$. Если $X=\mathbf{0}$ или система (4) имеет только нулевое решение, то утверждение тривиально. Поэтому допустим, что $N_+(X)\neq\emptyset$ и система (4) имеет ненулевое решение $\{z_j\}$, $j\in N_+(X)$, причем можно считать, что хотя бы одно z_j положительно. Введем вектор $X_1=(x_0^{(1)},x_1^{(1)},\dots,x_s^{(1)})$:

$$x_j^{(1)}=\begin{cases} x_j-\varepsilon z_j, & j\in N_+(X),\\ 0, & j\notin N_+(X),\end{cases}$$

где $\varepsilon=\min\limits_{\{j\in N_+(X)\mid z_j>0\}} x_j/z_j$. Нетрудно проверить, что $X_1\in\Omega$, $P(X_1,t)\equiv P(X,t)$ и вектор X_1 имеет хотя бы на одну положительную компоненту меньше, чем X.

К X_1 можно применить те же рассуждения, что и к X. Повторив эту процедуру конечное число раз, придем к базисному полиному, тождественно равному $P(X,t)$. Лемма доказана.

Теорема 1. Если $\Omega\neq\emptyset$, то существует базисный полином, наименее уклоняющийся от нуля.

Доказательство. Возьмем минимизирующую последовательность полиномов $\{P(X_r,t)\}$. На основании леммы можно считать, что все они базисные. Более того, перейдя, если нужно, к подпоследовательности, добьемся того, что базисы индексов всех X_r совпадут: $N_+(X_r)=N_+$ $\forall r=1,2,\dots$ Учитывая соотношения

$$\|\Sigma_{j\in N_+} x_j^{(r)} U_j\|\leq M,$$

$$\Sigma_{j\in N_+} x_j^{(r)} A_j=B,\quad x_j^{(r)}>0\quad \forall j\in N_+,$$

и то, что система

$$\Sigma_{j\in N_+} z_j U_j(t)=\mathbf{0}\quad \forall t\in D,$$

$$\Sigma_{j\in N_+} z_j A_j=\mathbf{0}$$

имеет только нулевое решение, с помощью обычных рассуждений заключаем, что все последовательности $\{x_j^{(r)}\}$, $j\in N_+$, ограничены. Значит, из последовательности $\{X_r\}$ можно выделить сходящуюся подпоследовательность $X_{r_i}\to X^*$. Очевидно, что $P(X^*,t)$ – базисный полином, наименее уклоняющийся от нуля. Теорема доказана.

Замечание. Этот результат является, по-видимому, новым и в случае $m=1$ (см. обзорную статью [153]).

3°. Допустим, что при всех $k\in 1:m$ в D_k заданы последовательности компактных подмножеств $\{D_k^{(r)}\}$, сходящиеся к D_k в метрике Хаусдорфа ([113]): $D_k^{(r)} \underset{r\to\infty}{\longrightarrow} D_k$. Положим $D^{(r)} = D_1^{(r)} \times D_2^{(r)} \times \ldots \times D_m^{(r)}$ и обозначим через X_r вектор коэффициентов базисного полинома, наименее уклоняющегося от нуля на $D^{(r)}$.

Теорема 2. Из $\{X_r\}$ можно выделить сходящуюся подпоследовательность $X_{r_i} \to X^*$. При этом $P(X^*,t)$ будет базисным полиномом, наименее уклоняющимся от нуля на D.

Доказательство. Если в $\{X_r\}$ есть нулевая подпоследовательность, то утверждение тривиально. Поэтому можно считать, что $X_r \neq 0$ и, более того, $N_+(X_r) = N_+$ для всех $r = 1,2,\ldots$ По определению система

$$\Sigma_{j\in N_+} z_j U_j(t) = 0 \quad \forall t\in D^{(r)},$$

$$\Sigma_{j\in N_+} z_j A_j = 0$$

при всех $r = 1,2,\ldots$ имеет только нулевое решение. Значит, только нулевое решение имеет и следующая система:

$$\Sigma_{j\in N_+} z_j U_j(t) = 0 \quad \forall t\in D,$$

$$\Sigma_{j\in N_+} z_j A_j = 0\,. \tag{5}$$

Учитывая соотношения

$$\left\| \Sigma_{j\in N_+} x_j^{(r)} U_j \right\|_{D^{(r)}} \le \rho(D),$$

$$\Sigma_{j\in N_+} x_j^{(r)} A_j = B, \quad x_j^{(r)} > 0 \quad \forall j\in N_+,$$

где $\rho(D)$ – величина наименьшего уклонения от нуля на D, и то, что система (5) имеет только нулевое решение, заключаем, что все последовательности $\{x_j^{(r)}\}$, $j\in N_+$, ограничены. Значит, из $\{X_r\}$ можно выделить сходящуюся подпоследовательность $X_{r_i} \to X^*$. Очевидно, что $P(X^*,t)$ – базисный полином на D. Осталось показать, что $P(X^*,t)$ наименее уклоняется от нуля на D. Имеем

$$\rho(D) \ge \left\| P(X_{r_i},\cdot) \right\|_{D^{(r_i)}} \ge \left\| P(X^*,\cdot) \right\|_{D^{(r_i)}} - \left\| P(X_{r_i},\cdot) - P(X^*,\cdot) \right\|_D. \tag{6}$$

Переходя в (6) к пределу при $r_i \to \infty$, получаем $\rho(D) \ge \|P(X^*,\cdot)\|_D$. Вместе с тем по определению $\rho(D) \le \|P(X^*,\cdot)\|_D$, откуда и следует, что $P(X^*,t)$ наименее уклоняется от нуля на D. Теорема доказана.

Замечание. Если все $D_k^{(r)}$ – дискретные множества, то построение полинома $P(X_r,t)$, наименее уклоняющегося от нуля на $D^{(r)}$, сводится к задаче линейного программирования. Поэтому $P(X_r,t)$ может быть найден точно.

4°. Вернемся к векторной задаче (I) и покажем, что ее можно свести к скалярной задаче.

Положим $G = \{Y = (k,t_k) \mid k\in 1:m,\ t_k\in D_k\}$ и определим на G скалярные функции $w(Y)$, $u_j(Y)$ следующим образом: если $Y = (k,t_k)$, то $w(Y) = w_k(t_k)$ и $u_j(Y) = u_{kj}(t_k)$, $j\in 0:n$. Очевидно, что G – компакт и функции $w, u_0, u_1, \ldots, u_n$ непрерывны на G. Кроме того, (I) равносильно обычной линейной задаче аппроксимации

$$\max_{Y\in G} w(Y) \left| \Sigma_{j=0}^n x_j u_j(Y) \right| \to \min,$$

$$\Sigma_{j=0}^n c_{ij} x_j \le b_i, \quad i\in 1:l.$$

Это дает возможность переформулировать все известные результаты, относящиеся к линейным задачам аппроксимации, для задач совместного приближения нескольких функций. Остановимся на рассмотрении только критерия оптимальности.

Введем обозначения

$$p_k(X,t_k) = \Sigma_{j=0}^n x_j u_{kj}(t_k),$$

$$Q(X) = \left\{ i\in 1:l \mid \Sigma_{j=0}^n c_{ij} x_j = b_i \right\}.$$

Пару (k,t_k), где $t_k\in D_k$, назовем экстремалью векторного полинома $P(X,t) = (p_1(X,t_1), \ldots, p_m(X,t_m))$, если

$$w_k(t_k) \left| p_k(X,t_k) \right| = \max_{k\in 1:m} \max_{t_k\in D_k} w_k(t_k) \left| p_k(X,t_k) \right|.$$

Справедлива (см. § 1.1) следующая теорема.

Теорема 3. Предположим, что ограничения задачи (I) непротиворечивы и величина наименьшего уклонения от нуля положительна. Для того чтобы допустимый полином $P(X^*,t)$ был решением задачи (I), необходимо и достаточно, чтобы нашлись его экстремали (k_r, t_{k_r}), $r\in 0:p$, индексы $i_r\in Q(X^*)$, $r\in 1:q$, и положительные числа $\alpha_0, \alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_q$, где $p\ge 0$, $q\ge 0$, $p+q\le n+1$, такие, что

$$\sum_{r=0}^{p}\alpha_r\xi_r u_{k_r j}(t_{k_r})+\sum_{r=1}^{q}\lambda_r c_{i_r j}=0 \quad \forall j\in 0:n.$$

Здесь $\xi_r=\operatorname{sign} p_{k_r}(X^*,t_{k_r})$.

Теореме 3 можно придать альтернансную форму. Предварительно введем определение.

Определение. Допустимый векторный полином $P(X^*,t)$ обладает r-точечным альтернансом, $r\in 2:n+2$, если существуют его экстремали $Y_r=(k_r,t_{k_r})$, $r\in 0:p$, и индексы $i_1,\dots,i_q$ из $Q(X^*)$, $p\geqslant 0$, $q\geqslant 0$, $p+q+1=r$, такие, что:

I. Любой минор порядка r матрицы

$$B=\begin{pmatrix} u_{k_0 0}(t_{k_0}) & u_{k_1 0}(t_{k_1}) & \dots & u_{k_p 0}(t_{k_p}) & c_{i_1 0} & \dots & c_{i_q 0} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ u_{k_0 n}(t_{k_0}) & u_{k_1 n}(t_{k_1}) & \dots & u_{k_p n}(t_{k_p}) & c_{i_1 n} & \dots & c_{i_q n} \\ p_{k_0}(X^*,t_{k_0}) & p_{k_1}(X^*,t_{k_1}) & \dots & p_{k_p}(X^*,t_{k_p}) & 0 & \dots & 0 \end{pmatrix},$$

не содержащий ее последнюю строку, равен нулю.

II. Существует ненулевой минор Δ порядка r матрицы B такой, что все определители Δ_i, получаемые из Δ вычеркиванием i-го столбца и последней строки, отличны от нуля, а величины $\xi_i\Delta_i$, где

$$\xi_i=\begin{cases} \operatorname{sign} p_{k_i}(X^*,t_{k_i}), & i\in 0:p, \\ 1, & i\in p+1:p+q, \end{cases}$$

имеют чередующиеся знаки.

Теорема 4. В условиях теоремы 3 допустимый полином $P(X^*,t)$ является решением задачи (I) тогда и только тогда, когда он обладает r-точечным альтернансом, где $r\in 2:n+2$.

5°. В заключение остановимся на задаче совместного приближения функции и ее производных.

Пусть $f(t)$ - вещественная функция, $k+1$ раз дифференцируемая на $[\alpha,\beta]$. Положим

$$\|f\|=\max_{t\in[\alpha,\beta]}|f(t)|.$$

Среди всех алгебраических полиномов $p(A,t)$ степени не выше n будем искать тот, который минимизирует величину $\max\limits_{i\in 0:k}\|p^{(i)}-f^{(i)}\|$.

Если ввести множество $G=\{Y=(i,t)\mid i\in 0:k,\ t\in[\alpha,\beta]\}$ и для $Y=(i,t)$ положить $F(A,Y)=p^{(i)}(A,t)-f^{(i)}(t)$, то задача сводится к линейной задаче аппроксимации

$$\max_{Y\in G}|F(A,Y)|\to\min_{A\in E_{n+1}}.$$

Справедлива следующая теорема единственности [143].

Теорема 5. Пусть $p(A_1,t)$ и $p(A_2,t)$ - полиномы наилучшего совместного приближения функции $f(t)$ и всех ее производных до k-го порядка включительно. Тогда

$$p^{(k)}(A_1,t)\equiv p^{(k)}(A_2,t).$$

При $k>0$ наличие $(k+1)$-й производной у $f(t)$ всюду на $[\alpha,\beta]$ существенно для справедливости этой теоремы.

Пример [131]. Пусть $[\alpha,\beta]=[-4.25,\ 4.25]$, $k=1$, $n=2$ и

$$f(t)=\begin{cases} (\operatorname{sign} t)\,t^2/2, & \text{если } 0\leqslant|t|\leqslant 1.5, \\ (\operatorname{sign} t)(1.5|t|-1.125), & \text{если } 1.5\leqslant|t|\leqslant 4.25. \end{cases}$$

Очевидно, что

$$f'(t)=\begin{cases} |t|, & \text{если } 0\leqslant|t|\leqslant 1.5, \\ 1.5, & \text{если } 1.5\leqslant|t|\leqslant 4.25. \end{cases}$$

Таким образом, в точках $t=0$, $t=\pm 1.5$ функция $f(t)$ не имеет второй производной.

В [131] методом от противного показано, что в этом случае первая производная полинома наилучшего совместного приближения определяется не единственным образом. Приведем прямое доказательство этого факта.

Рассмотрим полином $p_*(t)=t$. Его экстремалями будут пары $Y_0=(0,-4.25)$, $Y_1=(0,4.25)$, $Y_2=(1,0)$. Составим матрицу

$$B=\begin{pmatrix} 1 & 1 & 0 \\ -4.25 & 4.25 & 1 \\ (-4.25)^2 & (4.25)^2 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Нетрудно проверить, что Y_0, Y_1, Y_2 образуют трехточечный альтернанс полинома $p_*(t)$, поэтому он является полиномом наилучшего совместного приближения функции $f(t)$ и ее первой производной. При этом $\max\limits_{i\in 0:1} \| p_*^{(i)} - f^{(i)} \| = 1$. Однако у полиномов $p_\varepsilon(t) = \frac{\varepsilon}{2}t^2 + t - \frac{\varepsilon}{2}(4.25)^2$ при малых $\varepsilon > 0$ также $\max\limits_{i\in 0:1} \| p_\varepsilon^{(i)} - f^{(i)} \| = 1$. Неединственность доказана.

З а м е ч а н и е. Наличие полного альтернанса гарантирует единственность полинома наилучшего совместного приближения. На основании теоремы 5 заключаем, что в задаче совместного приближения функции и ее производных более характерной является ситуация, когда полином наилучшего приближения обладает неполным альтернансом.

§ 1.11. Аппроксимация с интерполяцией

1°. Рассмотрим задачу

$$\varphi(X) \stackrel{def}{=} \max_{t\in D} |f(X,t)| \to \min_{X\in\Omega}, \qquad (1)$$

где $\Omega = \{X\in E_n \mid f(X,\tau_j) = 0 \ \ \forall j\in 1:m\}$. Предположим, что $D\subset[a,b]$ - компакт, содержащий не менее $n+1$ точек, $\tau_j\in[a,b]$ $\forall j\in 1:m$ и $m < n$. Допустим также, что функция $f(X,t)$ непрерывна вместе с $f'_X(X,t)$ на $\Omega'\times[a,b]$, где $\Omega'\subset E_n$ - некоторое открытое множество, содержащее Ω, и

$$\inf_{X\in\Omega} \varphi(X) > 0. \qquad (2)$$

Основным результатом этого параграфа является теорема 3 характеризации точки строгого локального минимума функции φ на Ω.

В линейном случае, когда $f(X,t) = u_0(t) - \sum_{i=1}^n x_i u_i(t)$, задача (1) рассматривалась, например, в [167], [136], а в случае варисольвентного семейства $\{f(X,\cdot)\}$ - в [124].

2°. Для произвольной точки X из Ω положим

$$R(X) = \{t\in D \mid |f(X,t)| = \varphi(X)\},$$

$$H(X) = \{z = \operatorname{sign} f(X,t) f'_X(X,t) \mid t\in R(X)\},$$

$$H_1(X) = \{f'_X(X,\tau_1), \ldots, f'_X(X,\tau_m)\}.$$

Обозначим через $L(X)$ выпуклую оболочку множества $H(X)$, а через $K(X)$ - линейную оболочку множества $H_1(X)$. Очевидно, что $L(X)$ - выпуклый компакт.

Лемма 1. Если $X^*\in\Omega$ - точка локального минимума функции φ на Ω и векторы

$$f'_X(X^*,\tau_1), \ldots, f'_X(X^*,\tau_m) \qquad (3)$$

линейно-независимы, то

$$L(X^*)\cap K(X^*) \neq \emptyset. \qquad (4)$$

Д о к а з а т е л ь с т в о. Допустим противное, т.е. $L(X^*)\cap K(X^*) = \emptyset$. По теореме о строгой отделимости найдется вектор $g_0\in E_n$, $\|g_0\| = 1$, такой, что

$$\begin{aligned} &(g_0, V) = 0 \quad \forall V\in K(X^*), \\ &(g_0, z) < 0 \quad \forall z\in L(X^*). \end{aligned} \qquad (5)$$

В частности,

$$(f'_X(X^*,\tau_j), g_0) = 0 \quad \forall j\in 1:m. \qquad (6)$$

Вектор g_0, удовлетворяющий (6), при условии линейной независимости векторов (3) обладает следующим свойством: существует гладкая дуга $V = V(\lambda)$ такая, что $V(0) = X^*$, $V'(0) = g_0$ и $V(\lambda)$ принадлежит Ω при малых $\lambda > 0$ (см., например, [79, с. 37-38]). Пользуясь теперь неравенством (5), покажем, что при достаточно малых $\lambda > 0$

$$\varphi(V(\lambda)) < \varphi(X^*). \qquad (7)$$

Действительно, $V(\lambda) = X^* + \lambda(g_0 + \eta(\lambda))$, где $\eta(\lambda)\underset{\lambda\to+0}{\to} \mathbf{0}$. Далее на основании (5) при малых $\lambda > 0$

$$\max_{z\in H(X^*)} (z, g_0 + \eta(\lambda)) < 0. \qquad (8)$$

Воспользуемся разложением функции φ ([39, с. 237]):

$$\varphi(V(\lambda)) = \varphi(X^*) + \lambda \max_{z \in H(X^*)} \left(z, g_0 + \eta(\lambda) \right) + o\left(\lambda \| g_0 + \eta(\lambda) \| ; \frac{g_0 + \eta(\lambda)}{\| g_0 + \eta(\lambda) \|} \right),$$

где $o(\lambda; g)/\lambda \xrightarrow[\lambda \to +0]{} 0$ равномерно по g, $\|g\| = 1$. Отсюда с учетом (8) получаем (7). Но (7) противоречит тому, что X^* - точка локального минимума функции φ на Ω. Лемма доказана.

3°. Напомним определение и некоторые свойства чебышевских систем функций.

Определение. Пусть

$$u_1(t), \ldots, u_n(t) \qquad (*)$$

- система непрерывных на $[a,b]$ функций и $U(t) = (u_1(t), \ldots, u_n(t))$. Систему (*) (и вектор-функцию $U(t)$) называют чебышевской на $[a,b]$, если для любых попарно различных $t_1, \ldots, t_n$ из $[a,b]$ векторы $U(t_1), \ldots, U(t_n)$ линейно-независимы.

В этом параграфе потребуются следующие свойства чебышевских систем:

I. Для любого набора точек $t_1 < \ldots < t_n$ из $[a,b]$ определители, составленные из столбцов $U(t_1), \ldots, U(t_n)$, имеют одинаковые знаки (см., например, [60, с. 52]).

II. Пусть $t_0 < t_1 < \ldots < t_n$ - произвольный набор точек из $[a, b]$. Тогда система линейных уравнений $\sum_{i=0}^{n} \lambda_i U(t_i) = 0$ имеет единственное с точностью до множителя ненулевое решение $\lambda_0^*, \lambda_1^*, \ldots, \lambda_n^*$, при этом $\operatorname{sign} \lambda_i^* = -\operatorname{sign} \lambda_{i-1}^*$ для всех $i \in 1:n$ (доказательство непосредственно следует из свойства I и формул Крамера).

Теорема 1. Пусть $X^* \in \Omega$ - точка локального минимума функции φ на Ω и система функций

$$u_k(t) = f'_{x_k}(X^*, t), \quad k \in 1:n, \qquad (9)$$

является чебышевской на $[a, b]$. Тогда найдутся точки $t_0 < t_1 < \ldots < t_{n-m}$ из $R(X^*)$ такие, что для всех $i \in 1: n-m$

$$f(X^*, t_i) = (-1)^{s_i + 1} f(X^*, t_{i-1}), \qquad (10)$$

где s_i - число узлов интерполяции τ_j, лежащих в интервале (t_{i-1}, t_i).

Доказательство. Введем обозначение

$$\xi(t) = \begin{cases} \operatorname{sign} f(X^*, t), & \text{если } t \in R(X^*), \\ 1, & \text{если } t = \tau_j,\ j \in 1:m. \end{cases}$$

Из чебышевских свойств системы (9) следует линейная независимость векторов (3), поэтому справедливо заключение леммы I. На основании (4) существуют точки $t_0 < t_1 < \ldots < t_p$ из $R(X^*)$ и коэффициенты $\alpha_0, \alpha_1, \ldots, \alpha_p, \gamma_1, \ldots, \gamma_m$,

$$\alpha_i > 0 \ \ \forall i \in 0:p, \quad \sum_{i=0}^{p} \alpha_i = 1, \qquad (11)$$

для которых

$$\sum_{i=0}^{p} \alpha_i \xi(t_i) f'_X(X^*, t_i) + \sum_{j=1}^{m} \gamma_j f'_X(X^*, \tau_j) = 0. \qquad (12)$$

Заметим, что в силу (2) узлы интерполяции не принадлежат $R(X^*)$. Значит, все точки t_i, τ_j из (12) попарно различны. Учитывая чебышевские свойства системы функций (9), заключаем, что в левой части (12) не менее $n+1$ слагаемых, т.е. $p + m \geqslant n$. Покажем, что существует представление нуля, в котором $p + m = n$.

Пусть $p + m > n$. Тогда однородная система линейных уравнений

$$\sum_{i=0}^{p} \beta_i \xi(t_i) f'_X(X^*, t_i) + \sum_{j=1}^{m} \beta_{p+j} f'_X(X^*, \tau_j) = 0,$$

$$\sum_{i=0}^{p} \beta_i = 0$$

имеет ненулевое решение $\{\beta_i^*\}$, $i \in 0: p+m$. В силу линейной независимости векторов (3) $\sum_{i=0}^{p} |\beta_i^*| > 0$. При этом можно считать, что среди чисел $\beta_0^*, \beta_1^*, \ldots, \beta_p^*$ есть хотя бы одно положительное. Введем величины

$$\varepsilon = \min \left\{ \frac{\alpha_i}{\beta_i^*} \,\middle|\, \beta_i^* > 0,\ i \in 0:p \right\},$$

$$\alpha_i^* = \alpha_i - \varepsilon \beta_i^*, \ i \in 0:p; \qquad \gamma_j^* = \gamma_j - \varepsilon \beta_{p+j}^*, \ j \in 1:m.$$

Очевидно, что

$$\alpha_i^* \geqslant 0 \quad \forall i \in 0:p, \quad \sum_{i=0}^{p} \alpha_i^* = 1, \qquad (13)$$

$$\sum_{i=0}^{p} \alpha_i^* \xi(t_i) f'_X(X^*, t_i) + \sum_{j=1}^{m} \gamma_j^* f'_X(X^*, \tau_j) = 0 \qquad (14)$$

и хотя бы один из коэффициентов α_i^* обращается в нуль. Оставляя в (13), (14) только ненулевые α_i^*, получаем представление нуля, в котором число слагаемых меньше $p+m$. Повторив эту процедуру конечное число раз, придем к соотношениям вида (11), (12) с $p = n-m$. При этом автоматически все γ_j будут отличными от нуля.

Расположим точки t_i, τ_j в порядке возрастания, обозначив их $w_0 < w_1 < \ldots < w_n$. Тогда (12) можно переписать в виде

$$\Sigma_{k=0}^{n} \lambda_k \xi(w_k) f_X'(X^*, w_k) = \mathbf{0},$$

где $\lambda_k > 0$, если $w_k \in R(X^*)$. На основании свойства II чебышевских систем величины $\lambda_k \xi(w_k)$ имеют чередующиеся знаки. Отсюда очевидным образом следует (10). Теорема доказана.

4°. Положим $S = \{ g \in E_n \mid \|g\| = 1 \}$,

$$C(X) = \{ g \in S \mid (f_X'(X, \tau_j), g) = 0 \quad \forall j \in 1:m \}.$$

Лемма 2. Если в точке X^* из Ω

$$\min_{g \in C(X^*)} \max_{z \in H(X^*)} (z, g) > 0, \tag{15}$$

то X^* является точкой строгого локального минимума функции φ на Ω. Более того, найдутся $\delta > 0$ и $r > 0$ такие, что для всех $X \in \Omega$, $\|X - X^*\| \leqslant \delta$,

$$\varphi(X) \geqslant \varphi(X^*) + r\|X - X^*\|. \tag{16}$$

Доказательство. Введем обозначение

$$C_\varepsilon(X^*) = \{ g \in S \mid |(f_X'(X^*, \tau_j), g)| \leqslant \varepsilon \quad \forall j \in 1:m \}.$$

Нетрудно показать, опираясь на (15), что существует $\varepsilon^* > 0$, для которого

$$M^* \stackrel{\text{def}}{=} \min_{g \in C_{\varepsilon^*}(X^*)} \max_{z \in H(X^*)} (z, g) > 0.$$

Возьмем теперь некоторое $\delta > 0$ и рассмотрим множество $\Omega_\delta = \{X \in \Omega \mid \|X - X^*\| \leqslant \delta\}$. Пусть $X \in \Omega_\delta$, $X \neq X^*$. Тогда $X = X^* + \lambda g$, где $\lambda = \|X - X^*\|$, $g = (X - X^*)/\lambda$. Справедливы разложения

$$0 = f(X, \tau_j) - f(X^*, \tau_j) = \lambda (f_X'(X^*, \tau_j), g) + o_j(\lambda; g) \quad \forall j \in 1:m,$$

$$\varphi(X) - \varphi(X^*) = \lambda \max_{z \in H(X^*)} (z, g) + o_0(\lambda, g),$$

где $o_j(\lambda; g)/\lambda \underset{\lambda \to +0}{\longrightarrow} 0$ равномерно по $g \in S$ и $j \in 0:m$. Значит, найдется $\delta > 0$ такое, что для всех $\lambda \in (0, \delta]$ и $g \in S$

$$\max_{j \in 0:m} |o_j(\lambda; g)|/\lambda \leqslant r^* \stackrel{\text{def}}{=} \min\{\varepsilon^*, M^*/2\}.$$

Отсюда следует, что в случае $X \in \Omega_\delta$, $X \neq X^*$ вектор $g = (X - X^*)/\|X - X^*\|$ принадлежит $C_{\varepsilon^*}(X^*)$ и

$$\varphi(X) - \varphi(X^*) \geqslant \|X - X^*\| (M^* - r^*).$$

Положив $r = M^* - r^*$, получим (16). Лемма доказана.

Теорема 2. Пусть $X^* \in \Omega$. Если существуют $t_0 < t_1 < \ldots < t_{n-m}$ из $R(X^*)$ такие, что определители Δ_i, составленные из столбцов

$$f_X'(X^*; t_0), \ldots, f_X'(X^*; t_{i-1}), f_X'(X^*; t_{i+1}), \ldots, f_X'(X^*; t_{n-m}), f_X'(X^*; \tau_1), \ldots, f_X'(X^*; \tau_m),$$

для всех $i \in 0:n-m$ отличны от нуля и

$$f(X^*, t_i) \operatorname{sign} \Delta_i = -f(X^*, t_{i-1}) \operatorname{sign} \Delta_{i-1} \quad \forall i \in 1:n-m, \tag{17}$$

то выполняется неравенство (15) и, значит, справедливо заключение леммы 2.

Доказательство. Возьмем $g \in C(X^*)$ и положим

$$v = \max_{i \in 0:n-m} (\xi(t_i) f_X'(X^*; t_i), g).$$

Достаточно показать, что $v > 0$. Рассмотрим систему линейных уравнений относительно α_i:

$$\Sigma_{i=0}^{n-m} \alpha_i \xi(t_i) f_X'(X^*; t_i) + \Sigma_{j=1}^{m} \alpha_{n-m+j} f_X'(X^*; \tau_j) = \mathbf{0}, \quad \alpha_0 = 1. \tag{18}$$

Определитель этой системы равен $(-1)^n \xi(t_0) \prod_{i=0}^{n-m} \xi(t_i) \Delta_0 \neq 0$. Значит, система имеет единственное решение. По формулам Крамера

$$\alpha_i = (-1)^i \frac{\Delta_i \xi(t_i)}{\Delta_0 \xi(t_0)} \qquad \forall i \in 0:n-m.$$

Согласно (17) коэффициенты $\alpha_0, \alpha_1, \ldots, \alpha_{n-m}$ положительны. На основании (18) для $g \in C(X^*)$ получаем

$$\Sigma_{i=0}^{n-m} \alpha_i \left(\xi(t_i) f'_X(X^*, t_i), g \right) = 0 ,$$

откуда следует неравенство $v \geqslant 0$. Допустим, что $v = 0$. Тогда g удовлетворяет системе линейных однородных уравнений

$$(g, f'_X(X^*, t_i)) = 0 , \quad i \in 1:n-m ,$$

$$(g, f'_X(X^*, \tau_j)) = 0 , \quad j \in 1:m ,$$

определитель которой отличен от нуля. Но это противоречит условию $\| g \| = 1$. Значит, $v > 0$, что и требовалось доказать.

5°. Теперь без труда может быть установлена основная в этом параграфе теорема.

Теорема 3. Пусть в точке $X^* \in \Omega$ система функций (9) является чебышевской на $[\alpha, b]$. Для того чтобы X^* была точкой строгого локального минимума функции φ на Ω, необходимо и достаточно, чтобы существовали точки $t_0 < t_1 < \ldots < t_{n-m}$ из $R(X^*)$, в которых

$$f(X^*, t_i) = (-1)^{s_i+1} f(X^*, t_{i-1}) \quad \forall i \in 1:n-m, \qquad (19)$$

где s_i – число узлов интерполяции τ_j, лежащих в интервале (t_{i-1}, t_i).

Доказательство. Необходимость немедленно следует из теоремы 1.

Достаточность. Покажем, что выполнены условия теоремы 2. Для этого рассмотрим матрицу F, составленную из столбцов

$$f'_X(X^*, t_0), f'_X(X^*, t_1), \ldots, f'_X(X^*, t_{n-m}), f'_X(X^*, \tau_1), \ldots, f'_X(X^*, \tau_m).$$

В силу чебышевских свойств системы (9) определители Δ_i подматриц, получаемых из F исключением i-го столбца, отличны от нуля при всех $i \in 0:n$. Опираясь на свойство I чебышевских систем, нетрудно проверить, что для $i \in 1:n-m$

$$\operatorname{sign} \Delta_i = (-1)^{s_i} \operatorname{sign} \Delta_{i-1} . \qquad (20)$$

Из (19) и (20) следует (17). Остается сослаться на теорему 2.

§ 1.12. Аппроксимация с двусторонними ограничениями

1°. Допустим, что D, ω_s и ω_l – произвольные компакты в $[\alpha, \beta]$, причем D содержит не менее $n+1$ точек. Рассмотрим следующую задачу нелинейной аппроксимации непрерывной на компакте D функции $y(t)$:

$$\varphi(t) \stackrel{\text{def}}{=} \max_{t \in D} | f(X,t) - y(t) | \to \min_{X \in \Omega} ,$$

где $$\Omega = \left\{ X \in E_n \;\middle|\; \begin{array}{l} f(X,t) \leqslant s(t) \quad \forall t \in \omega_s \\ f(X,t) \geqslant l(t) \quad \forall t \in \omega_l \end{array} \right\} .$$

Будем считать, что $f(X,t)$ непрерывна вместе с $f'_X(X,t)$ на $E_n \times [\alpha, \beta]$ и

$$\inf_{X \in \Omega} \varphi(X) > 0 . \qquad (1)$$

Относительно граничных функций $s(t)$ и $l(t)$ сделаем предположение, что они непрерывны на ω_s и ω_l соответственно и

$$\begin{aligned} & l(t) < s(t) \quad \forall t \in \omega_l \cap \omega_s , \\ & l(t) \leqslant y(t) \quad \forall t \in \omega_l \cap D , \\ & s(t) \geqslant y(t) \quad \forall t \in \omega_s \cap D . \end{aligned} \qquad (2)$$

В этом параграфе доказываются теоремы характеризации точки локального минимума функции φ на Ω.

В линейном случае, когда $f(X,t) = \Sigma_{k=1}^{n} x_k u_k(t)$, задача рассматривалась в [176, 177] (см. также обзорную статью [153]). Нелинейный случай был рассмотрен в [141], где приведены результаты, касающиеся только необходимых альтернансных условий локального минимума. В работе [137] получены необходимые и достаточные альтернансные условия в случае, когда семейство аппроксимирующих функций является варисольвентным.

2°. Установим одно свойство неизолированных точек множества Ω, которым будем пользоваться в дальнейшем.

Положим $$\varphi_s(X) = \max_{t\in\omega_s} [f(X,t)-s(t)],$$

$$\varphi_l(X) = \max_{t\in\omega_l} [l(t)-f(X,t)].$$

Тогда $\Omega=\{X\in E_n \mid \varphi_s(X)\le 0, \varphi_l(X)\le 0\}$. Справедлива следующая лемма.

Лемма. Пусть X_0 - неизолированная граничная точка Ω. Тогда найдется направление $g_0\in E_n, \|g_0\|=1$, для которого

$$\partial\varphi_s(X_0)/\partial g_0 \le 0, \quad \text{если} \quad \varphi_s(X_0)=0,$$

$$\partial\varphi_l(X_0)/\partial g_0 \le 0, \quad \text{если} \quad \varphi_l(X_0)=0.$$

Доказательство. Для определенности будем считать, что $\varphi_s(X_0)=\varphi_l(X_0)=0$.

Так как X_0 неизолированная точка Ω, то найдется последовательность $X_k\in\Omega$ такая, что $X_0=\lim_{k\to\infty} X_k$. Полагая $\lambda_k=\|X_0-X_k\|$, $g_k=(X_k-X_0)/\lambda_k$, из разложений функций $\varphi_s(X_k)$ и $\varphi_l(X_k)$ для всех k имеем

$$\partial\varphi_s(X_0)/\partial g_k + o_s(\lambda_k;g_k)/\lambda_k \le 0,$$

$$\partial\varphi_l(X_0)/\partial g_k + o_l(\lambda_k;g_k)/\lambda_k \le 0,$$

где $o_s(\lambda;g)/\lambda\to 0$ и $o_l(\lambda;g)/\lambda\to 0$ при $\lambda\to+0$ равномерно по $g, \|g\|=1$. Из последовательности $\{g_k\}$ $(\|g_k\|=1)$ выбираем частичную, сходящуюся к g_0, $\|g_0\|=1$. Переходя к пределу в последних неравенствах, получаем требуемое. Лемма доказана.

3°. Введем обозначения

$$Q_s(X)=\{t\in\omega_s \mid f(X,t)=s(t)\},$$

$$Q_l(X)=\{t\in\omega_l \mid f(X,t)=l(t)\},$$

$$Q(X)=Q_s(X)\cup Q_l(X).$$

Положим далее

$$A_+(X)=Q_s(X)\cup\{t\in D \mid f(X,t)-y(t)=\varphi(X)\},$$

$$A_-(X)=Q_l(X)\cup\{t\in D \mid f(X,t)-y(t)=-\varphi(X)\},$$

$$A(X)=A_+(X)\cup A_-(X).$$

Легко проверить, что в предположениях (1), (2) множества $A_+(X)$ и $A_-(X)$ не пересекаются. Определим теперь функцию

$$\xi(X,t)=\begin{cases} 1, & \text{если } t\in A_+(X),\\ -1, & \text{если } t\in A_-(X),\end{cases}$$

и множество $\mathcal{H}(X)=\{z=\xi(X,t)f'_X(X,t) \mid t\in A(X)\}$.

Перейдем к установлению альтернансных свойств точки локального минимума.

Теорема 1. Пусть X^* - неизолированная точка локального минимума функции φ на Ω и $f'_X(X^*,t)$ - чебышевская вектор-функция на $[\alpha,\beta]$. Тогда существуют точки $t_0<t_1<\ldots<t_n$ из $A(X^*)$ (причем среди t_i есть точки из $R(X^*)$) такие, что

$$\xi(X^*,t_{i-1})=-\xi(X^*,t_i) \quad \forall i\in 1:n.$$

Доказательство. Докажем сначала, что $0\in\operatorname{co}(\mathcal{H}(X^*))$. Для этого допустим противное. Тогда по теореме о строгой отделимости существует вектор g_0, $\|g_0\|=1$, такой, что

$$(g_0,z)<0 \quad \forall z\in\mathcal{H}(X^*).$$

Отсюда следует выполнение неравенств

$$\frac{\partial\varphi(X^*)}{\partial g_0} = \max_{t\in R(X^*)} \xi(X^*,t)(f'_X(X^*,t),g_0) < 0, \qquad (3)$$

$$\frac{\partial\varphi_l(X^*)}{\partial g_0} < 0, \quad \text{если} \quad \varphi_l(X^*)=0,$$

$$\frac{\partial\varphi_s(X^*)}{\partial g_0} < 0, \quad \text{если} \quad \varphi_s(X^*)=0.$$

Поэтому при малых $\lambda>0$ имеем $\varphi_l(X^*+\lambda g_0)<0$ и $\varphi_s(X^*+\lambda g_0)<0$, т.е. $X^*+\lambda g_0\in\Omega$. В силу (3)

$$\varphi(X^*+\lambda g_0)<\varphi(X^*),$$

но это противоречит тому, что X^* – точка локального минимума φ на Ω.

Итак, $\mathbf{0} \in \operatorname{co}(\mathcal{H}(X^*))$. Это означает, что найдутся точки $t_0 < t_1 < \ldots < t_p$ из $A(X^*)$ и положительные $\alpha_0, \alpha_1, \ldots, \alpha_p$, $p \leqslant n$, такие, что

$$\Sigma_{i=0}^{p} \alpha_i \xi(X^*, t_i) f'_X(X^*, t_i) = \mathbf{0}. \qquad (4)$$

Из чебышевских свойств $f'_X(X^*, t)$ на $[\alpha, \beta]$ следует, что $p = n$ и $\alpha_i \xi(X^*, t_i)$ имеют чередующиеся знаки, откуда в силу положительности α_i

$$\xi(X^*, t_{i-1}) = -\xi(X^*, t_i) \quad \forall i \in 1:n.$$

Докажем теперь, что среди t_i есть точки из $R(X^*)$. По лемме существует ненулевой вектор $g \in E_n$, для которого

$$\xi(X^*, t_i)(f'_X(X^*, t_i), g) \leqslant 0 \quad \forall t_i \in Q(X^*).$$

Из (4) имеем

$$\Sigma_{i=0}^{n} \alpha_i \xi(X^*, t_i)(f'_X(X^*, t_i), g) = 0.$$

Если допустить, что все t_i принадлежат $Q(X^*)$, то из последнего равенства следует, что $(f'_X(X^*, t_i), g) = 0$ для всех $i \in 0:n$, но это невозможно в силу линейной независимости любых n векторов $f'_X(X^*, t_i)$. Теорема доказана.

4°. Установим теперь достаточность альтернансных условий.

Теорема 2. Пусть X^* – неизолированная точка множества Ω и вектор-функция $f'_X(X^*, t)$ является чебышевской на $[\alpha, \beta]$. Если в $A(X^*)$ существуют $t_0 < t_1 < \ldots < t_n$, для которых

$$\xi(X^*, t_{i-1}) = -\xi(X^*, t_i) \quad \forall i \in 1:n, \qquad (5)$$

то X^* есть точка строгого локального минимума функции φ на Ω. Более того, существуют положительные числа r и δ такие, что

$$\varphi(X) \geqslant \varphi(X^*) + r\|X - X^*\| \quad \forall X \in \Omega, \ \|X - X^*\| < \delta.$$

Доказательство. Рассмотрим систему линейных уравнений относительно λ_i:

$$\Sigma_{i=0}^{n} \lambda_i \xi(X^*, t_i) f'_X(X^*, t_i) = \mathbf{0}. \qquad (6)$$

Так как $f'_X(X^*, t)$ – чебышевская на $[\alpha, \beta]$ вектор-функция, то у системы (6) существует ненулевое решение, при этом

$$\lambda_i \lambda_{i-1} \xi(X^*, t_i) \xi(X^*, t_{i-1}) < 0 \quad \forall i \in 1:n.$$

Отсюда с учетом (5) следует, что система (6) имеет положительное решение, т.е. $\lambda_i > 0 \quad \forall i \in 0:n$.

Берем произвольный вектор $g \in E_n$, $\|g\| = 1$. Для него из (6), полагая $z_i = \xi(X^*, t_i) f'_X(X^*, t_i)$, получаем

$$\Sigma_{i=0}^{n} \lambda_i (z_i, g) = 0.$$

Так как любые n векторов z_i линейно-независимы, то

$$w(g) \stackrel{\text{def}}{=} \max_{i \in 0:n} (z_i, g) > 0.$$

Это означает, что для любого нормированного $g \in E_n$ хотя бы одна из производных $\partial\varphi(X^*)/\partial g$, $\partial\varphi_l(X^*)/\partial g$, $\partial\varphi_s(X^*)/\partial g$ положительна и, более того, превосходит величину r, где

$$r = \frac{1}{2} \min_{\|g\|=1} w(g).$$

Рассмотрим теперь $S_\delta(X^*)$ – окрестность точки X^*. Если $\varphi_s(X^*) < 0$ или $\varphi_l(X^*) < 0$, то при малых $\delta > 0$ для всех $X \in S_\delta(X^*)$ эти же ограничения будут "неактивными", и потому их можно исключить из рассмотрения. Для полноты картины будем считать, что $\varphi_s(X^*) = \varphi_l(X^*) = 0$.

В силу неизолированности X^* множество $\Omega \cap S_\delta(X^*)$ при любом $\delta > 0$ содержит точки, отличные от X^*. Полагая $\lambda = \|X - X^*\|$, $g = (X - X^*)/\lambda$, для всех $X \in \Omega \cap S_\delta(X^*)$ имеем

$$\varphi(X) - \varphi(X^*) = \lambda \left[\frac{\partial \varphi(X^*)}{\partial g} + \frac{o_0(\lambda; g)}{\lambda} \right],$$

$$\frac{\partial \varphi_s(X^*)}{\partial g} \leqslant \frac{|o_1(\lambda; g)|}{\lambda}, \qquad (7)$$

$$\frac{\partial \varphi_i(X^*)}{\partial g} \leqslant \frac{|o_2(\lambda; g)|}{\lambda},$$

где $|o_i(\lambda, g)| / \lambda \underset{\lambda \to +0}{\longrightarrow} 0$ равномерно по g для всех $i \in 0:2$. Легко видеть, что найдется $\delta > 0$ такое, что

$$\max_{i \in 0:2} |o_i(\lambda; g)| / \lambda \leqslant r \quad \forall \lambda \in (0, \delta).$$

Тогда для всех $X \in \Omega$, $\|X - X^*\| < \delta$, получим $\partial \varphi_s(X^*)/\partial g \leqslant r$, $\partial \varphi_i(X^*)/\partial g \leqslant r$ и в силу выбора r

$$\partial \varphi(X^*)/\partial g \geqslant 2r.$$

Теперь из (7) следует

$$\varphi(X) \geqslant \varphi(X^*) + \|X - X^*\| r,$$

что и доказывает теорему.

§ 1.13. О положительных полиномах

Во многих случаях проверка регулярности ограничений экстремальной задачи сводится к вопросу о существовании обобщенного полинома, положительного на некотором множестве. Обычно требуется найти ненулевой вектор g (вектор коэффициентов полинома) такой, что

$$(h'_X(X^*, \beta), g) > 0 \quad \forall \beta \in Q(X^*).$$

Условиям существования положительного полинома и посвящен данный параграф.

1°. Пусть Q - произвольный метрический компакт, содержащий не менее $n+1$ точек и $u_1, \ldots, u_n$ - система непрерывных на Q функций.

Определение. Говорят, что система $u_1, \ldots, u_n$ удовлетворяет условию $(\mathcal{K})$, если существует обобщенный полином $\sum_{i=1}^n a_i u_i$, положительный всюду на Q.

Введем обозначение $U = (u_1, \ldots, u_n)$.

Теорема 1. Условие $(\mathcal{K})$ выполняется тогда и только тогда,

когда для любых попарно различных точек $t_0, t_1, \ldots, t_n$ из Q несовместна система

$$\sum_{j=0}^n x_j U(t_j) = \mathbf{0},$$
$$\sum_{j=0}^n x_j = 1, \quad x_j \geqslant 0 \quad \forall j \in 0:n. \tag{1}$$

Доказательство. Необходимость. Допустим, что при некоторых a_i

$$\sum_{i=1}^n a_i u_i(t) > 0 \quad \forall t \in Q. \tag{2}$$

Тогда для произвольных $t_0, t_1, \ldots, t_n$ из Q и любых неотрицательных $x_0, x_1, \ldots, x_n$ таких, что $\sum_{j=0}^n x_j = 1$, в силу (2) имеем

$$0 < \sum_{j=0}^n x_j \sum_{i=1}^n a_i u_i(t_j) = \sum_{i=1}^n a_i \sum_{j=0}^n x_j u_i(t_j).$$

Отсюда следует, что $\sum_{j=0}^n x_j u_i(t_j) \neq 0$ хотя бы при одном $i \in 1:n$. Тем самым доказана несовместность системы (1).

Достаточность. Обозначим через $\mathcal{Z}$ выпуклую оболочку, натянутую на компактное множество $H = \{Z = U(t) \mid t \in Q\} \subset E_n$. Как известно, $\mathcal{Z}$ является выпуклым компактным множеством (простое доказательство этого факта имеется, например, в [39, с. 307]). Из несовместности системы (1) на основании теоремы Каратеодори следует, что $\mathbf{0} \notin \mathcal{Z}$. По теореме отделимости найдется вектор $A = (a_1, \ldots, a_n)$ такой, что $(A, Z) > 0$ $\forall Z \in \mathcal{Z}$. В частности, $(A, U(t)) > 0$ $\forall t \in Q$. Последнее же равносильно выполнению условия $(\mathcal{K})$. Теорема доказана.

Замечание. В случае $Q = [\alpha, \beta]$ утверждение, эквивалентное теореме 1, установлено в [1].

2°. В дальнейшем нам понадобится один достаточный признак выполнения условия $(\mathcal{K})$, который легко следует из теоремы 1. Введем обозначение

$$\Delta(z_1, \ldots, z_n) = \begin{vmatrix} u_1(z_1) & \ldots & u_1(z_n) \\ \ldots & \ldots & \ldots \\ u_n(z_1) & \ldots & u_n(z_n) \end{vmatrix}.$$

Лемма. Для выполнения условия $(\mathcal{K})$ достаточно, чтобы при любых попарно различных $t_0, t_1, \ldots, t_n$ из Q в числовой последовательности $\{(-1)^j \Delta_j\}_{j=0}^n$, где $\Delta_j = \Delta(t_0, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n)$, после выбрасывания нулевых членов имелась хотя бы одна перемена

знака.

Доказательство. Зафиксируем $t_0, t_1, \dots, t_n$ и покажем, что система (1) несовместна. Допустим, что система (1) имеет решение $x_0, x_1, \dots, x_n$. По условию найдутся индексы k, l из $0:n$, на которых

$$\operatorname{sign}\{(-1)^k \Delta_k\} = -\operatorname{sign}\{(-1)^l \Delta_l\}. \qquad (3)$$

Имеем

$$\sum_{\substack{j=0 \\ j\neq k}}^{n} x_j U(t_j) = -x_k U(t_k). \qquad (4)$$

Определитель системы (4) равен Δ_k и по условию отличен от нуля. Отсюда следует, в частности, что $x_k \neq 0$.

На основании формулы Крамера

$$x_l = (-1)^{k-l} x_k \frac{\Delta_l}{\Delta_k}, \quad \text{или} \quad \{(-1)^k \Delta_k\} x_l = \{(-1)^l \Delta_l\} x_k.$$

Учитывая (3), получаем $\operatorname{sign} x_l = -\operatorname{sign} x_k$. Но это противоречит (1). Лемма доказана.

3°. Переходим к основному результату этого параграфа, установленному И.К.Даугаветом.

Теорема 2. Пусть Q – компакт, на котором может быть задана хоть одна чебышевская система, состоящая из n непрерывных функций*. Для того чтобы по любой такой системе можно было построить положительный на Q полином, необходимо и достаточно, чтобы компакт Q состоял не более чем из n связных компонент.

Доказательство. Достаточность. Пусть $u_1, \dots, u_n$ – чебышевская система на Q. Покажем, что существует положительный по этой системе полином. Для этого возьмем произвольные попарно различные точки $t_0, t_1, \dots, t_n$ из Q и установим, что в последовательности $\{(-1)^j \Delta_j\}$ имеется хотя бы одна перемена знака.

Поскольку Q состоит не более чем из n связных компонент, то в одну такую компоненту Q_i попадают по крайней мере две точки, скажем t_{j_1} и t_{j_2}. Соединим их путем, целиком лежащим в Q_i. Если на этом пути окажутся другие точки из $\{t_j\}$, то берем две соседние. Пусть это будут t_k и t_l $(k<l)$.

* В этом случае (при $n>1$) компакт Q гомеоморфен окружности либо ее части (см., например, [60, с.51]).

Введем функцию

$$\eta(t) = \Delta(t_0, \dots, t_{k-1}, t, t_{k+1}, \dots, t_{l-1}, t_{l+1}, \dots, t_n).$$

Очевидно, что $\eta(t_k) = \Delta_l$, $\eta(t_l) = (-1)^{l-k-1} \Delta_k$. Но $\eta(t_k)\eta(t_l) > 0$, ибо в противном случае на пути, соединяющем t_k и t_l, по непрерывности нашлась бы точка $\bar{t}$, в которой $\eta(\bar{t}) = 0$, что противоречит чебышевским свойствам системы $u_1, \dots, u_n$. Отсюда

$$\{(-1)^l \Delta_l\}\{(-1)^k \Delta_k\} < 0.$$

Достаточность доказана.

Необходимость. Допустим, вопреки утверждению, что Q содержит по крайней мере $n+1$ связных компонент $Q_0, Q_1, \dots, Q_n$. По условию на Q может быть задана чебышевская система функций $u_1, \dots, u_n$. Построим другую чебышевскую систему, которая не будет удовлетворять условию $(\mathcal{K})$.

Возьмем по одной точке t_j в каждом Q_j. Очевидно, что линейная система уравнений

$$\sum_{j=0}^{n} x_j U(t_j) = 0$$

имеет ненулевое решение. Более того, в силу чебышевских свойств системы $u_1, \dots, u_n$ все x_j отличны от нуля. Перепишем (5) в виде

$$\sum_{j=0}^{n} |x_j| \operatorname{sign} x_j U(t_j) = 0.$$

Теперь введем новую систему функций

$$v_i(t) = \begin{cases} (\operatorname{sign} x_j) u_i(t) & \forall t \in Q_j, \\ u_i(t) & \forall t \in Q \setminus \bigcup_{j=0}^{n} Q_j. \end{cases}$$

Очевидно, что функции $v_1, \dots, v_n$ непрерывны и образуют чебышевскую на Q систему. Положив $V = (v_1, \dots, v_n)$, $y_j = |x_j| \Big/ \sum_{j=0}^{n} |x_j|$, получим

$$\sum_{j=0}^{n} y_j V(t_j) = 0,$$

$$\sum_{j=0}^{n} y_j = 1, \quad y_j > 0 \quad \forall j \in 0:n.$$

В силу теоремы 1 по системе $v_1, \dots, v_n$ нельзя построить положительный на Q полином, что противоречит допущению. Теорема доказана.

§ 1.14. О чувствительности величины наилучшего приближения

В этом параграфе исследуется вопрос об изменении величины наилучшего приближения при варьировании аппроксимируемой функции.

1°. Пусть $u, u_1, \ldots, u_n$ - элементы банахова пространства B. Обозначим через $E(u)$ величину наилучшего приближения элемента u с помощью линейных комбинаций элементов $u_1, \ldots, u_n$:

$$E(u) = \min_{X \in E_n} \| u - \Sigma_{i=1}^{n} x_i u_i \| ,$$

где $X = (x_1, \ldots, x_n)$. Для произвольного $g \in B$ установим существование и выясним вид производной по направлению

$$\frac{\partial E(u)}{\partial g} = \lim_{\lambda \to +0} \frac{E(u+\lambda g) - E(u)}{\lambda} .$$

Через B^* будем обозначать сопряженное B пространство линейных функционалов со слабой* топологией. Положим

$$M = \{ F \in B^* \mid F(u_i) = 0 \quad \forall i \in 1:n, \ \|F\| \leqslant 1 \} .$$

Как известно [82],

$$E(u) = \max_{F \in M} F(u), \tag{1}$$

причем максимум достигается, ибо M - слабо* компактное множество.

Предложение. Для произвольного $g \in B$ существует производная по направлению и $\partial E(u)/\partial g = \max_{F \in M^*} F(g)$. Здесь M^* - множество функционалов, на которых достигается максимум в (1).

Доказательство немедленно следует из (1) и правила дифференцирования функции максимума (см., например, [97, с. 62]).

2°. Пусть Q - компакт на вещественной оси, $B = C(Q)$ и система $u_1, \ldots, u_n$ является чебышевской на некотором отрезке $[\alpha, \beta]$, содержащем Q. Зафиксируем $u \in C(Q)$ и предположим, что $E(u) > 0$. Тогда у функции u существует единственный полином наилучшего приближения $p^0 = \Sigma_{i=1}^{n} x_i^0 u_i$. Обозначим через T^0

множество точек максимального уклонения:

$$T^0 = \{ t \in Q \mid \ | u(t) - p^0(t) | = E(u) \} .$$

Теорема. Если T^0 состоит ровно из $n+1$ точек, то для любой функции $g \in C(Q)$ справедливо равенство

$$\left| \frac{\partial E(u)}{\partial g} \right| = E(g, T^0),$$

где $E(g, T^0)$ - величина наилучшего приближения функции g на T^0 с помощью полиномов $p = \Sigma_{i=1}^{n} x_i u_i$.

Доказательство. Учитывая (1) и теорему Рисса (см., например, [53, с. 28]), получаем

$$E(u) = \max_{\mu \in M} \int_Q u \, d\mu , \tag{2}$$

где M - множество регулярных борелевских мер, удовлетворяющих условиям

$$\int_Q u_i \, d\mu = 0 \quad \forall i \in 1:n, \quad |\mu|(Q) = 1 . \tag{3}$$

Если обозначить через M^* множество мер, на которых достигается максимум в (2), то согласно предложению

$$\frac{\partial E(u)}{\partial g} = \max_{\mu \in M^*} \int_Q g \, d\mu .$$

В силу (3) для любой меры $\mu \in M^*$

$$\int_Q u \, d\mu = \int_Q (u - p^0) \, d\mu = E(u) . \tag{4}$$

Но $| u(t) - p^0(t) | = E(u)$ при $t \in T^0$ и $| u(t) - p^0(t) | < E(u)$ вне T^0. Учитывая, кроме того, что $|\mu|(Q) = 1$, заключаем, что всякая мера $\mu \in M^*$ сосредоточена на T^0.

По условию T^0 состоит ровно из $n+1$ точек $t_0 < t_1 < \ldots < t_n$, а система $\{u_i\}$ - чебышевская на $[\alpha, \beta]$. Поэтому t_k являются точками альтернанса:

$$u(t_k) - p^0(t_k) = \sigma(-1)^k E(u), \quad k \in 0:n, \quad \sigma = \pm 1 . \tag{5}$$

Меры $\mu \in M^*$ порождаются нагрузками μ_k в точках t_k, причем на основании (4) и (5)

$$\Sigma_{k=0}^{n}\sigma(-1)^k E(u)\mu_k = E(u),\quad \Sigma_{k=0}^{n}|\mu_k| = 1 .$$

Отсюда следует, что $\mu_k = \sigma(-1)^k|\mu_k| \quad \forall k \in 0:n$.

Далее, согласно (3)

$$\Sigma_{k=0}^{n} u_i(t_k)(-1)^k|\mu_k| = 0 \quad \forall i \in 1:n,\ \Sigma_{k=0}^{n}|\mu_k| = 1. \qquad (6)$$

Так как система $\{u_i\}$ чебышевская, то числа $|\mu_k|$ однозначно определяются из (6) и необходимо являются положительными. Значит, M^* состоит из единственной меры μ^* и

$$E(u) = \int_Q u\, d\mu^* = \sigma \Sigma_{k=0}^{n}(-1)^k|\mu_k| u(t_k),$$

$$\frac{\partial E(u)}{\partial g} = \int_Q g\, d\mu^* = \sigma \Sigma_{k=0}^{n}(-1)^k|\mu_k| g(t_k).$$

Вместе с тем если приближать функцию g на множестве T^0, то все точки из T^0 будут точками альтернанса. Учитывая (6), получаем

$$E(g, T^0) = \left| \Sigma_{k=0}^{n}(-1)^k|\mu_k| g(t_k) \right| = \left| \frac{\partial E(u)}{\partial g} \right| .$$

Теорема доказана.

З а м е ч а н и е. Пусть Q - дискретное множество на вещественной оси, а T^0 состоит из $n+1$ точек $t_0 < t_1 < \ldots < t_n$. Обозначим через $g_k(t)$ функцию, равную единице при $t = t_k$ и нулю в остальных точках $t \in Q$. Тогда

$$\frac{\partial E(u)}{\partial g_k} = \sigma(-1)^k|\mu_k| = \mu_k .$$

Тем самым выяснен смысл чисел μ_k.

§ 1.15. Одна задача со связанными ограничениями

Расширение области практического применения теории экстремальных задач приводит к усложнению вида минимизируемых функций, а также множеств, на которых они определяются. В настоящем параграфе рассматривается одна из подобных задач, сводящаяся к минимизации функции максимума по множеству, определяемому в результате

нахождения максимума другой функции. Показывается, что в типичной постановке указанная задача некорректна, в связи с чем вводится новая критериальная функция, обладающая необходимыми для минимизации свойствами (непрерывностью и дифференцируемостью по направлениям). Приводятся геометрическая интерпретация необходимого условия минимума и метод нахождения направления наискорейшего спуска.

Задачи рассматриваемого вида возникают в теории игр [108], при многокритериальной оптимизации, а также в теории чувствительности и допусков электрических цепей [20].

Пусть
$$\varphi(X) = \max_{i \in I} f_i(X), \qquad (1)$$

где $I = 1:N$, $f_i(X)$ - непрерывно дифференцируемые функции на открытом множестве $U \subset E_n$. Обозначим, как обычно, множество максимумов в (1) через

$$R(X) = \{ i \in I \mid f_i(X) = \varphi(X) \} \qquad (2)$$

и рассмотрим функцию

$$\varphi_1(X) = \max_{i \in R(X)} F_i(X), \qquad (3)$$

где $F_i(X)$ - непрерывные функции, причем

$$F_i(X) > 0 \quad \forall X \in U . \qquad (4)$$

Требуется найти

$$\inf_{X \in \Omega} \varphi_1(X), \qquad (5)$$

где $\Omega \subset U$.

В частном случае функцией $F_i(X)$ может являться норма градиента функции $f_i(X)$.

Задача (5) может не быть корректной. Приведем простейший пример, в котором функция $\varphi_1(X)$ оказывается разрывной.

П р и м е р 1. Пусть $x \in E_1$, $I = 1:2$. Рассмотрим функции $f_1(x) = x^2$, $f_2(x) = 1{,}5x - 0{,}5x^2$. Тогда

$$\varphi(x) = \max\{ f_1(x), f_2(x) \}$$

и

$$R(X)=\begin{cases}\{1\} & \text{при } x\bar{\in}[0,1],\\ \{1,2\} & \text{при } x=0,\ x=1,\\ \{2\} & \text{при } x\in(0,1),\end{cases}$$

$$\varphi(x)=\begin{cases}1{,}5x-0{,}5x^2 & \text{при } x\in[0,1],\\ x^2 & \text{при } x\bar{\in}[0,1],\end{cases}$$

$$\varphi_1(x)=\begin{cases}-2x & \text{при } x<0,\\ 1{,}5-x & \text{при } x\in[0,1),\\ 2x & \text{при } x\geqslant 1.\end{cases}$$

Положим $F_i(x)=|f_i'(x)|$, $i=1,2$. Тогда $F_1(x)=|2x|$ и $F_2(x)=|1{,}5-x|$. Графики функций $\varphi(x)$ и $\varphi_1(x)$ изображены на рис. 4 и 5.

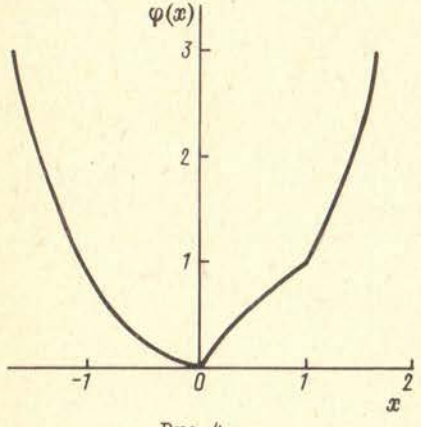


Рис.4.

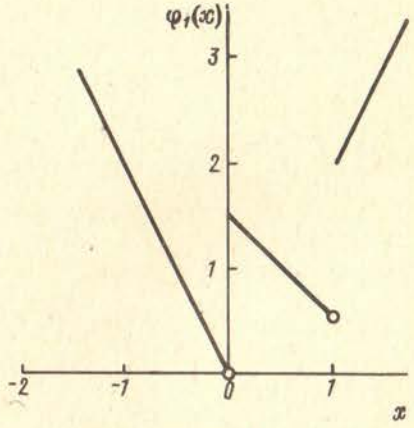


Рис.5.

Ясно, что $\inf \varphi_1(x)=0$, однако этот инфимум не достигается: при небольших изменениях переменной в окрестности точки $x'=0$ (так же, как и точки $x''=1$) функция $\varphi_1(x)$ может резко изменяться. Таким образом, в поставленном виде задача минимизации $\varphi_1(X)$ оказывается некорректной.

В практических задачах X является набором параметров физической системы, подлежащим выбору. Поэтому постановка экстремальной задачи должна обеспечивать по крайней мере непрерывность целевой функции при изменении X. Один из возможных методов реализации этого условия рассматривается далее.

Возьмем $\varepsilon>0$ и рассмотрим функцию

$$\varphi_{1\varepsilon}(X)=\max_{i\in R_\varepsilon(X)} F_i(X)\left[1-\frac{\varphi(X)-f_i(X)}{\varepsilon}\right]=\max_{i\in R_\varepsilon(X)} H_i(X), \tag{6}$$

где

$$R_\varepsilon(X)=\{i\in I \mid \varphi(X)-f_i(X)\leqslant\varepsilon\},$$
$$H_i(X)=F_i(X)\left[1-\frac{\varphi(X)-f_i(X)}{\varepsilon}\right]. \tag{7}$$

Поскольку в (6) $R_\varepsilon(X)$ зависит от X, то функция $\varphi_{1\varepsilon}(X)$ называется функцией максимума при связанных ограничениях.

Следует отметить, что при любом фиксированном $X\in U$ найдется такое $\varepsilon_0>0$, что $\varphi_{1\varepsilon}(X)=\varphi_1(X)$ для всех $\varepsilon\in(0,\varepsilon_0]$. Это свойство очевидно, ибо, как легко заметить, для любого X при достаточно малых $\varepsilon>0$ окажется $R_\varepsilon(X)=R(X)$.

Представляет интерес рассмотрение некоторых свойств функции $\varphi_{1\varepsilon}(X)$.

Лемма 1. При любом фиксированном $\varepsilon>0$ функция $\varphi_{1\varepsilon}(X)$ непрерывна по X.

Доказательство. Зафиксируем X_0. Если $i\in R_\varepsilon(X)$ таково, что $\varphi(X)-f_i(X)=\varepsilon$, то выражение в квадратных скобках в (6) равно нулю. Поэтому найдется такое $\delta>0$, что

$$\varphi_{1\varepsilon}(X)=\max_{i\in R_\varepsilon(X_0)} H_i(X) \tag{8}$$

для всех $X\in S_\delta(X_0)\overset{\text{def}}{=}\{X \mid \|X-X_0\|<\delta\}$, так как отображение $R_\varepsilon(X)$ полунепрерывно сверху. Из (8) теперь очевидна непрерывность функции $\varphi_{1\varepsilon}(X)$.

Пример 2. Пусть по-прежнему $f_1(x)=x^2$, $f_2(x)=1{,}5x-0{,}5x^2$ и взято $\varepsilon=0{,}1$. Тогда путем элементарных расчетов находим

$$R_\varepsilon(X)=\begin{cases}\{1\} & \text{при } x\bar{\in}[x_1,x_4],\\ \{1,2\} & \text{при } x\in[x_1,x_2]\cup[x_3,x_4],\\ \{2\} & \text{при } x\in(x_2,x_3),\end{cases}$$

где $x_1=\frac{1}{3}\left(1{,}5-\sqrt{2{,}85}\right)\approx -0{,}0627$, $x_2=\frac{1}{3}\left(1{,}5-\sqrt{1{,}65}\right)\approx 0{,}0718$,

$x_3=\frac{1}{3}\left(1{,}5-\sqrt{1{,}65}\right)\approx 0{,}9283$, $x_4=\frac{1}{3}\left(1{,}5+\sqrt{2{,}85}\right)\approx 1{,}0627$,

и
$$\varphi_{1\varepsilon}(x)=\begin{cases}|2x| & \text{при } x<x_1,\ x>1,\\ \max\left\{|2x|,\ (1{,}5-x)\left(1-\dfrac{1{,}5x^2-1{,}5x}{0{,}1}\right)\right\} & \text{при } x\in[x_1,0],\\ 1{,}5-x & \text{при } x\in[0,x_3],\\ \max\left\{(1{,}5-x),\ 2x\left(1-\dfrac{1{,}5x-1{,}5x^2}{0{,}1}\right)\right\} & \text{при } x\in[x_3,1].\end{cases}$$

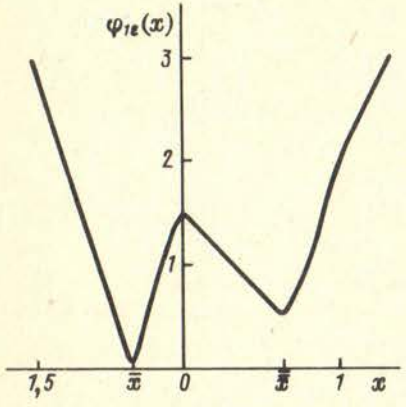


Рис. 6.

График функции $\varphi_{1\varepsilon}(x)$ изображен на рис. 6. Через $\bar{x}$ обозначен первый отрицательный корень уравнения

$$(1{,}5-x)(1-15x^2+15x)+2x=0,$$

$$\bar{x}\approx -0{,}058,$$

а через $\bar{\bar{x}}$ - первый меньше единицы корень уравнения

$$2x(1-15x+15x^2)+x-1{,}5=0,$$

$$\bar{\bar{x}}\approx 0{,}949.$$

Уравнения не упрощались для сохранения наглядности их формирования. Построенная кривая иллюстрирует отсутствие разрывов у функции $\varphi_{1\varepsilon}(x)$.

Лемма 2. Функция (6) дифференцируема по направлениям, причем

$$\frac{\partial\varphi_{1\varepsilon}(X)}{\partial g}\overset{def}{=}\lim_{\alpha\to+0}\frac{\varphi_{1\varepsilon}(X+\alpha g)-\varphi(X)}{\alpha}=$$

$$=\max_{i\in R_{2\varepsilon}(X)}\min_{j\in R(X)}\left\{\left(\frac{\partial F_i(X)}{\partial X}\left[1-\frac{\varphi(X)-f_i(X)}{\varepsilon}\right]+\right.\right.$$

$$\left.\left.+\frac{1}{\varepsilon}F_i(X)\left(f'_{iX}(X)-f'_{jX}(X)\right),\ g\right)\right\},\qquad(9)$$

где $R_{2\varepsilon}(X)=\{i\in R_\varepsilon(X)\mid H_i(X)=\varphi_{1\varepsilon}(X)\}$.

Доказательство. Если функция $\Phi_i(y)$ - непрерывно дифференцируемая функция одной переменной, а $\varphi(X)$ - дифференцируемая по направлениям функция, то и функция $u(X)=\max_{i\in I}\Phi_i(\varphi(X))$ также дифференцируема по направлениям, при этом

$$\frac{\partial u(X)}{\partial g}\overset{def}{=}\lim_{\alpha\to+0}\frac{u(X+\alpha g)-u(X)}{\alpha}=\max_{i\in Q(X)}\Phi'_{iX}(\varphi(X))\frac{\partial\varphi(X)}{\partial g},\qquad(10)$$

где $Q(X)=\{i\in I\mid u(X)=\Phi_i(\varphi(X))\}$.

Соотношение (10) непосредственно следует из представления

$$\Phi_i(\varphi(X+\alpha g))=\Phi_i\left(\varphi(X)+\alpha\frac{\partial\varphi(X)}{\partial g}+o(\alpha)\right)=$$

$$=\Phi_i(\varphi(X))+\alpha\Phi'_{iX}(\varphi(X))\frac{\partial\varphi(X)}{\partial g}+o_i(\alpha).$$

Применяя (10) к функции (6), легко получить (9).

Для определения условий минимума функции $\varphi_{1\varepsilon}(X)$ на U следует воспользоваться леммой 2, в которой установлено, что функция $\varphi_{1\varepsilon}(X)$ является дифференцируемой по направлениям. Это свойство позволяет сформулировать необходимое условие минимума в следующем виде.

Для того чтобы в точке $X^*\in U$ функция $\varphi_{1\varepsilon}(X)$ достигала минимального значения, необходимо, чтобы

$$\inf_{\|g\|=1}\frac{\partial\varphi_{1\varepsilon}(X^*)}{\partial g}\geqslant 0.\qquad(11)$$

Точка X^*, удовлетворяющая (11), называется стационарной точкой функции максимума. Если в точке X_0 условие (11) еще не выполнено, то направление $g(X_0)$, для которого

$$\frac{\partial \varphi_{1\varepsilon}(X_0)}{\partial g(X_0)} = \inf_{\|g\|=1} \frac{\partial \varphi_{1\varepsilon}(X_0)}{\partial g}, \tag{12}$$

называется направлением наискорейшего спуска.

Дадим геометрическую интерпретацию условия (11) и покажем, как можно находить направление наискорейшего спуска.

Введем обозначения

$$A_i(X) = F'_{iX}(X)\left[1 - \frac{\varphi(X) - f_i(X)}{\varepsilon}\right] + \frac{1}{\varepsilon} F_i(X) f_{iX}(X),$$

$$B_j(X) = -f_{jX}(X), \quad c_i(X) = \frac{1}{\varepsilon} F_i(X), \tag{13}$$

$$I_1(X) = R_{2\varepsilon}(X), \quad I_2(X) = R(X),$$

в которых $A_i(X), B_j(X)$ - векторы, $c_i(X)$ - положительные числа, $I_1(X)$ и $I_2(X)$ - множества индексов. Теперь формулу (9) можно переписать в виде

$$\frac{\partial \varphi_{1\varepsilon}(X)}{\partial g} = \max_{i \in I_1(X)} \min_{j \in I_2(X)} \left(A_i(X) + c_i(X) B_j(X),\ g\right). \tag{14}$$

Лемма 3. Условие (11) эквивалентно условию

$$0 \in \cap_{j \in I_2(X^*)} L_j(X^*), \tag{15}$$

где $L_j(X) = \operatorname{co}\{A_i(X) + c_j B_j(X) \mid i \in I_1(X)\}$.

Доказательство. Пусть выполнено условие (11), или, что то же,

$$\min_{\|g\|=1} \max_{i \in I_1} \min_{j \in I_2} \left(A_i + c_i B_j,\ g\right) \geqslant 0 \tag{16}$$

(здесь и далее в выражениях (13) аргументы опущены, т.е. $A_i = A_i(X)$ и т.д.). Покажем, что в этом случае выполнено (15). Допустим противное. Тогда найдется $j_0 \in I_2$ такое, что $0 \notin L_{j_0}(X^*)$. По теореме отделимости найдутся такое число $\alpha > 0$ и такой вектор

$g_0 \in E_n$, $\|g_0\| = 1$, что

$$\left(A_i + c_i B_{j_0}, g_0\right) \leqslant -\alpha < 0 \quad \forall i \in I_1.$$

Тогда тем более

$$b_i = \min_{j \in I_2} \left(A_i + c_i B_j,\ g_0\right) \leqslant -\alpha \quad \forall i \in I_1.$$

Значит, и $\max_{i \in I_1} b_i \leqslant -\alpha$, что противоречит (16). Итак, из (16) следует (15).

Пусть теперь выполнено (15). Докажем, что в этом случае справедливо (16). Допустим противное. Тогда существует такое $\bar{g} \in E_n$, $\|\bar{g}\| = 1$, что

$$\max_{i \in I_1} \min_{j \in I_2} \left(A_i + c_i B_j,\ \bar{g}\right) = -\alpha < 0. \tag{17}$$

Так как

$$\min_{j \in I_2} \left(A_i + c_i B_j, \bar{g}\right) = \left(A_i, \bar{g}\right) + c_i \min_{j \in I_2} \left(B_j, \bar{g}\right) = \left(A_i, \bar{g}\right) + c_i \left(B_{j_0}, \bar{g}\right),$$

где j_0 одинаково для всех i, то из (17) следует

$$\max_{i \in I_1} \left(A_i + c_i B_{j_0}, \bar{g}\right) = -\alpha < 0.$$

Таким образом, $0 \notin L_{j_0}(X^*)$, что противоречит (15). Лемма доказана.

Теперь ясно, что для нахождения направления наискорейшего спуска $g(X_0)$ необходимо найти

$$\max_{j \in I_2(X_0)} \min_{z \in L_j(X_0)} \|z\| = \min_{z \in L_{j_0}(X_0)} \|z\| = \|z_0\|, \tag{18}$$

где z - точка выпуклой оболочки $L_j(X)$, и взять

$$g(X_0) = \frac{-z_0}{\|z_0\|}. \tag{19}$$

Описанная операция сводится к решению нескольких задач квадратичного программирования. Их количество определится числом точек во множестве $I_2(X_0) = R(X_0)$. Из (18) следует также, что направление наискорейшего спуска может оказаться не единственным.

Замечание 1. Для нахождения (19) можно применять обычные вычислительные методы, например, методы типа методов наискорейшего спуска или симплекс-метод – для квадратичного программирования (см. § 2.7 и 2.8).

Замечание 2. Ранее предполагалось, что в (5) $\Omega = U$. Если Ω – замкнутое выпуклое множество, заданное неравенствами, то все изложенное переносится с соответствующими изменениями и на этот случай.

Замечание 3. Естественным образом полученные результаты могут быть распространены на случай, когда множество I в (1) состоит из бесконечного числа точек.

§ 1.16. Целочисленная аппроксимация задачи оптимального размещения прямоугольников в трапеции

1°. Постановка задачи. Дана трапецеидальная полость (рис. 7), которую в дальнейшем будем называть трапецией. Обозначим через α угол отклонения ее боковых сторон от вертикали, а через l – длину ее основания. В этой трапеции нужно разместить n прямоугольников заданного размера таким образом, чтобы их основания (которые фиксированы) были параллельны основанию трапеции и прямоугольники не перекрывали друг друга. Очевидно, что таких размещений существует бесконечное множество. Обозначим через h расстояние от самого удаленного верхнего основания прямоугольника до основания трапеции. Из всевозможных размещений нужно выбрать те, при которых достигается $\min h$. Не исключено, что таких размещений может быть несколько или даже бесконечное множество.

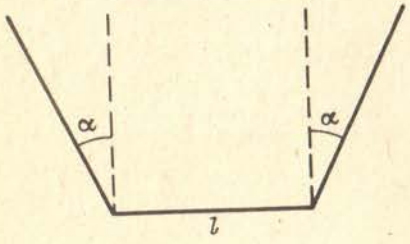


Рис. 7.

В этом параграфе рассмотрены три аспекта сформулированной проблемы. Во-первых, установлена оценка снизу для минимального значения высоты h трапеции при заданном наборе прямоугольников (п.3°). Во-вторых, с помощью операции, которую мы в дальнейшем называем "построением сетки" (п.2°), задача редуцирована к случаю, когда все размеры прямоугольников и основание l трапеции – целые числа, а угол α равен 45°. Наконец, доказано (п.4°) существование

оптимального расположения прямоугольников с вершинами в узлах построенной сетки.

2°. Построение сетки. Даны числа $b_1, b_2, \ldots, b_n$ и ε. Пусть

$$\min_i b_i > \varepsilon > 0 . \qquad (1)$$

Число d будем называть наибольшим общим делителем (НОД) чисел $b_1, b_2, \ldots, b_n$ с точностью аппроксимации ε, если d – наибольшее из чисел, для которого справедливы равенства

$$b_i = d \times k_i + \varepsilon_i , \qquad i \in 1:n, \qquad (2)$$

где k_i – натуральные числа, а

$$\max_i |\varepsilon_i| \leqslant \varepsilon . \qquad (3)$$

В этом случае будем писать $d = \text{НОД}(b_1, b_2, \ldots, b_n)_\varepsilon$. Очевидно, что $d > \varepsilon$.

Если d таково, что возможно неоднозначное представление некоторых чисел b_j в виде (2), например $b_j = d \times k'_j + \varepsilon'_j$ и $b_j = d \times k''_j + \varepsilon''_j$, то в качестве k_j берем то из чисел k'_j и k''_j, которое соответствует меньшему из $|\varepsilon'_j|$ и $|\varepsilon''_j|$. В случае равенства $|\varepsilon'_j|$ и $|\varepsilon''_j|$ в качестве k_j берем меньшее из чисел k'_j и k''_j.

Пусть $a_i \times b_i$ – размеры i-го прямоугольника (a_i – ширина, b_i – высота) и пусть

$$d_b = \text{НОД}(b_1, b_2, \ldots, b_n)_{\varepsilon_b} .$$

Тогда если за ε_b взять наибольшую погрешность по вертикали, которой можно пренебречь, то каждое из чисел b_i может быть заменено натуральным числом k_i, если за единицу масштаба по вертикали взять d_b.

Пусть $d_a = \text{НОД}(a_1, a_2, \ldots, a_n, l, l_1)_{\varepsilon_a}$. Тогда

$$a_i = d_a \times m_i + \varepsilon_{a_i}, \quad |\varepsilon_{a_i}| \leqslant \varepsilon_a,$$
$$l = d_a \times \lambda + \varepsilon_l, \quad |\varepsilon_l| \leqslant \varepsilon_a,$$
$$l_1 = d_a \times \lambda_1 + \varepsilon_{l_1}, \quad |\varepsilon_{l_1}| \leqslant \varepsilon_a,$$

где m_i, λ, λ_1 – натуральные числа, а l_1 – длина горизонтального отрезка, параллельного основанию трапеции и отстоящего от основания на расстояние d_b.

Если за ε_a взять наибольшую погрешность по горизонтали, ко-

торой можно пренебречь, а за d_a - единицу масштаба по горизонтали, то имеем $a_i = m_i$, $l = \lambda$, $l_1 = \lambda_1$.

Построим трапецию с основанием λ и длиной отрезка, отстоящего от основания на d_b, равной λ_1. Тогда угол наклона боковой стенки трапеции к вертикали будет равен

$$\beta = \operatorname{arctg} \frac{(\lambda_1 - \lambda) d_a}{2 d_b} .$$

Предположим, что наклон стенок трапеции должен быть выдержан с точностью ε_α. Тогда, если $|\alpha - \beta| > \varepsilon_\alpha$, необходимо, уменьшив ε_a, пересчитать НОД d_a, который может только уменьшиться.

Этот процесс следует продолжать до тех пор, пока не будет выполнено условие

$$|\alpha - \beta| \leqslant \varepsilon_\alpha . \qquad (4)$$

Условие (4) при любом как угодно малом ε_α должно выполняться, если ε_a станет достаточно малым, ибо $\beta \to \alpha$ при $\varepsilon_a \to 0$.

В дальнейшем вместо прямоугольников с размерами $a_i \times b_i$ будем рассматривать прямоугольники с размерами $m_i \times k_i$, где m_i - ширина i-го прямоугольника, если за единицу масштаба по горизонтали взять d_a, а k_i - его высота, если за единицу масштаба по вертикали взять d_b. Эти прямоугольники будем размещать в трапеции с основанием λ и углом наклона ее боковой стенки к вертикали β. Тогда горизонтальный отрезок, отстоящий от основания на расстояние d_b, будет иметь длину λ_1. Будем предполагать также, что величина $\lambda_1 - \lambda$ - число четное. Этого можно добиться при вычислении d_a. Для этого нужно уменьшать ε_a до тех пор, пока не будет выполнено и это условие.

Теперь трапеция может быть покрыта сеткой (рис. 8). Расстояние между горизонтальными линиями равно d_b, а между вертикальными - d_a. Каждый следующий горизонтальный ряд сетки содержит на $\lambda_1 - \lambda$ клеток больше, чем предыдущий. Построение сетки закончено.

Несложное доказательство нижеследующей теоремы мы опускаем.

Теорема 1. Если $d = \text{НОД}\,(b_1, b_2, \ldots, b_n)_\varepsilon$, то справедливы равенства

$$\text{НОД}(k_1, k_2, \ldots, k_n) = 1 , \qquad (5)$$

$$\min_i \varepsilon_i = -\varepsilon . \qquad (6)$$

Условия (5) и (6) служат необходимыми условиями того, что $d = \text{НОД}\,(b_1, b_2, \ldots, b_n)_\varepsilon$, но не являются достаточными.

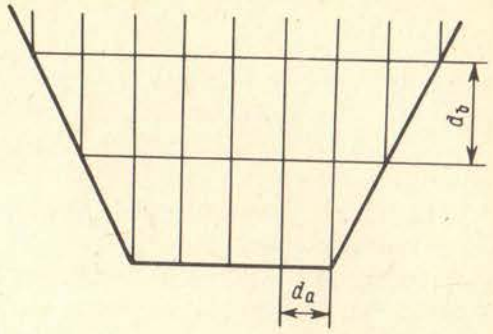


Рис. 8.

Перейдем к описанию алгоритма вычисления $\text{НОД}(b_1, b_2, \ldots, b_n)$ с точностью аппроксимации ε. Очевидно, не ограничивая общности, можно считать, что равных чисел среди $b_1, b_2, \ldots, b_n$ нет.

Пусть $d = \text{НОД}\,(b_1, b_2, \ldots, b_n)_\varepsilon$. Тогда

$$b_i = d \times k_i + \varepsilon_i , \quad i \in 1{:}n , \qquad \min_i \varepsilon_i = -\varepsilon . \qquad (7)$$

Из равенств (7) следует $d = (b_i - \varepsilon_i)/k_i$. Предположим, что $\varepsilon_{i_0} = -\varepsilon$. Тогда $d = (b_{i_0} + \varepsilon)/k_{i_0}$.

Таким образом, алгоритм вычисления d может быть следующим. Положим вначале $i_0 = 1$. В качестве k_1 возьмем наименьшее число натурального ряда, для которого выполняется условие

$$\max_i |\varepsilon_i| \leqslant \varepsilon , \quad i \in 1{:}n . \qquad (8)$$

Подобрать такое k_1 можно, присваивая ему последовательно значения 1, 2, 3, ... Очевидно, что процесс этот конечен. Обозначим соответствующие d через d_1. Затем положим $i_0 = 2$. И снова возьмем в качестве k_2 наименьшее число натурального ряда, для которого выполняется условие (8). Соответствующее ему значение d обозначим через d_2.

Будем продолжать этот процесс до тех пор, пока не переберем все b_i. В результате получим $d_1, d_2, \ldots, d_n$. Очевидно, что $d = \max(d_1, d_2, \ldots, d_n)$.

3°. Оценка нижней границы оптимальной высоты. Вычислим нижнюю границу оптимальной высоты. Если эта граница близка к точной нижней границе, то, сравнивая ее с высотами, полученными в процессе решения, вычисления можно прекратить, когда будет получена высота, достаточно близкая, с практической точки зрения, к нижней границе оптимальной высоты.

Грубо оценить точную нижнюю границу можно следующим образом. Вычислить S – сумму площадей всех прямоугольников $S=\sum_{i=1}^{n} a_i b_i$ и затем найти высоту трапеции h с основанием l и углом наклона боковых сторон к вертикали α, площадь которой равна S. Очевидно, что $h=\left(\sqrt{l^2+4S\operatorname{tg}\alpha}-l\right)/2\operatorname{tg}\alpha$. Это значение h и является одной из нижних границ оптимальной высоты $H_{опт}$.

Далее приводится алгоритм вычисления более точной нижней границы оптимальной высоты.

Построим два набора прямоугольных треугольников – одни с основаниями a_i, другие с высотами b_i, как показано на рис. 9.

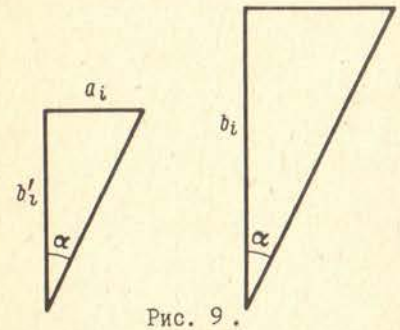


Рис. 9.

Высоты треугольников с основаниями a_i обозначим через $b_i'=a_i\operatorname{ctg}\alpha$.

Расположим числа b_i и b_i' в порядке их возрастания, что соответствует возрастанию площадей треугольников, высотами которых они являются, и обозначим их в этом порядке через h_j ($j\in 1:2n$). Таким образом, $h_1\leqslant h_2\leqslant\ldots\leqslant h_{2n}$. Впишем в трапецию ступенчатую фигуру следующим образом (рис. 10). В концах основания трапеции P_1, Q_1 восставим перпендикуляры к основанию длиной h_1. Через верхние концы этих перпендикуляров R_1 и S_1 проведем горизонтальные линии до пересечения с боковыми сторонами трапеции в точках P_2 и Q_2. Из точек P_2, Q_2 отложим внутрь трапеции отрезки, перпендикулярные основанию длиной h_2. Через верхние концы этих отрезков R_2 и S_2 проведем горизонтальные линии до пересечения с боковыми сторонами трапеции в точках P_3 и Q_3 и т.д. Наконец, из точек P_{2n}, Q_{2n} отложим внутрь трапеции отрезки, перпендикулярные основанию, длиной h_{2n}, верхние концы которых обозначим R_{2n}, S_{2n}. Таким образом, ступенчатая фигура получена из трапеции последовательным отсечением указанным способом $2n$ пар прямоугольных треугольников, высоты которых равны $h_1, h_2, \ldots, h_{2n}$. Такую ступенчатую фигуру будем называть в дальнейшем основной ступенчатой фигурой.

Будем перемещать точки P_i, Q_i вверх вдоль боковых сторон трапеции таким образом, чтобы нижнее основание ступенчатой фигуры $P_1 Q_1$ оставалось параллельным основанию трапеции, а расстояния

$P_i P_{i+1}$, $Q_i Q_{i+1}$ сохранялись неизменными. Тогда высота h части основной ступенчатой фигуры, площадь которой равна S, будет монотонно убывать. Обозначим расстояние от основания ступенчатой фигуры до основания трапеции через $\tilde{h}$.

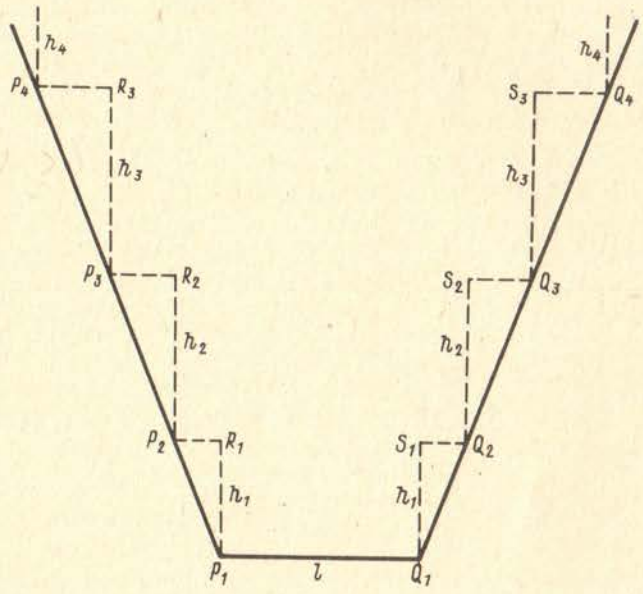


Рис. 10.

Теорема 2. Пусть сумма площадей всех прямоугольников равна S. Тогда за нижнюю границу оптимальной высоты можно брать величину $H_н$, вычисленную по формуле $H_н=\min_{\tilde{h}}(h+\tilde{h})$, где h – высота той части основной ступенчатой фигуры, площадь которой равна S, а $\tilde{h}$ – расстояние от основания трапеции до нижнего основания ступенчатой фигуры (см. рис. 11).

Предварительно докажем несколько лемм.

Лемма 1. Площадь ступенчатой фигуры $P_1 Q_1 \ldots P_{2n} Q_{2n}$ больше S при любых $l\geqslant 0$ и $\alpha\in(0,\pi/2)$.

Доказательство. В какой бы последовательности мы ни брали высоты h_j, площади ступенчатых фигур, построенных описанных выше способом, одинаковы. Действительно, все ступенчатые фигуры имеют одинаковую высоту, равную $\sum_{j=1}^{2n} h_j$. Площадь

каждой ступенчатой фигуры равна площади трапеции с той же высотой без суммы площадей всех треугольников.

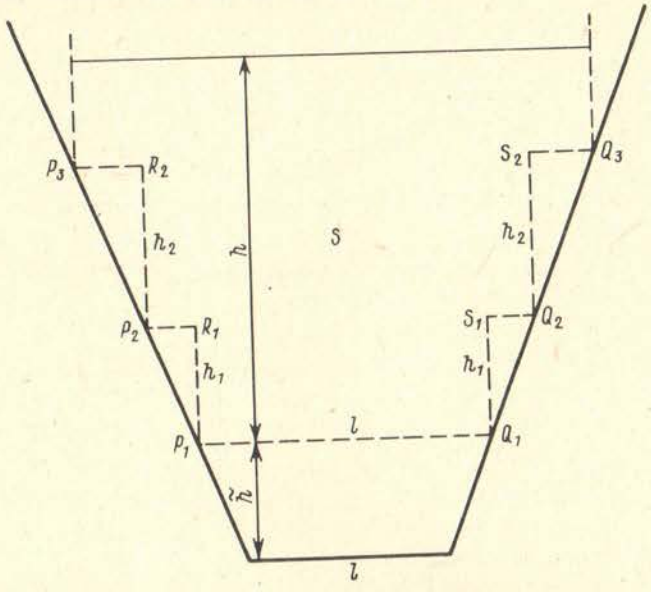


Рис. 11.

Расположим числа h_j в порядке $b_1, b_2, \ldots, b_n, b'_1, b'_2, \ldots, b'_n$ и построим соответствующую ступенчатую фигуру. Ее площадь, равную площади фигуры $P_1Q_1 \ldots P_{2n}Q_{2n}$, вычислим по формуле

$$S_{P_1Q_1\ldots P_{2n}Q_{2n}} = l b_1 + [l+2b_1\operatorname{tg}\alpha]\, b_2 + \ldots + [l+2(b_1+b_2+\ldots+b_{n-1})\operatorname{tg}\alpha]\, b_n +$$

$$+ [l+2(b_1+b_2+\ldots+b_n)\operatorname{tg}\alpha]\, b'_1 + [l+2(b_1+b_2+\ldots+b_n+b'_1)\operatorname{tg}\alpha]\, b'_2 + \ldots$$

$$\ldots + [l+2(b_1+b_2+\ldots+b_n+b'_1+b'_2+\ldots+b'_{n-1})\operatorname{tg}\alpha]\, b'_n >$$

$$> (b_1b'_1 + b_2b'_2 + \ldots + b_nb'_n)\operatorname{tg}\alpha .$$

Но $b'_j \operatorname{tg}\alpha = a_i$. Следовательно, $S_{P_1Q_1\ldots P_{2n}Q_{2n}} > \sum_{i=1}^{n} a_i b_i = S$. Лемма доказана.

Составим всевозможные перестановки из чисел $h_1, h_2, \ldots, h_{2n}$ и построим описанным способом соответствующие ступенчатые фигуры с одинаковым основанием l. В каждой ступенчатой фигуре проведем горизонтальную линию, отсекающую, считая от основания, часть фигуры, площадь которой равна S. Легко проверяется нижеследующее утверждение.

Лемма 2. Среди всех ступенчатых фигур рассматриваемого вида с площадью S наименьшую высоту имеет основная ступенчатая фигура.

Лемма 3. Пусть при некотором расположении прямоугольников в трапеции они образуют фигуру высотой $H \leqslant \sum_{j=1}^{2n} h_j$, причем нижнее основание хотя бы одного прямоугольника расположено на основании трапеции. Тогда площадь этой фигуры не превосходит площади той части основной ступенчатой фигуры, высота которой равна H.

Доказательство. Любая фигура может быть получена из трапеции отсечением прямоугольных треугольников и трапеций, примыкающих к ее боковым сторонам. Высоты отсекаемых прямоугольных треугольников не могут быть меньше высот h_j, а отсекаемые трапеции представляют собой либо верхние части прямоугольных треугольников с высотами h_j, либо верхние части подобных им треугольников с бóльшими высотами, либо равнобочные трапеции. Таким образом, любой набор отсекаемых прямоугольных треугольников и трапеций, сумма высот которых равна H, по площади не меньше суммы площадей прямоугольных треугольников, отсекаемых в последовательности $h_1 \leqslant h_2 \leqslant \ldots \leqslant h_k, \Delta h_{k+1}, h_1 = \min_j h_j$, общей высотой H. Следовательно, и площадь основной ступенчатой фигуры высотой H не меньше площади любой другой фигуры с такой же высотой. Лемма доказана.

Из леммы следует, что высота той части основной ступенчатой фигуры, площадь которой равна площади любой другой ступенчатой фигуры, не превосходит высоты последней, если их нижние основания совпадают. Таким образом, нижнюю границу оптимальной высоты можно вычислять по формуле, приведенной в теореме: $H_т = \min_{\tilde{h}} (h + \tilde{h})$.

Теорема 2 доказана.

Пусть при некотором значении $\tilde{h}$ величина h равна $h = H_k + \Delta h_{k+1}$, где $H_k = h_1 + h_2 + \ldots + h_k$, а основание ступенчатой фигуры равно $\tilde{l}$ (см. рис. 11). Тогда площадь ступенчатой фигуры вычисляется по формуле

$$S = h^2 \operatorname{tg}\alpha + h\tilde{l} - G_k^2 \operatorname{tg}\alpha - \Delta h_{k+1}^2 \operatorname{tg}\alpha , \qquad (9)$$

где $\tilde{l} = l_1 + 2\tilde{h}\operatorname{tg}\alpha$, $G_k^2 = h_1^2 + h_2^2 + \ldots + h_k^2$. Предположим, что площадь S при изменении $\tilde{h}$ остается постоянной. Тогда величина h является функцией от $\tilde{h}$. Введем H по формуле $H = h(\tilde{h}) + \tilde{h}$. Покажем, что H с ростом $\tilde{h}$ может убывать.

Предположим, что h меняется в таких пределах, при которых k не меняется. Тогда $H_k = \text{const}$, $G_k^2 = \text{const}$, $\Delta h_{k+1} \leqslant h_{k+1}$. С помощью равенства $\Delta h_{k+1} = h - H_k$ выражение для S можно привести к виду

$$S = h[l + 2(H_k + \tilde{h})\operatorname{tg}\alpha] - (H_k^2 + G_k^2)\operatorname{tg}\alpha, \qquad (10)$$

откуда

$$h = \frac{S + (H_k^2 + G_k^2)\operatorname{tg}\alpha}{l + 2(H_k + \tilde{h})\operatorname{tg}\alpha}.$$

Следовательно,

$$\frac{dH}{d\tilde{h}} = 1 + \frac{dh}{d\tilde{h}} = 1 - \frac{2[S + (H_k^2 + G_k^2)\operatorname{tg}\alpha]\operatorname{tg}\alpha}{[l + 2(H_k + \tilde{h})\operatorname{tg}\alpha]^2}.$$

Заменяя в правой части S по формуле (10), окончательно получаем

$$\frac{dH}{d\tilde{h}} = 1 - \frac{2h\operatorname{tg}\alpha}{l + 2(H_k + \tilde{h})\operatorname{tg}\alpha}.$$

Из условия $dH/d\tilde{h} < 0$ имеем $\Delta h_{k+1} > \frac{1}{2} l \operatorname{ctg}\alpha + \tilde{h}$.

Таким образом, если при $\tilde{h} = 0$ выполняется неравенство $\Delta h_{k+1} > l \operatorname{ctg}\alpha$, то $dH/d\tilde{h} < 0$ и величина H с ростом $\tilde{h}$ убывает.

Алгоритм вычисления H_r. Пусть при $\tilde{h} = 0$ $k = k_0$, $0 < \Delta h_{k_0+1} \leqslant h_{k_0+1}$. С ростом $\tilde{h}$ величина k будет последовательно принимать значения $k_0, k_0 - 1, \ldots, 2, 1, 0$. Предположим, что при $k = i$ величина $\tilde{h}$ принимает все значения из промежутка $[\tilde{h}_{i+1}, \tilde{h}_i)$. Например, при $k = k_0$ она принимает все значения из полуинтервала $[0, \tilde{h}_{k_0})$.

Пусть $\Delta h_{k_0+1} > \frac{1}{2} \operatorname{ctg}\alpha$. Тогда $(dH/d\tilde{h})_{\tilde{h}=0} < 0$. Поэтому при увеличении $\tilde{h}$ функция $H(\tilde{h})$ начинает убывать. Покажем, что в этом случае на промежутке $[0, \tilde{h}_{k_0})$ существует минимум H. Действительно, при $\tilde{h} \to \tilde{h}_{k_0} - 0$ величина $\Delta h_{k_0+1} \to 0$, следовательно, $h \to H_{k_0}$. Но

$$\lim_{\tilde{h} \to \tilde{h}_{k_0} - 0} \frac{dH}{d\tilde{h}} = 1 - \frac{2H_{k_0}\operatorname{tg}\alpha}{l + 2(H_{k_0} + \tilde{h}_{k_0})\operatorname{tg}\alpha} > 0.$$

В силу непрерывности функции $dH/d\tilde{h}$ на промежутке $[0, \tilde{h}_{k_0})$ существует точка $\tilde{\tilde{h}}_{k_0}$, в которой $dH/d\tilde{h} = 0$. Решая это уравнение относительно $\tilde{h}$, получаем

$$\tilde{\tilde{h}}_{k_0} = \left[\frac{1}{2} S \operatorname{ctg}\alpha + \frac{1}{2}\left(H_{k_0}^2 + G_{k_0}^2\right)\right]^{1/2} - \frac{1}{2} l \operatorname{ctg}\alpha - H_{k_0}.$$

Так как $H(\tilde{\tilde{h}}_{k_0}) = h(\tilde{\tilde{h}}_{k_0}) + \tilde{\tilde{h}}_{k_0}$, то

$$H(\tilde{h}_{k_0}) = \left[2S\operatorname{ctg}\alpha + 2\left(H_{k_0}^2 + G_{k_0}^2\right)\right]^{1/2} - \frac{1}{2} l \operatorname{ctg}\alpha - H_{k_0} = 2\tilde{\tilde{h}}_{k_0} + H_{k_0} + \frac{1}{2} l \operatorname{ctg}\alpha.$$

Будем теперь менять $\tilde{h}$ в пределах $[\tilde{h}_{k_0}, \tilde{h}_{k_0-1})$. Начальное значение Δh_{k_0} равно h_{k_0}. Если $h_{k_0} > \frac{1}{2} l \operatorname{ctg}\alpha + \tilde{h}_{k_0}$, то снова

$$\lim_{\tilde{h} \to \tilde{h}_{k_0} + 0} \frac{dH}{d\tilde{h}} = -\frac{2H_{k_0}\operatorname{tg}\alpha}{l + 2(H_{k_0-1} + \tilde{h}_{k_0})\operatorname{tg}\alpha} < 0.$$

Но

$$\lim_{\tilde{h} \to \tilde{h}_{k_0-1} - 0} \frac{dH}{d\tilde{h}} = 1 - \frac{2H_{k_0}\operatorname{tg}\alpha}{l + 2\left(H_{k_0-1} + \tilde{h}_{k_0+1}\right)\operatorname{tg}\alpha} > 0.$$

Следовательно, и на промежутке $[\tilde{h}_{k_0}, \tilde{h}_{k_0-1})$ функция $H(\tilde{h})$ имеет минимум. Вычисления дают

$$H(\tilde{\tilde{h}}_{k_0-1}) = 2\tilde{\tilde{h}}_{k_0-1} + H_{k_0-1} + \frac{1}{2} l \operatorname{ctg}\alpha.$$

Так как $H_{k_0} > H_{k_0-1}$, а $\tilde{\tilde{h}}_{k_0} < \tilde{\tilde{h}}_{k_0-1}$, то сказать заранее, какой из минимумов меньше, нельзя.

Таким образом, выводим следующее правило для нахождения наименьшего значения H.

1. Вычисляем $H(0) = h(0)$ по формуле

$$H(0) = h(0) = \frac{S + (H_k^2 + G_k^2)\operatorname{tg}\alpha}{l + 2H_k \operatorname{tg}\alpha},$$

причем k подбираем из условия $h - H_k \leqslant h_{k+1}$.

2. Из уравнения $h(\tilde{h}_i) = H_i$ $(i = k, k-1, \ldots, 1)$ определяем границы промежутков $[\tilde{h}_{i+1}, \tilde{h}_i)$, причем $\tilde{h}_{k+1} = 0$,

$$\tilde{h}_i = \frac{S - lH_i + (H_i^2 + G_i^2)\operatorname{tg}\alpha}{2H_i \operatorname{tg}\alpha}.$$

3. Находим h'_j по формуле

$$h'_j = \begin{cases} h_j, & j = 1, 2, \ldots, k, \\ h(0) - H_k, & j = k+1. \end{cases}$$

4. Если $h'_j > 1/2\, l \operatorname{ctg}\alpha + h'_j$ $(j=k+1,k,\ldots,1)$, то на соответствующем промежутке есть минимум $\widetilde{H}_{j-1}$, который вычисляется по формуле

$$\widetilde{H}_{j-1} = \left[2S \operatorname{ctg}\alpha + 2\left(H_{j-1}^2 + G_{j-1}^2\right)\right]^{1/2} - \frac{1}{2} l \operatorname{ctg}\alpha - H_{j-1}.$$

5. Получаем H_r по формуле $H_r = \min\left(H(0), \widetilde{H}_k, \widetilde{H}_{k-1}, \ldots, \widetilde{H}_1\right)$.

4°. Существование оптимального расположения прямоугольников с вершинами в узлах сетки. В силу сказанного в пп. 1° и 2° можем считать, что в трапеции уже построена сетка (см. рис. 8). В масштабе этой сетки все стороны прямоугольников и основание трапеции являются целыми числами, а угловой коэффициент наклона боковой стороны к основанию равен 1.

В настоящем параграфе показано, что можно построить оптимальное решение, при котором вершины всех прямоугольников расположены в узлах исходной сетки либо в узлах сетки, полученной из исходной параллельным переносом на вектор (1/2, 1/2). Это утверждение непосредственно вытекает из теорем 3 и 4.

Теорема 3. Предположим, что основание трапеции и стороны всех прямоугольников – четные целые числа. Тогда существует оптимальное решение, при котором вершины прямоугольников расположены в узлах сетки.

Доказательство теоремы проведем в несколько этапов.

1. Определения и обозначения. Будем говорить, что прямоугольники Δ и Δ' связаны горизонтальной (вертикальной) цепочкой, если найдутся такие прямоугольники $\Delta_1 = \Delta, \Delta_2, \ldots, \Delta_n = \Delta'$, что Δ_i и Δ_{i+1} $(i = 1, \ldots, n-1)$ касаются по части боковых сторон (по части оснований). Нижними будем называть прямоугольники, стоящие на основании трапеции.

Прямоугольник называется правым (левым) боковым, если одна из его вершин находится на правой (левой) стороне трапеции. Некоторые прямоугольники могут быть одновременно и правыми боковыми и левыми боковыми.

Будем считать сетку координатной. Начало координат выберем в левом нижнем углу трапеции. Координатами прямоугольника Δ будем называть координаты его правого верхнего угла. Обозначим их через x_Δ и y_Δ.

2. Несколько очевидных утверждений.

Лемма 1. Координаты бокового прямоугольника являются либо все целыми либо все нецелыми.

Лемма 2. Если прямоугольники Δ и Δ' связаны горизонтальной (вертикальной) цепочкой, то $x_\Delta - x_{\Delta'}$ $(y_\Delta - y_{\Delta'})$ – целые числа. В частности, если прямоугольник Δ связан вертикальной цепочкой с нижним прямоугольником, то y_Δ – целое число.

Лемма 3. Если левый боковой Δ и правый боковой Δ' прямоугольники связаны горизонтальной цепочкой, то $y_\Delta + y_{\Delta'}$ – четное целое число.

3. Первая модификация оптимального решения. Пусть расположение прямоугольников оптимальное. Их можно сдвинуть так, чтобы выполнялось следующее условие: каждый прямоугольник Δ является боковым либо соединен с боковыми горизонтальной и вертикальной цепочками одновременно (разумеется, эти цепочки связывают Δ с разными боковыми прямоугольниками).

Пусть первая модификация уже произведена. Тогда для того чтобы установить целочисленность координат всех прямоугольников, достаточно в силу леммы 2 показать, что координаты боковых прямоугольников целочисленны (а в силу леммы 1 достаточно установить целочисленность одной из координат каждого бокового прямоугольника).

4. Определение множеств Λ^Δ, Π^Δ, $B\Lambda^\Delta$, $B\Pi^\Delta$. Пусть Δ – боковой прямоугольник. Для определенности будем предполагать его правым. Обозначим через Λ_1^Δ множество левых боковых прямоугольников, которые можно связать с Δ горизонтальной цепочкой, а через Π_1^Δ – множество правых боковых прямоугольников, которые связаны горизонтальными цепочками с прямоугольниками из Λ_1^Δ (очевидно, что $\Delta \in \Pi_1^\Delta$). Далее, аналогично по множеству Π_1^Δ построим множество Λ_2^Δ левых прямоугольников, связанных горизонтальными цепочками с прямоугольниками из Π_1^Δ, и т.д. В результате на некотором шаге получим множество $\Lambda_k^\Delta = \Lambda^\Delta$ левых боковых прямоугольников и множество $\Pi_k^\Delta = \Pi^\Delta$ правых боковых прямоугольников, которые удовлетворяют следующим условиям.

Свойство 1. Любой боковой прямоугольник, связанный горизонтальной цепочкой с каким-либо прямоугольником из $\Lambda^\Delta \cup \Pi^\Delta$, принадлежит $\Lambda^\Delta \cup \Pi^\Delta$.

Свойство 2. Любые два прямоугольника из $\Lambda^\Delta \cup \Pi^\Delta$ можно соединить горизонтальной цепочкой.

Обозначим далее через $B\Lambda^\Delta$ множество прямоугольников, связанных с прямоугольниками из Λ^Δ вертикальными цепочками. Аналогично построим множество $B\Pi^\Delta$ прямоугольников, связанных вертикальными цепочками с прямоугольниками из Π^Δ. Очевидно, что $\Lambda^\Delta \subset B\Lambda^\Delta$ и $\Pi^\Delta \subset B\Pi^\Delta$.

Положим далее $y_\Lambda = \max\limits_{\Delta' \in B\Lambda^\Delta} y_{\Delta'}$, $y_\Pi = \max\limits_{\Delta' \in B\Pi^\Delta} y_{\Delta'}$.

5. Свойства множеств Λ^Δ, Π^Δ, $B\Lambda^\Delta$ и $B\Pi^\Delta$. 1. Если хотя бы один из прямоугольников из $\Lambda^\Delta \cup \Pi^\Delta$ имеет целые координаты, то все прямоугольники из $\Lambda^\Delta \cup \Pi^\Delta$ также имеют целые координаты.

Утверждение непосредственно вытекает из лемм 1, 2 и свойств 1 и 2 п. 4.

2. Если y_Π и y_Λ – целые числа, то все прямоугольники из $\Lambda^\Delta \cup \Pi^\Delta$ расположены в узлах сетки.

Доказательство вытекает из определения множеств Λ^Δ и Π^Δ и лемм 1, 2.

3. Если $\Delta', \Delta'' \in B\Lambda^\Delta$ (или $B\Pi^\Delta$), то разность $y_{\Delta'} - y_{\Delta''}$ – четное целое число.

Утверждение вытекает из определения $B\Lambda^\Delta$ и $B\Pi^\Delta$, леммы 2 и четности высот всех прямоугольников.

4. $y_\Pi + y_\Lambda$ – четное число. В частности, если $y_\Pi = y_\Lambda$, то y_Π и y_Λ – целые числа. Это утверждение вытекает из леммы 3 и четности высот прямоугольников.

5. Если y_Π (а следовательно, и y_Λ) – не целое число, то множества $B\Pi^\Delta$ и $B\Lambda^\Delta$ не пересекаются.

Действительно, в этом случае дробные части чисел $y_\Pi/2$ и $y_\Lambda/2$ не равны 0 и 1/2. Но в силу свойства 4 п. 5 $\{y_\Pi/2\} + \{y_\Lambda/2\} = 0$ или 1. Следовательно, $\{y_\Pi/2\} \neq \{y_\Lambda/2\}$. Для любых прямоугольников $\Delta' \in B\Pi^\Delta$ и $\Delta'' \in B\Lambda^\Delta$ согласно свойству 3 п. 5 выполнены равенства

$$\left\{\frac{y_{\Delta'}}{2}\right\} = \left\{\frac{y_\Pi}{2}\right\}, \quad \left\{\frac{y_{\Delta''}}{2}\right\} = \left\{\frac{y_\Lambda}{2}\right\}.$$

Поэтому $\{y_{\Delta'}/2\} \neq \{y_{\Delta''}/2\}$ и, значит, $\Delta' \neq \Delta''$.

6. Вторая модификация оптимального решения. Пусть боковой прямоугольник Δ имеет нецелые координаты. Тогда на основании п. 5 связанные с ним числа y_Π и y_Λ также не являются целыми и не равны между собой, а множества $B\Pi^\Delta$ и $B\Lambda^\Delta$ не содержат общих прямоугольников. Кроме того, заметим, что согласно лемме 2 множества $B\Pi^\Delta$ и $B\Lambda^\Delta$ не содержат нижних прямоугольников. Предположим для определенности, что $y_\Lambda < y_\Pi$.

Для достаточно малого $\varepsilon > 0$ можно сдвинуть все прямоугольники из $B\Pi^\Delta$ вверх и влево на ε, а затем, не меняя положения этих прямоугольников, провести первую модификацию. При этом прямоугольники из $B\Pi^\Delta$ опустятся на ε вниз. Этот процесс можно продолжить до тех пор, пока не окажется выполненным одно из условий:

1. $y_\Lambda = y_\Pi$;

2. В множества Λ^Δ, Π^Δ, $B\Lambda^\Delta$ или $B\Pi^\Delta$ окажутся включенными новые прямоугольники.

В случае выполнения условия 1 на основании свойств 4 и 2 п. 5 можем заключить, что все прямоугольники из Λ^Δ и Π^Δ имеют целые координаты.

В случае выполнения условия 2 нужно повторить рассуждения пп. 5 и 6 для новых систем прямоугольников.

В результате этой процедуры мы добьемся целочисленности координат прямоугольника Δ (а одновременно и целочисленности всех связанных с ним цепочками боковых прямоугольников). Если после этой процедуры останутся боковые прямоугольники с нецелыми координатами, то ее нужно повторить. После того как все боковые прямоугольники будут расположены в узлах решетки, остальные прямоугольники в результате первой модификации расположатся в узлах решетки автоматически в силу свойств этой модификации и леммы 2. Теорема 3 доказана.

Теорема 4. В условиях теоремы 3 предположим, что найдено оптимальное расположение прямоугольников системы Ω в узлах решетки и при этом $y_{max} = \max\limits_{\Delta \in \Omega} y_\Delta$ – нечетное число. Тогда существует такое оптимальное решение, при котором координаты всех прямоугольников – нечетные (четные) числа.

Доказательство. Пусть для определенности число y_{max} нечетно. Сдвинем все прямоугольники с четными координатами на единицу вверх. Очевидно, что при этом взаимные пересечения прямоугольников не образуются. Для всех $\Delta \in \Omega$ координаты y_Δ станут нечетными и по-прежнему будут не больше y_{max}.

Сдвинем теперь каждый прямоугольник влево, насколько это возможно, не поменяв его вертикальной координаты.

Таким образом, каждый прямоугольник будет связан с левым боковым горизонтальной цепочкой, а следовательно, в силу лемм 1 и 2, горизонтальные координаты всех прямоугольников будут также нечетными. Теорема 4 доказана.

Часть вторая

ЭЛЕМЕНТЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ МИНИМАКСНЫХ ЗАДАЧ

В этой части книги приводится ряд АЛГОЛ-процедур решения задач равномерной аппроксимации и общих минимаксных задач. Все параграфы имеют одинаковую структуру: постановка задачи, метод решения, смысл формальных параметров процедуры, описание ее на языке АЛГОЛ-60, контрольный пример, некоторые рекомендации по использованию процедуры.

В отдельных случаях знаки логических операций, а также операций отношения заменены символами в соответствии с приведенными в следующей таблице:

Знаки операций	Символы
$\neg$	not
$\wedge$	and
$\vee$	or
$\neq$	ne (not equal)
$<$	lt (less than)
$\leqslant$	le (less or equal)
$>$	gt (greater than)
$\geqslant$	ge (greater or equal)

§ 2.1. Решение линейных задач аппроксимации при ограничениях

1°. Постановка задачи. Процедура approximation предназначена для решения следующей задачи. Пусть $F_1, F_2, \ldots, F_n$ – табличные значения некоторой функции F и $\psi_{k1}, \psi_{k2}, \ldots, \psi_{kn}$ – табличные значения в тех же узлах функций ψ_k, $k \in 1:p-1$. Требуется найти коэффициенты $x_1, x_2, \ldots, x_{p-1}$, удовлетворяющие ограничениям вида

$$\Sigma_{k=1}^{p-1} x_k \alpha_{ki} + \alpha_{pi} = 0, \quad i \in 1:m, \qquad (1)$$

$$\Sigma_{k=1}^{p-1} x_k g_{kj} + g_{pj} \geqslant 0, \quad j \in 1:r, \qquad (2)$$

для которых величина

$$\varphi_n(x_1, \ldots, x_{p-1}) = \max_{i \in 1:n} \left| F_i - \Sigma_{k=1}^{p-1} x_k \psi_{ki} \right| \qquad (3)$$

минимальна.

Задачу (1)-(3) можно интерпретировать как сеточную для задачи аппроксимации непрерывной на отрезке $[\alpha, \beta]$ функции $F(t)$ обобщенным полиномом $\Sigma_{k=1}^{p-1} x_k \psi_k(t)$ по заданной системе функций $\psi_k(t)$ с дополнительными условиями на коэффициенты (1),(2).

Предполагается, что все величины, входящие в описание задачи (1)-(3), кроме искомых коэффициентов, либо заданы, либо могут быть определены в процессе счета.

2°. Метод решения. Задача (1)-(3) очевидным образом сводится к задаче линейного программирования, причем в теле процедуры решается двойственная к ней задача, в которой число ограничений равно $p+1$, что, как правило, существенно меньше числа ограничений прямой задачи линейного программирования ($2n+m+r$). Для решения двойственной задачи применены симплексный метод обратных матриц и метод искусственного базиса для построения начального плана (см., например, [18]). При этом процедура не требует хранения матрицы ограничений в памяти машины, ее элементы могут вычисляться в процессе счета. Эти обстоятельства позволяют использовать процедуру со сколь угодно большими значениями параметров n, m, r, значение же параметра p невелико по смыслу самой задачи.

3°. Смысл формальных параметров. Входными параметрами процедуры являются p, n, m, r (смысл их тот же, что и в постановке задачи) и параметр eps, значение которого должно быть равно малому положительному числу, определяющему точность решения задачи линейного программирования (например, eps = 10^{-6}).

Параметры it, e, x, art – выходные. Значение переменной it равно числу проделанных итераций симплексного метода. Если переменная art в результате обращения к процедуре получила значение true, то ограничения (1),(2) несовместны, и задача решения не имеет. Если же art = false, то задача решение имеет, и оно содержится в массиве $x[0:p]$. При этом значения $x[k]$ для $k \in 1:p-1$ равны искомым коэффициентам ($x[p]=1$ и не играет роли

в формировании решения), а значение компоненты $x[0]$ равно $\varphi_n(x_1,\ldots,x_{p-1})$ на полученном решении.

Массив $e[1:p+1]$ содержит некоторую дополнительную информацию, связанную с альтернансными свойствами полученного решения. Компоненты массива e расположены в порядке возрастания модуля. Первые компоненты массива, которые по модулю не превосходят n, определяют точки максимума целевой функции на полученном решении, т.е. для $j_k = |e[k]| \leqslant n$

$$F_{j_k} - \Sigma_{i=1}^{p-1} x[i]\,\varphi_{ij_k} = \operatorname{sign}(e[k])\,x[0].$$

Следующие компоненты массива, для которых $1 \leqslant e[k]-n-m \leqslant r$, определяют те ограничения-неравенства из (2), которые на полученном решении обращаются в равенства, номера этих ограничений равны $e[k]-n-m$. Компонента $e[p+1]$ имеет значение $n+m+r+1$. Если есть и другие компоненты $e[k]$, равные $e[p+1]$, то в оптимальном базисе остались искусственные векторы условий, количество которых равно числу компонент из $e[1:p]$ со значением $e[p+1]$.

Параметр f определяет процедуру с описанием

```
procedure f(j, y); value j;
integer j; array y;
   < тело процедуры >
```

Тело процедуры по номеру $j \in 1:n+m+r$ восстанавливает столбец $y[1:p]$ матрицы ограничений решаемой задачи линейного программирования, при этом вектор $y[1:p]$ формируется по правилу: для всех $k \in 1:p$

$$y[k] = \begin{cases} \varphi_{kj}, & \text{если } j \in 1:n, \\ a_{k,j-n}, & \text{если } j \in n+1:n+m, \\ g_{k,j-n-m}, & \text{если } j \in n+m+1:n+m+r, \end{cases}$$

где $\varphi_{pj} = -F_j$ при $k = p$.

4°. Описание процедуры.

```
procedure approximation(p,n,m,r,f,eps,it,e,x,art);
value p,n,m,r,eps; procedure f;
integer p,n,m,r,it; boolean art;
real eps; array x; integer array e;
  begin integer array num [0:p];
  array u,y [0:p],g[1:p],b[0:p,0:p];
```



```
  integer av,i,k,j,l,q,s,w,w1,h,h1;
  real d,u0,d1,d2;
  av:=it:=0; q:=n+m+r+1; h:=1;
  for k:=0 step 1 until p do
  begin u[k]:=sign(k-p); num[k]:=q;
    for i:=0 step 1 until p do b[k,i]:=0;
    b[k,k]:=1; e[k+1]:=q end;
  art:=true;
  opt:l:=0; d:=-eps; w:=0;
  u0:=if art then u[0] else b[p,0];
  for k:=1 step 1 until av+1 do
  begin w1:=w; w:=abs(e[k]);
  for j:=w1+1 step 1 until w-1 do
  begin f(j,y); d1:=d2:=0;
    for i:=1 step 1 until p do
    begin d1:=d1+u[i]×y[i];
      d2:=d2+b[p,i]×y[i] end;
    if art then d2:=d1;
    if j le n+m then
    begin d1:=-abs(d2);
    if j le n then d1:=d1+u0;
    if d1 lt d then
      begin d:=d1; l:=if d2 gt 0 then -j else j;
      go to mkg end
    end
    else if d2 lt d then
      begin l:=j; d:=d2; go to mkg end;
    go to mkj;
mkg: s:=k; for i:=1 step 1 until p do g[i]:=y[i];
mkj: end j end k;
  if abs(e[1]) lt n+.5 and 2×u0 lt d then
  begin d:=2×u0; l:=-e[1]; s:=abs(l);
    f(s,g); s:=1 end;
  if l ne 0 then go to place;
  if art then begin art:=not art; h:=-1; go to opt end;
  for k:=0 step 1 until p do x[k]:=b[p,k];
  go to fin;
place: j:=sign(l); q:=-1; d1:=0;
  w:=if j×l le n then 1 else 0;
  for k:=0 step 1 until p do
  begin y[k]:=b[k,0]×w;
```

```
      for i:=1 step 1 until p do
      y[k]:=y[k]+b[k,i]×g[i]×j;
      h1:=h+(1-h)×sign(e[p+1]-num[k]);
      if y[k]×h1 ge 10-10 and k lt p then
      begin d2:=b[k,0]/y[k];
         if d2 lt d1 or q=-1 then
         begin q:=k; d1:=d2 end
      end
   end;
   it:=it+1;
   if q=-1 then begin art:=true; go to fin end;
   d2:=abs(num[q]); num[q]:=1; w:=0;
mke:w:=w+1;
   if d2 gt abs(e[w])+.5 then go to mke;
   if e[w]=e[p+1] then av:=av+1;
   k:=sign(w-s+.5); w1:=s-(1-k)/2;
   for i:=w step -k until w1+k do
   e[i]:=e[i-k]; e[w1]:=1;
   for i:=0 step 1 until p do
   begin b[q,i]:=b[q,i]/y[q];
      if art then u[i]:=u[i]-b[q,i]×d;
      for k:=0 step 1 until p do
      if k ne q then b[k,i]:=b[k,i]-b[q,i]×y[k]
   end;
   if av=p then art:=false; go to opt;
fin: end approximation
```

5°. Контрольный пример. Построить алгебраический полином

$$P(t) = x_1 + x_2 t^2 + x_3 t^4 + x_4 t^6 + x_5 t^8 + t^{10},$$

наименее уклоняющийся от нуля на промежутке $[0,0.8]$ при условиях

$$\begin{aligned} 128\,x_2 + 128\,x_3 + 120\,x_4 + 112\,x_5 + 105 &= 0, \\ 128\,x_1 - 16\,x_3 - 20\,x_4 - 21\,x_5 - 21 &= 0, \end{aligned} \qquad (4)$$

$$p'(t) \geqslant 0, \quad p''(t) \geqslant 0, \quad t \in [0.8, 1].$$

На отрезках $[0, 0.8]$ и $[0.8, 1]$ соответственно были выбраны системы точек

$$t_j = (j-1)0.05, \quad j \in 1{:}17,$$

$$t'_j = (j-1)0.05 + 0.8, \quad j \in 1{:}5.$$

Решалась следующая сеточная задача:

$$\begin{aligned} &\varphi_n(x_1,\ldots,x_5) = \max_{j\in 1:17} \left| t_j^{10} + \Sigma_{k=1}^{5} x_k t_j^{2k-2} \right| \to \min, \\ &\Sigma_{k=1}^{5} a_{ik} x_k + a_{i6} = 0, \quad i = 1,2, \\ &\Sigma_{k=2}^{5} (2k-2)(t'_j)^{2k-3} x_k + 10(t'_j)^9 \geqslant 0, \ j \in 1{:}5, \qquad (5) \\ &\Sigma_{k=2}^{5} (2k-2)(2k-3)(t'_j)^{2k-4} x_k + 90(t'_j)^8 \geqslant 0, \ j \in 1{:}5. \end{aligned}$$

Элементы матрицы $a[1{:}2,\ 1{:}6]$ определяются ограничениями (4). Значения входных параметров для этой задачи $p = 6$, $n = 17$, $m = 2$, $r = 10$, $eps = 10^{-5}$. Описание процедуры f :

```
procedure f(j,g); value j;
integer j; array g;
   begin integer i; real t,t1;
   if j ⩽ 17 then begin t:=(j-1)/20;
      g[1]:=1; t:=t×t;
      for i:=2 step 1 until 6 do g[i]:=g[i-1]×t;
      go to mkf end;
   j:=j-17;
   if j ⩽ 2 then begin
      for i:=1 step 1 until 6 do g[i]:=a[j,i];
   go to mkf end;
   j:=j-2; g[1]:=0;
   if j ⩽ 5 then begin t:=(j-1)×0.05+0.8;
      t1:=t; t:=t×t;
      for i:=2 step 1 until 6 do begin
         g[i]:=t1×(2×i-2); t1:=t1×t end;
```

```
      go to mkf end;
   j:=j-5;  t:=(j-1)×0.05+0.8;
   t1:=1; t:=t×t;
   for i:=2 step 1 until 6 do begin
      g[i]:=t1×(2×i-2)×(2×i-3);
      t1:=t1×t end;
mkf:end f
```

Теперь приведем текст всей программы для решения рассматриваемого примера, в котором описания процедур approximation и f будут даны условно.

```
begin
   integer n,m,r,p,it; real eps;
   read (n,m,r,p,eps);
   begin array  x[0:p], a[1:m, 1:p];
      integer array e[1:p+1]; boolean art;
      <описание процедуры  f >;
      <описание процедуры  approximation > ;
      read(a);
      approximation (p,n,m,r,f,eps,it,e,x,art);
      if art then print    ('ограничения задачи
         несовместны') else
      print (it,x,e)
   end
end
```

Вычисления по этой программе проводились на машине Odra-1204. Время счета составило 1,5 мин. Приведем результаты счета. Максимальное уклонение $\varphi_n = x[0] = 0.00531$. Искомые коэффициенты $x[1] = -0.002342$, $x[2] = 0.1584$, $x[3] = -1.050$, $x[4] = 2.485$, $x[5] = -2.581$. Число итераций 18. Компоненты вектора $e[1:p+1]$ имеют значения $e[1] = -8$, $e[2] = 16$, $e[3] = 17$, $e[4] = -18$, $e[5] = 19$, $e[6] = 25$, $e[7] = 30$. Отсюда видно, что в оптимальном базисе нет искусственных векторов условий. Точками максимума целевой функции являются $t_1 = (|e[1]| - 1) \times 0.05 = 0.35$, $t_2 = (|e[2]| - 1) \times 0.05 = 0.75$, $t_3 = (|e[3]| - 1) \times 0.05 = 0.8$. Значение $e[6]$, равное 25, указывает, что ограничения неравенства реализуются как равенства в точке $t = 0.8$, а именно: $P''(0.8) = 0$.

6°. Следует отметить, что процедура approximation позволяет решать задачу линейного программирования и с вырожденной матрицей ограничений. Это означает, что в исходной задаче функции φ_1, φ_2, ..., φ_{p-1}, а также ограничения (1), (2) могут быть линейно-зависимыми.

§ 2.2. Построение алгебраического полинома наилучшего приближения

1°. Постановка задачи. Задана таблица значений $y[1:m]$ функции $y(x)$ в узлах $x[1:m]$ (предполагается, что узлы упорядочены по возрастанию). Требуется определить вектор коэффициентов $a[0:n]$ алгебраического полинома наилучшего приближения функции $y(x)$, т.е. найти решение следующей задачи:

$$\max_{k \in 1:m} \left| y[k] - \Sigma_{i=0}^{n} a[i](x[k])^i \right| \to \min . \qquad (1)$$

Разумеется, для решения задачи (1) можно воспользоваться процедурой approximation из § 2.1. Однако приводимая ниже процедура remez более полно учитывает специфику задачи (1) и потому более эффективна.

2°. Метод решения. Процедура remez, реализующая R-алгоритм построения алгебраического полинома наилучшего приближения [39, с. 34-38], представляет собой усовершенствованный вариант программы из [69].

3°. Смысл формальных параметров. Параметры m, n, x, y те же, что и в (1). В массивах $in, jn[1:n+2]$ типа integer помещаются упорядоченные по возрастанию индексы точек, входящих соответственно в начальный и экстремальный базисы. Среди выходных параметров it - счетчик итерации, $rh0$ - величина наилучшего приближения, $a[0:n]$ - вектор коэффициентов полинома наилучшего приближения, $ax[1:n+2]$ - экстремальный базис ($ax[i] = x[jn[i]]$).

4°. Описание процедуры.

```
procedure remez(m,n,x,y,in) results:(it,rho,a,ax,jn);
value m,n,in; integer m,n,it; real rho;
integer array in,jn; array a,ax,x,y;
   begin integer i,j,j1,j2,k,r1,r2,iz;
   real h,z,w,kmax,kmin,min1,max1,u;
   boolean d; array ay,ah[1:n+2];
   rho:=0;
   it:=0;
   comment чебышевская интерполяция;
start:j:=-1; it:=it+1;
```

```
for i:=1 step 1 until n+2 do
begin ax[i]:=x[in[i]];
   ay[i]:=y[in[i]];
   j:=-j; ah[i]:=j  end;
for k:=2 step 1 until n+2 do
for i:=n+2 step -1 until k do
begin u:=ax[i]-ax[i-k+1];
   ay[i]:=(ay[i]- ay[i-1])/u;
   ah[i]:=(ah[i]-ah[i-1])/u  end;
h:=ay[n+2]/ah[n+2];
for i:=1 step 1 until n+1 do
   a[i-1]:=ay[i]-h×ah[i];
for i:=n-1 step -1 until 0 do
begin u:=ax[i+1];
   for k:=i step 1 until n-1 do
      a[k]:=a[k]-a[k+1]× u
end;
comment  проверка условия конца;
u:=rho; rho:=abs(h);
if rho le u then go to fin;
comment  преобразование базиса;
in[1]:=1; in[n+2]:=m; d:=true;
z:=0; max1:=h/2; min1:=-max1;
for i:=1 step 1 until n+1 do
begin kmax:=kmin:=0; j1:=r1:=r2:=in[i]; j2:=in[i+1];
   for j:=j1 step 1 until j2 do
   begin w:=0; u:=x[j];
      for k:=n step -1 until 0 do
      w:=w×u+a[k]; w:=w-y[j];
      if w gt kmax then begin kmax:=w; r1:=j end;
      if w lt kmin then begin kmin:=w; r2:=j end
   end j;
   if r1 gt r2 then
   begin j:=r1; r1:=r2; r2:=j;
      u:=kmax; kmax:=kmin; kmin:=u end;
   if sign(min1)=sign(kmax) then
   begin if abs(min1) lt abs(kmax) then
      begin jn[i]:=r1; jn[i+1]:=r2;
      max1:=kmax; min1:=kmin;
      if abs(z) gt abs(max1) then jn[i-1]:=iz;
```



```
      if d then begin h:=kmax; d:=false end
      end
      else
      begin max1:=min1;
      if abs(z) gt abs(kmin) then
      begin jn[i+1]:=iz; min1:=z end
      else
      begin jn[i+1]:=r2; min1:=kmin end
      end; z:=0
   end
   else
   begin max1:=min1;
      if abs(z) gt abs(kmax) then
      begin jn[i+1]:=iz; min1:=z end
      else
      begin jn[i+1]:=r1; min1:=kmax end;
      iz:=r2; z:=kmin
   end
end i;
if abs(z) gt abs(h) then
begin for j:=1 step 1 until n+1 do
      jn[j]:=jn[j+1];
   jn[n+2]:=iz
end;
for i:=1 step 1 until n+2 do
   in[i]:=jn[i];
go to start;
fin: end remez
```

5°. Контрольный пример. Построим полином наилучшего приближения седьмой степени функции $y(x)=\mathrm{abs}(x)$ на сетке $x[i]=-1+2(i-1)/1000$, $i\in 1{:}1001$.

В качестве начального базиса индексов возьмем массив $in[i]=125(i-1)+1$, $i\in 1{:}9$. Полный текст программы в этом случае будет выглядеть следующим образом:

```
begin integer m,n,it,i; real rho;
   read (m,n);
   begin integer array in,jn [1:n+2];
      array x,y[1:m], a[0:n], ax[1:n+2];
         <описание процедуры remez>;
      for i:=1 step 1 until m do
```

```
    begin x[i]:=-1+.002x(i-1);
      y[i]:=abs(x[i]) end;
    for i:=1 step 1 until n+2 do
      in[i]:=125x(i-1)+1;
    remez(m,n,x,y,in,it,rho,a,ax,jn);
    print(it,rho,a,ax)
  end
end
```

Результаты вычислений: $it = 4$; $rh0 = 0.0459284$;
$a[1] = a[3] = a[5] = a[7] = 0$,
$a[0] = 0.0459284$, $a[2] = 2.86804$,
$a[4] = -4.17842$, $a[6] = 2.31038$;
$ax[1] = -ax[9] = -1.000$,
$ax[2] = -ax[8] = -0.884$,
$ax[3] = -ax[7] = -0.572$,
$ax[4] = -ax[6] = -0.196$,
$ax[5] = 0$.

6°. В качестве начального базиса индексов $in[1:n+2]$ можно взять любой упорядоченный по возрастанию набор чисел из $1:m$, если только величина наилучшего приближения на соответствующем базисе $\{x[in[i]]\}$, $i \in 1:n+2$, положительна.

С помощью процедуры remez были составлены таблицы коэффициентов полиномов наилучшего приближения [73].

§ 2.3. Ускоренное выравнивание

1°. Постановка задачи. Рассматриваемая в этом параграфе процедура equalization позволяет построить алгебраический полином $P_n(t) = \sum_{i=0}^{n} b[i]\,t^i$ наилучшего приближения функции $y(t)$ на отрезке $[\alpha, \beta]$. В качестве начального приближения берется некоторое сеточное решение, которое можно получить с помощью процедуры remez из § 2.2.

2°. Метод решения. Процедура equalization реализует метод ускоренного выравнивания, описанный в п. 3° § 1.3.

3°. Смысл формальных параметров: n – степень полинома, iq – число неизвестных точек альтернанса (соответствует параметру r в формуле (1) § 1.3), $ix[1:iq]$ – индексы неизвестных точек альтернанса ($ix[k] \in 1:n+2$, $ix[k] < ix[k+1]$), $a[0:n]$ – начальные

значения коэффициентов полинома, $ax[1:n+2]$ – начальный базис (массивы a и ax формирует процедура remez), eps – точность вычисления коэффициентов полинома $P_n(t)$, it – счетчик итераций, h – число, модуль которого равен величине наилучшего приближения, $b[0:n]$ – вектор коэффициентов полинома наилучшего приближения, $bx[1:n+2]$ – экстремальный базис.

В теле процедуры equalization используются процедура-функция $y(t)$ и процедура $dy(t, y1, y2)$, присваивающая переменной $y1$ значение $y'(t)$, а переменной $y2$ – значение $y''(t)$.

4°. Описание процедуры.

```
procedure equalization(n,iq,ix,a,ax,eps)
results:(it,h,b,bx);
value n,iq,eps; integer n,iq,it; real eps,h;
integer array ix; array a,ax,b,bx;
  begin integer i,j,k; real t,p1,p2,y1,y2;
  array aw[0:n];
  procedure chebint;
  begin array cx,cy,ch[1:n+2];
    j:=-1;
    for i:=1 step 1 until n+2 do
    begin t:=bx[i]; j:=-j;
      cx[i]:=t; cy[i]:=y(t);
      ch[i]:=j
    end;
    for k:=2 step 1 until n+2 do
    for i:=n+2 step -1 until k do
    begin t:=cx[i]-cx[i-k+1];
      cy[i]:=(cy[i]-cy[i-1])/t;
      ch[i]:=(ch[i]-ch[i-1])/t
    end;
    h:=cy[n+2]/ch[n+2];
    for i:=1 step 1 until n+1 do
      b[i-1]:=cy[i]-hxch[i];
    for i:=n-1 step-1 until 0 do
    begin t:=cx[i+1];
      for k:=i step 1 until n-1 do
      b[k]:=b[k]-b[k+1]x t
    end
  end chebint;
  procedure dp;
  begin p1:=nxb[n]; p2:=0;
```

```
      for j:=n-1 step-1 until 1 do
      begin p1:=p1×t+j×b[j];
         p2:=p2×t+j×(j+1)×aw[j+1]
      end
   end dp;
   for i:=0 step 1 until n do
   begin b[i]:=aw[i]:=a[i];
      bx[i+1]:=ax[i+1] end;
   bx[n+2]:=ax[n+2]; it:=0;
start: it:=it+1;
   for i:=1 step 1 until iq do
   begin k:=ix[i]; t:=bx[k];
      dp; dy(t,y1,y2);
      t:=t-(y1-p1)/(y2-p2);
      bx[k]:=t  end;
   for i:=0 step 1 until n do  aw[i]:=b[i];
   chebint;
   for i:=0 step 1 until n do
      if abs(b[i]-aw[i]) gt eps then
      go to start
   end equalization
```

5°. Контрольный пример. Построим полином наилучшего приближения седьмой степени функции $y(t)=0.5\,(|t-0.5|+|t+0.5|)$ на отрезке $[-1,\ 1]$ с точностью $eps=_{10}-6$.

Начальное приближение было получено с помощью процедуры remez на равномерной сетке, состоящей из 21 точки. Массив $ax[1:n+2]$ оказался следующим: (-1.0, -0.9, -0.7, -0.5, -0.3, 0.5, 0.7, 0.9, 1.0). Нетрудно понять, что точки -1,0, -0.5, 0.5, 1.0 входят в экстремальный базис. Поэтому в данном случае $iq = 5$, $ix[1:iq] = (2,3,5,7,8)$. Мы не будем приводить полную программу. Ограничимся лишь описанием процедуры-функции $y(t)$ и процедуры $dy(t, y1, y2)$:

```
real procedure y(t); value t; real t;
y:=.5×(abs(t-.5)+abs(t+.5));
procedure dy(t,y1,y2); value t; real t,y1,y2;
begin y1:=if abs(t)<.5 then 0 else sign(t);
   y2:=0 end
```

Результаты вычислений: $it = 4$; $|h| = 0.02039$;

$b[1] = b[3] = b[5] = b[7] = 0$,
$b[0] = 0.49455$, $b[2] = -0.34658$,
$b[4] = 2.12263$, $b[6] = -1.29099$;

$bx[1] = -bx[9] = -1.000$, $bx[2] = -bx[8] = -0.898$,
$bx[3] = -bx[7] = -0.656$, $bx[4] = -bx[6] = -0.500$,
$bx[5] = -0.298$.

Интересно отметить, что нуль не является точкой альтернанса.

§ 2.4. Нахождение разбиений с равными уклонениями

1°. Постановка задачи. Пусть f – непрерывная на отрезке $[a, b]$ функция, n, m – натуральные числа. Обозначим через T совокупность разбиений $\tau = (\xi_0, \xi_1, \ldots, \xi_m)$ отрезка $[a,b]$, $a = \xi_0 < \xi_1 < \ldots < \xi_m = b$, а через $E(\xi_{i-1}, \xi_i)$ – наилучшее приближение функции f на $[\xi_{i-1}, \xi_i]$ алгебраическими полиномами степени не выше n:

$$E(\xi_{i-1}, \xi_i) = \min_{P} \max_{t\in[\xi_{i-1},\xi_i]} |f(t) - P(t)| =$$

$$= \max_{t\in[\xi_{i-1},\xi_i]} |f(t) - P_i(t)|.$$

Положим $E(\tau) = \max\limits_{i\in 1:m} E(\xi_{i-1}, \xi_i)$, $E^* = \inf\limits_{\tau\in T} E(\tau)$. Разбиение $\tau^* \in T$ называется оптимальным, если $E(\tau^*) = E^*$. Разбиение τ называется разбиением с равными уклонениями, если

$$\Phi_i(\xi_{i-1}, \xi_i, \xi_{i+1}) \stackrel{def}{=} E(\xi_{i-1}, \xi_i) - E(\xi_i, \xi_{i+1}) = 0 \quad \forall i \in 1:m-1. \quad (1)$$

Известно [11, 148], что для любой непрерывной функции f разбиение с равными уклонениями существует и является оптимальным. Требуется найти такое разбиение. Если рассматривать (1) как систему уравнений относительно $\xi_1, \ldots, \xi_{m-1}$ ($\xi_0 = a, \xi_m = b$), то задача по существу сводится к решению этой системы.

2°. Метод решения. Процедура equideviation реализует метод Ньютона решения системы (1). Для применимости метода Ньютона в данном случае необходимо выполнение следующих дополнительных условий: функция f непрерывно дифференцируема на $[a, b]$, разделенные разности $f[x_0, x_1, \ldots, x_{n+1}]$ неотрицательны при любых попарно различных $x_0, x_1, \ldots, x_{n+1}$ из $[a, b]$ и $E^* > 0$.

Опишем более подробно схему вычислений. Выбираем начальное разбиение $\tau_0 = (\xi_0, \xi_1, \ldots, \xi_m)$ и строим полиномы наилучшего приближения $P_i(t)$ функции f на $[\xi_{i-1}, \xi_i]$, $i \in 1:m$. При этом находим числа

$$a_i = \frac{\partial \Phi_i}{\partial \xi_{i-1}} = \frac{\partial E(\xi_{i-1}, \xi_i)}{\partial \xi_{i-1}},$$

$$b_i = \frac{\partial \Phi_i}{\partial \xi_i} = \frac{\partial E(\xi_{i-1}, \xi_i)}{\partial \xi_i} - \frac{\partial E(\xi_i, \xi_{i+1})}{\partial \xi_i},$$

$$c_i = \frac{\partial \Phi_i}{\partial \xi_{i+1}} = -\frac{\partial E(\xi_i, \xi_{i+1})}{\partial \xi_{i+1}},$$

$$\Phi_i = E(\xi_{i-1}, \xi_i) - E(\xi_i, \xi_{i+1}).$$

Далее методом прогонки решаем трехдиагональную систему уравнений

$$a_i \eta_{i-1} + b_i \eta_i + c_i \eta_{i+1} = -\Phi_i, \quad i \in 1:m-1,$$

$$\eta_0 = \eta_m = 0.$$

Компоненты $\xi_i^{(1)}$ нового разбиения τ_1 вычисляем по формуле $\xi_i^{(1)} = \xi_i + \eta_i$, $i \in 0:m$. Затем процесс повторяется. Вычисления заканчиваются, когда

$$|E(\xi_{i-1}, \xi_i) - E(\xi_i, \xi_{i+1})| < eps1 \quad \forall i \in 1:m-1. \qquad (2)$$

Полиномы наилучшего приближения находим следующим образом. На $[\xi_{i-1}, \xi_i]$ вводится равномерная сетка T_q, состоящая из $q+1$ точек, и решается дискретная задача наилучшего приближения с помощью алгоритма Валле-Пуссена [39, с. 33-34]. При этом учитывается, что концы отрезка входят в экстремальный базис [11]. Затем осуществляется выравнивание максимумов (см. § 1.3).

После того как найдены полином $P_i(t)$ и его экстремальный базис $\xi_{i-1} = x_0 < x_1 < \ldots < x_{n+1} = \xi_i$, вычисляем производные $\partial E/\partial \xi_{i-1}$ и $\partial E/\partial \xi_i$ по следующим формулам [11, 77]:

$$\frac{\partial E(\xi_{i-1}, \xi_i)}{\partial \xi_{i-1}} = \left\{ \prod_{k=1}^{n+1} (x_k - x_0)\, z[x_0, x_1, \ldots, x_{n+1}] \right\}^{-1} [f'(\xi_{i-1}) - P_i'(\xi_{i-1})],$$

$$\frac{\partial E(\xi_{i-1}, \xi_i)}{\partial \xi_i} = \left\{ \prod_{k=0}^{n} (x_k - x_{n+1})\, z[x_0, x_1, \ldots, x_{n+1}] \right\}^{-1} [f'(\xi_i) - P_i'(\xi_i)],$$

где $z[x_0, x_1, \ldots, x_{n+1}]$ - разделенная разность $(n+1)$-го порядка функции $z(x_k) = (-1)^k$, $k \in 0:n+1$.

3°. Смысл формальных параметров: m - число отрезков $[\xi_{i-1}, \xi_i]$, $i \in 1:m$; n - степень полиномов; q - число точек сетки T_q, уменьшенное на единицу (см. п. 2°); eps - параметр, определяющий окончание процесса выравнивания максимумов (выравнивание прекращается, когда $|f'(x_k) - P'(x_k)| < eps \ \forall k \in 1:n$); eps1 - параметр из (2); ksi, ksi1[0:m] - оптимальное и начальное разбиения; e[1:m] - массив величин наилучшего приближения при оптимальном разбиении ($e[i] = E(\xi_{i-1}, \xi_i)$); f, f1, f2 - идентификаторы процедур-функций с заголовком: < идентификатор функции > (u); value u; real u; эти процедуры-функции определяют соответственно аппроксимируемую функцию, ее первую и вторую производные; it - счетчик итераций.

4°. Описание процедуры.

```
procedure equideviation (m,n,q,eps,eps1,ksi1,f,f1,f2)
   results:(it,ksi,e); value m,n,q,eps,eps1;
integer m,n,q,it; real eps,eps1;
array ksi1,ksi,e; real procedure f,f1,f2;
   begin integer n1,i1,i,j,k,s;
   real h,rho,u,v,c,d,d1,b,b1,b2,b3,r1,l,x0;
   array x,y,z [0:n+1], p,r [1:m]; it:=0;n1:=n+1;
   for i:=0 step 1 until m do ksi[i]:=ksi1[ i];
mk:it:=it+1; i1:=1;
10:c:=ksi[i1-1]; d:=ksi[i1]; l:=(d-c)/q; h:=(d-c)/n1;
   for i:=0 step 1 until n1 do x[i]:=c+i×h;
   rho:=0; s:=1; go to t;
   comment алгоритм Валле-Пуссена;
11: u:=rho; rho:=abs(h); if rho≤u then go to 12:
   v:=0; d1:=d+1/2;
   for u:=c step l until d1 do
   begin b:=0;
      for k:=n step -1 until 0 do b:=b×u+y[k];
      b:=b-f(u);
      if abs(b)>abs(v) then begin v:=b;x0:=u end
   end;
   j:=sign(h);
   for i:=0 step 1 until n do
   begin j:=-j;
      if x[i]≤x0∧x0≤x[i+1] then
      begin if j×v>0 then x[i]:=x0 else x[i+1]:=x0  end
```

```
   end; go to t;
   comment выравнивание максимумов;
12: v:=0;
   for i:=1 step 1 until n do
   begin u:=x[i]; b2:=0; b1:=y[n]; b:=b1×u+y[n-1];
      for k:=n-2 step -1 until 0 do
      begin b2:=b2×u+b1; b1:=b1×u+b; b:=b×u+y[k] end;
      b:=f1(u)-b1; x[i]:=x[i]-b/(f2(u)-2×b2);
      if abs(b)>v then v:=abs(b)
   end;
   s:=if v ≥ eps then 2 else 0;
   comment чебышевская интерполяция;
t: j:=-1;
   for i:=0 step 1 until n1 do
   begin y[i]:=f(x[i]); j:=-j; z[i]:= j end;
   for k:=1 step 1 until n1 do
   for i:=n1 step -1 until k do
   begin u:=x[i]-x[i-k];
      y[i]:=(y[i]- y[i-1])/u; z[i]:=(z[i]-z[i-1])/u  end;
   h:=y[n1]/z[n1];
   for i:=0 step 1 until n1 do y[i]:=y[i]-h×z[i];
   for i:=n-1 step -1 until 0 do
   for k:=i step 1 until n-1 do
      y[k]:=y[k]-y[k+1]× x[i];
   if s=2 then go to 12; if s=1 then go to 11;
   e[i1]:=abs(h);
   comment нахождение производных функции E;
   b2:=b:=y[n]; b3:=b1:=0;
   for k:=n-1 step -1 until 0 do
   begin b1:=b1×c+b; b:=b×c+y[k];
      b3:=b3×d+b2; b2:=b2×d+y[k] end;
   u:=v:=1;
   for i:=1 step 1 until n1 do
   begin u:=u×(x[i]-c); v:=v×(x[i-1]-d) end;
   p[i1]:=(f1(d)-b3)/v/z[n1]; r[i1]:=(f1(c)-b1)/u/z[n1];
   i1:=i1+1; if i1 ≤ m then go to 10;
   comment метод прогонки и проверка критерия конца процедуры;
   r1:=b1:=v:=0;
   for i:=1 step 1 until m-1 do
   begin b:=e[i+1]-e[i]; if abs(b) > v then v:=abs(b);
```



```
      d:=r[i]× r1+p[i]-r[i+1]; b1:=(b-r[i]× b1)/d; r1:=p[i+1]/d;
      p[i]:=b1; r[i]:=r1  end;
   b1:=0;
   for i:=m-1 step -1 until 1 do
   begin b1:=r[i]× b1+p[i]; ksi[i]:=ksi[i]+b1  end;
   if v ≥ eps1  then go to mk
   end equideviation
```

5°. Контрольный пример. Рассмотрим задачу наилучшей кусочно-полиномиальной аппроксимации функции $f(t)=t^5$ на отрезке [0, 1] полиномами степени $n = 3$. Положим $m = 4$, $q = 20$, eps = $5_{10}-4$, eps1 = $_{10} - 7$, $ksi1$ = (0, 0.25, 0.50, 0.75, 1). Рабочая программа для определения разбиения с равными уклонениями следующая:

```
begin integer m,n,q,it; read(m,n,q);
   begin real eps, eps1; array ksi, ksi1[0:m], e[1:m];
      real procedure f(t); value t; real t; f:=t↑5;
      real procedure f1(t); value t; real t; f1:=5×t↑4;
      real procedure f2(t); value t; real t; f2:=20×t↑3;
      <описание процедуры equideviation>;
      read(eps,eps1,ksi1);
      equideviation(m,n,q,eps,eps1,ksi1,f,f1,f2,it,ksi,e);
      print(it,ksi,e)
   end
end
```

Результаты вычислений: it = 5, ksi = (0, 0.3I99, 0.5683, 0.79I6, I), e = (0.66I$_{10}$-4, 0.66I$_{10}$-4, 0.66I$_{10}$-4, 0.66I$_{10}$-4).

§ 2.5. Решение задачи дробно-рациональной аппроксимации

1°. Постановка задачи. Пусть $f_1, \ldots, f_{nb}$ - табличные значения некоторой функции f, а $\varphi_{j1}, \ldots, \varphi_{j,nb}$ - табличные значения функций φ_j, $j \in 0:m+n+1$, в тех же узлах.

Требуется найти коэффициенты $\alpha_0, \ldots, \alpha_{m+n+1}$, для которых

$$\mu = \max_{i \in 1:nb} \left| f_i - \frac{\Sigma_{j \in 0:m} \alpha_j \varphi_{ji}}{\Sigma_{j \in m+1:m+n+1} \alpha_j \varphi_{ji}} \right| \qquad (1)$$

минимально.

Введем нормировку искомой дроби:

$$\max\{|\alpha_j| \mid j\in 0:m+n+1\} = 1 . \qquad (2)$$

Задачу (1),(2) можно, в частности, интерпретировать как сеточную для аппроксимации непрерывной на отрезке функции $f(x)$ дробью

$$R_{m,n}(\alpha,x) = \frac{\Sigma_{j\in 0:m}\alpha_j\varphi_j(x)}{\Sigma_{j\in m+1:m+n+1}\alpha_j\varphi_j(x)} .$$

2°. Метод решения. Задача (1), (2) сводится к последовательности задач линейного программирования. Процедура approxr осуществляет решение этих задач с помощью модифицированного алгоритма метода последовательного улучшения плана с использованием искусственного базиса.

3°. Смысл формальных параметров. Входными параметрами процедуры approxr являются целые m , n , nb (смысл их ясен из постановки задачи) и вещественные неотрицательные $mb1$, $mb2$, eps , eps1 . Параметры $mb1$ и $mb2$ задают оценки соответственно снизу и сверху для μ , отвечающего решению задачи (1), (2). Очевидно, что $mb1 \geqslant 0$, а $mb2 \leqslant \max\limits_{i\in 1:nb} |f_i|$. В процессе работы процедуры эти оценки уточняются. Параметр eps1 задает требуемую точность совпадения оценок. Значение параметра eps - абсолютная погрешность при решении задач линейного программирования.

Процедура ef по номеру $i\in 1:nb$ формирует вектор $fi[0:m+n+1]$ по правилу $fi[j]=\varphi_{ji}$ и присваивает идентификатору $f1$ значение f_i . Процедура ef имеет заголовок

```
ef(i,f1,fi); value i; integer i; real f1; array fi;
```

Выходными параметрами процедуры являются max,it и массивы $\alpha[0:m+n+1]$ и $nu[-1:m+n+1]$, it - число итераций метода последовательного улучшения плана, необходимых для решения задачи. Однако если после обращения к процедуре значение it оказывается отрицательным, то следует заново обратиться к процедуре, предварительно увеличив eps ; max - значение μ на решении; массив α содержит решение задачи; массив nu служит для выяснения альтернансных свойств решения. Компонента $nu[k]$, не превосходящая по абсолютной величине nb , определяет номер узла $i=|nu[k]|$, для которого

$$f_i - \frac{\Sigma_{j\in 0:m}\alpha_j\varphi_{ji}}{\Sigma_{j\in m+1:m+n+1}\alpha_j\varphi_{ji}} = \operatorname{sign}(nu[k])\,max .$$

Остальные $nu[k]$ определяют номера j коэффициентов $\alpha[j]$, $j\in 0:m+n+1$, которые по модулю равны 1. При этом $j=|nu[k]|-nb-2$.

4°. Описание процедуры.

```
procedure approxr(n,m,nb,mb1,mb2,eps,eps1,
     ef,max,a,nu,it);
value n,m,nb,mb1,mb2,eps,eps1;
integer n,m,nb,it;
real mb1,mb2,eps,eps1,max;
array a; integer array nu;
procedure ef;
  begin integer i,i1,j,k,l,l1,n1,n2,c,r,p,gr,gr1;
  real u,x,w,f,s,t,mn1,mn2; boolean bul;
  n2:=m+n+1; n1:=n2+1;
  begin array bm[-1:n1,-1:n2], fi[0:n2], y[-1:n1];
    integer array nu1[-2:n1+1];
    procedure start;
    begin for i:=-1 step 1 until n2 do
      begin bm[n1,i]:=-1;
        nu[i]:=nu 1[i]:=nb+2+i;
        for j:=-1 step 1 until n2 do bm[i,j]:=0;bm[i,i]:=1 end;
      nu1[n1+1]:=c:=nu1[-2]:=nu[-1]:=0;
      bm[n1,-1]:=0; gr:=-1;
      gr1:=n1; nu1[gr1]:=nb+m+n+4
    end;
    procedure comp(st);
    value st; integer st;
    begin integer j; s:=t:=0;
      for j:=0 step 1 until m do
        t:=t+bm[st,j]×fi[j];
      for j:=m+1 step 1 until n2 do
        s:=s+bm[st,j]×fi[j]
    end;
    real procedure qual(st,i);
```

```
value st,i; integer st,i;
   qual:=bm[st,-1]+t×sign(i)+
      s×(max-sign(i)×f);
real procedure qual1(st,i);
value st,i; integer st,i;
begin integer j; j:=abs(i);
   qual1:=c+sign(i)×bm[st,j-nb-2]
end;
procedure opt(k,u);
real u; integer k;
begin u:=-eps; k:=0; l1:=1;
   for j:=-2 step 1 until gr-1 do
   begin i:=abs(nu1[j]); r:=abs(nu1[j+1]);
      if l1=0 then
      begin if i=r then
      begin l1:=1; go to en end else
      begin ef(i,f,fi); comp(p);
      optn(qual,-nu1[j]) end end else l1:=0;
      i:=i+1; r:=r-1;
      for i:=i step 1 until r do
      begin ef(i,f,fi); comp(p);
      for l:=-i,i do optn(qual,l) end;
en:end;
   j:=0; if gr gt -1 then j:=abs(nu1[gr-2]);
   if j ne nb then
      begin j:=gr-1; ef(nb,f,fi); comp(p);
      if nu1[j] ne -nb then optn(qual,-nb)
      else optn(qual,nb) end;
   for j:=gr step 1 until gr1-1 do
   for i:=abs(nu1[j])+1 step 1 until
      abs(nu1[j+1])-1 do
      begin optn(qual1,i); optn(qual1,-i)  end
end;
procedure optn(f,i);
value i; integer i; real procedure f;
begin w:=f(p,i);
   if bul then w:=-w;
   if w lt u then
      begin u:=w; k:=i; i1:=j+1 end
```



```
   end;
   procedure min;
   begin l:=-2; w:=0;
      if r le nb then ef(r,f,fi);
      for i:=-1 step 1 until n1 do
      begin if r le nb then
         begin comp(i); y[i]:=qual(i,k) end
         else y[i]:=qual1(i,k)-(if bul then 0 else 1);
         if y[i] lt 10-10 or i=n1 then go to en;
         s:=bm[i,-1]/y[i];
         if s lt w or l lt-1 then
         begin w:=s; l:=i end;
      en: end;
      if r gt nb then y[n1]:=y[n1]+1;
      if l=-2 then begin it:=-1; go to fin end
   end;
   procedure rec;
   begin for i:=-1 step 1 until n2 do
      begin bm[l,i]:=bm[l,i]/y[l];
         for j:=-1 step 1 until n1 do
         if j ne l then bm[j,i]:=bm[j,i]-bm[l,i]×y[j]
      end end;
   mn1:=mb1; max:=mn2:=mb2; it:=0;
mkstart: bul:=true; p:=-1; start;
mkopt: opt(k,u); it:=it+1;
   if u ge-eps then
   begin if gr1 ne n1+1 then
      begin it:=-2; go to fin end else
         if abs(mn1-mn2) le eps1 and bm[n1,-1]lt -eps then
         go to fin
         else if bm[n1, -1]lt -eps  then mn2:=max
         else mn1:=max;
      max:=(mn1+mn2)/2; go to mkstart
   end;
   r:=abs(k); min;
   if bul and l=-1 then
      begin bul:=false; c:=1; p:=n1 end;
   rec; l1:=abs(nu[l]);
   if l1=r then i1:=i1-1;
   if l1=0 then j:=gr1:=gr1+1 else
```

```
        begin j:=-2;
        mk0:j:=j+1; i:=abs(nu1[j]);
           if l1 gt i then go to mk0;
           if j le n1 then begin if nu1[j+1]=nu[l]
           then j:=j+1 end
        end;
     nu[l]:=k; k:=sign(i1-j-0.5);
     if k=1 or k=-1 and entier(abs(nu1[i1-1]))=r
        then i1:=i1-1;
     for i:=j step k until i1-k do
        nu1[i]:=nu 1[i+k];
     nu1[i1]:=nu[l];
     if i1 le gr and j gt gr or i1 ge gr and j it gr
        then gr:=gr-k;
     go to mkopt;
  fin: for i:=0 step 1 until n2 do
        a[i]:=bm[n1,i]
  end end approxr
```

5°. Контрольный пример. Будем аппроксимировать на $[c, d] = [-1, 1]$ функцию $f(x)=|x|$ дробью

$$R_{3,3}(\alpha,x) = \frac{\alpha_0 x^3+\alpha_1 x^2+\alpha_2 x+\alpha_3}{\alpha_4 x^3+\alpha_5 x^2+\alpha_6 x+\alpha_7}\ .$$

Выберем на промежутке $[-1, 1]$ сетку: $x_i = -1 + (i-1)\cdot 0.1$, $i \in 1:21$. Входные параметры процедуры approxr для решения этой сеточной задачи имеют значения $m = n = 3$, $nb = 21$, $mb1 = 0$, $mb2 = 0.2$, $eps = {}_{10}-5$, $eps1 = {}_{10}-3$. Полный текст программы в этом случае выглядит следующим образом:

```
begin integer q,n,m,nb,it,q1;
   real c,d,mb1,mb2,eps,eps1,max,h;
   read(c,d,n,m,mb1,mb2,nb,eps,eps1);
   q:=m+n+1;
   begin array a[0:q]; integer array nu[-1:q];
      real procedure ixs(i); value i; integer i;
         ixs:=c+(i-1)×h;
      real procedure f(x); value x; real x;
         f:=abs(x);
      procedure fi1(x,u); value x; real x; array u;
      begin integer i;
         u[m]:=u[q]:=1;
```



```
         for i:=q-1 step -1 until 0 do
            if i≠ m then u[i]:= u[i+1]×x
      end;
      procedure ef(i,f1,fi); value i; integer i;
      real f1; array fi;
      begin real x;
         x:=ixs(i); f1:=f(x); fi1(x,fi)
      end;
      <описание процедуры approxr>;
      h:=(d-c)/(nb-1);
      approxr(n,m,nb,mb1,mb2,eps,eps1,ef,max,a,nu,it);
      print(it,nu,a,max)
   end
end
```

Результаты счета: α =(-0.999,1,-0.013,0.013,-0.748,0.748, -0.311, 0.311), nu = (9, 13, -17, -5, -24, -21, 1, 21, -11), $max = 0.044$, $it = 238$.

Следовательно, для выбранной сетки максимальное отклонение $|f(x) - R_{3,3}(\alpha,x)|$, равное 0.044, достигается в узлах $x_1 = -1$, $x_5 = -0.6$, $x_9 = -0.2$, $x_{11} = 0$, $x_{13} = 0.2$, $x_{17} = 0.6$, $x_{21} = 1$.

Заметим, что присутствие номера 21 в массиве nu дважды (с разными знаками) указывает на случай вырождения: $x = 1$ является корнем числителя и знаменателя дроби.

6°. В [133, 145, 179] приведены другие программы нахождения наилучшего дробно-рационального приближения.

§ 2.6. Решение задачи синтеза электрического фильтра

1°. Постановка задачи (см., например, [64, 120]). Пусть $\overline{D}_j$, $j \in 1: nt_1$, - полосы пропускания, $\overline{D}_j$, $j \in nt_1+1: nt_1+nt_2$, - полосы задержания. В каждой полосе введена сетка и соответствующие сеточные области обозначены через D_j, $j \in 1: nt_1 + nt_2$. Пусть, кроме того, задана весовая функция $g(x)$, принимающая неотрицательные значения для всех точек этих сеточных областей.

Требуется определить коэффициенты дроби

$$R(cf,x)=\frac{P(cf,x)}{g(x)Q(cf,x)}=\frac{cf_0x^{m_1}+\ldots+cf_{m_1}}{g(x)(cf_{m_1+1}x^{m_2}+\ldots+cf_{m_1+m_2+1})},$$

для которой отношение

$$\mu^2=\frac{\max\limits_{\substack{x\in D_j\\ j\in 1:nt_1}}|R(cf,x)|}{\max\limits_{\substack{x\in D_j\\ j\in nt_1+1:nt_1+nt_2}}|R(cf,x)|} \qquad (1)$$

минимально.

Среди дробей, являющихся решением (1), будем искать ту, для которой

$$\max\{\,|cf_i|\;|\;i\in 0:m_1+m_2+1\,\}=1\,. \qquad (2)$$

2°. Метод решения. Решением задачи (1), (2) является дробь $R(cf^*,x)$, у которой знаменатель $Q(cf^*,x)$ имеет постоянный знак в каждой полосе пропускания, а числитель $P(cf^*,x)$ – постоянный знак в каждой полосе задержания. Не ограничивая общности, можно считать, что $Q(cf,x)>0$ в первой полосе пропускания, а $P(cf,x)>0$ в первой полосе задержания. Поскольку для оптимального решения комбинация знаков $Q(cf,x)$ и $P(cf,x)$ во всех полосах заранее не известна, приходится перебирать все возможные комбинации, число которых равно $2^{nt_1+nt_2-2}$. Например, в случае двух полос пропускания D_1, D_2 и одной полосы задержания D_3 возможны комбинации:

	D_1	D_2	D_3
1.	+	+	+
2.	+	–	+

Будем решать задачу (1), (2) с дополнительным ограничением (3), которое состоит в требовании знакопостоянства $Q(cf,x)$ и $P(cf,x)$ в соответствующих полосах при некоторой фиксированной комбинации знаков. Например, для приведенного ранее случая трех полос ограничение (3) для последней знаковой комбинации имеет вид

$$Q(cf,x)>0\,,\quad x\in D_1,$$

$$Q(cf,x)<0\,,\quad x\in D_2,$$

$$P(cf,x)>0\,,\quad x\in D_3\,.$$

Процедура filter предназначена для решения задачи (1) – (3). Для решения всей задачи следует $2^{nt_1+nt_2-2}$ раз воспользоваться процедурой, чтобы перебрать все возможные ограничения (3). Среди полученных дробей нужно выбрать ту, которой соответствует наименьшее значение μ^2, и взять ее в качестве решения исходной задачи.

Задача (1) – (3) решается сведением к последовательности задач линейного программирования. Для их решения используется модифицированный алгоритм метода последовательного улучшения плана с применением метода искусственного базиса.

3°. Смысл формальных параметров. Входными параметрами процедуры являются:

а) массивы $m[1:2]$, $nt[1:2]$, смысл которых ясен из постановки задачи;

б) массив $nbt[1:nt[1]+nt[2]]$, первые $nt[1]$ компонент которого задают число точек в каждой полосе пропускания, а последующие $nt[2]$ компонент – в каждой полосе задержания;

в) массив $sgt[1:nt[1]+nt[2]]$, первые $nt[1]$ компонент которого совпадают со знаком $Q(cf,x)$ в соответствующей полосе пропускания, а последующие $nt[2]$ компонент – со знаком $P(cf,x)$ в полосах задержания;

г) массив $mbt[1:2]$, у которого $mbt[1]$ представляет собой оценку снизу, а $mbt[2]$ – оценку сверху для μ (компоненты массива неотрицательны). В процессе работы процедуры эти оценки уточняются;

д) вещественные **epsI** и **eps**, определяющие соответственно требуемую точность совпадения этих оценок и погрешность решения задач линейного программирования;

е) процедура-функция g, которая служит для вычисления $g(x)$, с заголовком g(x); value x; real x;

ж) процедура-функция ixs, которая служит для вычисления очередной точки сетки x из D_i, $i\in 1:nt_1+nt_2$. Функция имеет заголовок

```
ixs(i,j,j1); value i,j,j1; integer i,j,j1;
```

здесь j - номер точки; i - номер полосы, $i \in 1:nt_1+nt_2$; $1+j1$ - номер первой точки в i-й полосе. Причем сначала нумеруются точки полос пропускания, затем точки полос задержания. Точки нумеруются подряд, но при переходе из одной полосы в другую один номер пропускается. Таким образом, точки полосы i имеют номера

от $\Sigma_{k\in 1:i-1} nbt[k]+i$ до $\Sigma_{k\in 1:i} nbt[k]+i-1$.

В частности, в рассмотренном ранее примере точки сетки имеют номера:

для D_1 от 1 до $nbt[1]$,

для D_2 от $nbt[1]+2$ до $nbt[1]+nbt[2]+1$,

для D_3 от $nbt[1]+nbt[2]+3$ до $nbt[1]+nbt[2]+nbt[3]+2$.

Выходными параметрами процедуры являются:

а) массив $cf[0:m[1]+m[2]+1]$, компоненты которого равны коэффициентам дроби $R(cf^*, x)$;

б) max - значение μ на решении;

в) it - число шагов метода последовательного улучшения плана, необходимых для решения задачи (1) - (3). Однако после обращения к процедуре filter значение параметра it может оказаться отрицательным. Это означает, что выбранное значение параметра eps мало́. Для получения решения задачи (1)-(3) следует увеличить eps и снова обратиться к процедуре;

г) массив $nu[-1:m[1]+m[2]+1]$, который дает дополнительную информацию об альтернансных свойствах решения. Компоненты $nu[k]$, не превосходящие по абсолютной величине $N=\Sigma_{k\in 1:nt[1]+nt[2]} nbt[k]+nt[1]+nt[2]$, определяют точку $x\in D_j$, $j\in 1:nt[1]+nt[2]$, в которой $|R(cf^*, x)|=max$, если эта точка из полосы пропускания, и $|1/R(cf^*, x)|=max$, если она из полосы задержания. Компоненты $nu[k]$, превосходящие N по абсолютной величине, указывают, какие коэффициенты $cf[i]$, $i\in 0:m[1]+m[2]$, по модулю равны 1. При этом номер i коэффициента равен $|nu[k]|-N-1$.

4°. Описание процедуры.

```
procedure filter(m,nt,nbt,sgt,mbt,eps,eps1,
   g,ixs,it,max,cf,nu);
value eps,eps1;
integer it; real max,eps,eps1;
array mbt,cf; integer array m,nt,nbt,sgt,nu;
```



```
real procedure ixs,g;
   begin boolean bul;
   real mn1,mn2,u,x,w,s,s1;
   integer n1,n2,n3,n4,i,i1,i2,j,j1,j2,
      k,k1,k2,l,l1,p,p1,q,r,gr,nb,c;
   q:=if m[1]gt m[2] then m[1] else m[2];
   n1:=m[1]+m[2]+2; n2:=nt[1]+nt[2]+1;
   n3:=n1+n2+1; n4:=n2-1; p1:=n1-1;
   nb:=n2; for i:=1 step 1 until n4 do
         nb:=nb+nbt[i];
   begin array u1[0:q],y[-1:n1],
         bm[-1:n1,-1:p1];
      integer array nu1[0:n3],rz[1:n2];
      procedure comp;
      begin integer i;
         u1[0]:=1; for i:=1 step 1 until q do
            u1[i]:=u1[i-1]×x
      end;
      procedure comp1(st);
      value st; integer st;
      begin integer i; s:=s1:=0;
         for i:=0 step 1 until m[1]do
            s:=s+bm[st,i]×u1[m[1]-i];
         for i:=0 step 1 until m[2]do
            s1:=s1+bm[st,m[1]+i+1]×u1[m[2]-i];
         s1:=s1×g(x)
      end;
      real procedure f(i0,i1,i,st);
      value i0,i1,i,st; integer i,i0,i1,st;
         f:=bm[st,-1]+sgt[i1]×
            (if i0=1 then sign(i)× s+max×s1
            else sign(i)× s1+max×s);
      procedure est(t,i0);
      value t,i0; integer t,i0;
      begin
         if t gt 0 then
            begin comp; comp1(p) end;
         w:=f(i,i1,i0,p)× sign(p);
         if w lt u then
            begin u:=w; k:=i0; r:=j; i2:=i1; k2:=j2
```

```
            end end;
        procedure rec(d1,d2,e1,e2);
        value d1,d2,e1,e2; integer d1,d2,e1,e2;
        begin
          for i:=i1 step k until i2-k do
            nu1[i]:=nu1[i+k];
          if d1 lt nb then
          begin j1:=e1; if d2 lt nb then j2:=e2-1
            else begin gr:=gr-k; j2:=n2 end;
            for i:=j1 step 1 until j2 do
            rz[i]:=rz[i]-k
          end end;
      mn2:=max:=mbt[2]; mn1:=mbt[1]; it:=0;
mkstart: bul:=true; gr:=n4; p:=-1;
      for i:=-1 step 1 until p1 do
      begin nu1[gr+i+2]:=nb+i+1;
        nu[i]:=nb+i; bm[n1,i ]:=-1;
        for j:=-1 step 1 until p1 do
          bm[i,j]:=0; bm[i,i]:=1
      end;
      bm[n1,-1]:=0; nu1[gr+p1+3]:=nb+n1+1;
      nu[-1]:=nu1[0]:=rz[1]:=c:=0;
      for i:=1 step 1 until n4 do
        begin rz[i+1]:=i; nu1[i]:=nbt[i]+nu1[i-1]+1
        end;
mkopt:u:=-eps; k:=j2:=0;
      l1:=nt[1]; l:=1;
      for i:=1,2 do
      begin for i1:=l step 1 until l1 do
      begin for j:=rz[i1]step 1 until rz[i1+1]-1 do
        for j1:=abs(nu1[j])+1 step 1 until abs(nu1[j+1])-1 do
        begin x:=ixs(i1,j1,j2);
          for k1:=j1,-j1 do est(k1,k1)
        end;
        for j:=rz[i1]+1 step 1 until rz[i1+1]-1 do
        begin k1:=abs(nu1[j]); j1:=abs(nu1[j+1]);
          if k1 ne j1 then
          begin x:=ixs(i1,k1,j2);
          est(1,-nu1[j]) end else j:=j+1
        end;
```



```
        j2:=j2+nbt[i1]+1
      end;
      l:=nt[1]+1; l1:=n4
      end;
      for j:=gr step 1 until n3-1 do
      for i:=abs(nu1[j])+1 step 1 until abs(nu1[j+1])-1 do
      begin x:=bm[p,i-nb];
        for l:=i,-i do
        begin w:=(c+sign(l)×x)×sign(p);
          if w lt u then begin u:=w; k:=l; r:=j end
        end end;
      if u ge -eps then
      begin if nu[-1]=0 then
        begin it:=-1; go to fin end else
        if abs(mn1-mn2) le eps1 and bm[n1,-1]lt -eps then go to fin
          else if bm[n1,-1]lt-eps then mn2:=max else
          mn1:=max; max:=(mn1+mn2)/2;
        go to mkstart
      end;
      k1:=abs(k); if k1 lt nb then
        begin x:=ixs(i2,k1,k2); comp end;
      l:=-2; i1:=if i2 le nt[1]then 1 else 2;
      for j:=-1 step 1 until n1 do
      begin if k1 ge nb then y[j]:=bm[j,k1-nb]×sign(k)
        else begin comp1(j); y[j]:=f(i1,i2,k,j) end;
        if y[j] ge 10^-10 and j ne n1 then
          w:=bm[j,-1]/y[j] else go to mk1;
        if l=-2 or w lt u then
          begin u:=w; l:=j end;
      mk1: end;
      if l=-2 then begin it:=-2; go to fin end;
      if k1 ge nb then y[n1]:=y[n1]+1;
      if bul and l=-1 then
        begin bul:=false; c:=1; p:=n1 end;
mkrec: for j:=-1 step 1 until p1 do
      begin bm[l,j]:=bm[l,j]/y[l];
        for i:=-1 step 1 until n1 do
        if i ne l then bm[i,j]:=bm[i,j]-bm[l,j]×y[i]
      end;
```

```
      k2:=abs(nu[l]); j2:=if k1 lt nb then i2+1 else 0;
      i2:=if k1=k2 then r else r+1;
      j1:=0; if k2 ne 0 then
      begin i1:=0; mk2:i1:=i1+1;
         i:=abs(nu1[i1]); if k2 gt i then go to mk2;
         if i1 lt n3 then begin
            if nu1[i1+1]=nu[l] then i1:=i1+1 end;
         if k2 lt nb then
         begin j1:=1; mk3:j1:=j1+1;
            if i1 gt rz[j1] then go to mk3 end
      end else
      begin i1:=n3; k2:=nb+n1+1 end;
      it:=it+1; nu[l]:=k;
      k:=sign(k1-k2-0.5);
      if k=1 then
         begin i2:=i2-1; rec(k2,k1,j1,j2) end else
            rec(k1,k2,j2,j1);
      nu1[i2]:=nu[l]; go to mkopt;
fin: for i:=0 step 1 until p1 do
      cf[i]:=bm[n1,i]
end end filter
```

5°. Контрольный пример. Рассмотрим задачу с двумя полосами пропускания D_1 и D_2 и двумя полосами задержания D_3 и D_4. Зададим точки сетки:

в D_1: $x_j = 1.3 + (j - 1) \times 0.1$, $j \in 1:8$,
в D_2: $x_j = 4.5 + (j - 10) \times 0.1$, $j \in 10:20$,
в D_3: $x_j = 0.5 + (j - 22) \times 0.1$, $j \in 22:27$,
в D_4: $x_j = 2.5 + (j - 29) \times 0.1$, $j \in 29:39$.

Будем искать оптимальную дробь среди дробей вида

$$R(cf,x) = \frac{P(cf,x)}{g(x)Q(cf,x)} = \frac{cf_0x^3+cf_1x^2+cf_2x+cf_3}{\sqrt{x}\,(cf_4x^2+cf_5x+cf_6)}.$$

Для $Q(cf,x)$ в D_1 и D_2 и $P(cf,x)$ в D_3, D_4 возможны следующие комбинации знаков:

	D_1	D_2	D_3	D_4
1.	+	+	+	+
2.	+	-	+	+

3.	+	+	+	-
4.	+	-	+	-

Следовательно, для решения задачи (1), (2) придется четыре раза воспользоваться процедурой filter с разными значениями массива $sgt[1:4]$.

Для примера покажем, как воспользоваться процедурой filter в случае первой комбинации знаков. Входные параметры имеют значения $m[1:2] = (3, 2)$, $nt[1:2] = (2, 2)$, $nbt[1:4] = (8, 11, 6, 11)$, $mbt[1:2] = (0, 1)$, $sgt[1:4] = (1, 1, 1, 1)$, $eps = {}_{10}{-4}$, $eps1 = {}_{10}{-2}$.

Приведем полный текст программы для этого случая. В программе используется массив $c[1:4]$, значения компонент которого совпадают со значениями первых точек каждой полосы: $c = (1.3, 4.5, 0.5, 2.5)$.

```
begin
   <описание процедуры filter>;
   integer array m,nt[1:2],
      nbt,sgt[1:4],nu[-1:6];
   array mbt[1:2],c[1:4],cf[0:6];
   real procedure ixs(i,j,j1);
   value i,j,j1; integer i,j,j1;
      ixs:=c[i]+0.1×(j-j1-1);
   real procedure g(x); value x; real x;
      g:=sqrt(x);
   real max,eps,eps1; integer it;
   read(c,nt,m,nbt,sgt,mbt,eps,eps1);
   filter(m,nt,nbt,sgt,mbt,eps,eps1,g,
      ixs,it,max,cf,nu);
print(max,cf,nu,it) end
```

В результате работы программы получено решение:

1.

$$cf = \begin{pmatrix} -0.0532 \\ 0.4379 \\ -1.0000 \\ 0.6868 \\ 0.0149 \\ -0.0541 \\ 0.0773 \end{pmatrix}, \qquad nu = \begin{pmatrix} 4 \\ -29 \\ -10 \\ 20 \\ -8 \\ 43 \\ -27 \\ -1 \end{pmatrix},$$

$max = 0.80$, $it = 191$.

Меняя только значение параметра sgt, получаем решения и для оставшихся комбинаций знаков:

2.

$$cf = \begin{pmatrix} -0.0517 \\ 0.4293 \\ -1.0000 \\ 0.7039 \\ 0.0284 \\ -0.1171 \\ 0.1312 \end{pmatrix}, \quad nu = \begin{pmatrix} -27 \\ 5 \\ -8 \\ 20 \\ -1 \\ 43 \\ -39 \\ -10 \end{pmatrix},$$

$max = 0.52$, $it = 202$.

3.

$$cf = \begin{pmatrix} 0.1689 \\ -1.0000 \\ 0.5641 \\ 1.0000 \\ 0.0416 \\ -0.1307 \\ 0.6115 \end{pmatrix}, \quad nu = \begin{pmatrix} 42 \\ 8 \\ 29 \\ 10 \\ -20 \\ -44 \\ -27 \\ -1 \end{pmatrix},$$

$max = 0.73$, $it = 192$.

4.

$$cf = \begin{pmatrix} 0.0465 \\ -0.4114 \\ 0.4534 \\ 0.1334 \\ -0.2681 \\ 1.0000 \\ -0.6223 \end{pmatrix}, \quad nu = \begin{pmatrix} -1 \\ -27 \\ -10 \\ -39 \\ -46 \\ 22 \\ 8 \\ 29 \end{pmatrix},$$

$max = 0.59$, $it = 166$.

Следовательно, решением задачи (1), (2) является дробь $R(cf,x) = \dfrac{-0.0517\,x^3 + 0.4293\,x^2 - x + 0.7039}{\sqrt{x}\,(0.0284\,x^2 - 0.1171\,x + 0.1312)}$, полученная для второй комбинации знаков, так как ей соответствует наименьшее значение $max = 0.52$.

§ 2.7. Определение расстояния от конуса до многогранника

1°. Постановка задачи. Пусть в пространстве E_n заданы многогранник L и конус K:

$$L = \{ U = \Sigma_{j=1}^{s} x_j z_j \mid x_j \geq 0, \ \Sigma_{j=1}^{s} x_j = 1 \},$$

$$K = \{ V = -\Sigma_{j=s+1}^{m} x_j z_j \mid x_j \geq 0 \}, \quad s \leq m,$$

где $z_j = (z_{1j}, z_{2j}, \ldots, z_{nj})$, $j \in 1{:}m$. Требуется определить

$$\min_{U \in L,\, V \in K} \| U - V \|^2. \qquad (1)$$

Эта задача равносильна следующей задаче квадратичного программирования:

$$x_0 \stackrel{def}{=} \Sigma_{k=1}^{n} \left(\Sigma_{j=1}^{m} x_j z_{kj} \right)^2 \to \min,$$

$$\Sigma_{j=1}^{s} x_j = 1, \qquad (2)$$

$$x_j \geq 0 \quad \forall j \in 1{:}m.$$

Допустим, что $x_0^*, x_1^*, \ldots, x_m^*$ — решение задачи (2). Если $x_0^* > 0$, то множества L и K не пересекаются. В этом случае можно определить вектор $g^* = -g/\|g\|$, где $g = \Sigma_{j=1}^{m} x_j^* z_j$. Вектор g^* используется при решении минимаксных задач как направление наискорейшего спуска.

2°. Метод решения. Приводимая ниже процедура distkm позволяет решать поставленную задачу. В случае $s = m$ distkm определяет квадрат расстояния от нуля до многогранника L.

В процедуре используется метод Вулфа (см. [63]), причем специфика задачи позволила значительно упростить алгоритм метода.

3°. Смысл параметров процедуры. Входными параметрами являются n, m, s, вещественный массив $z[1{:}n, 1{:}m]$, смысл которых тот же, что и в постановке задачи, и параметр eps, который определяет точность вычислений (например, $eps = 10^{-6}$).

Выходные параметры: it — число итераций, $x[0{:}m]$ — вещественный массив, значения компонент которого при $k \in 1{:}m$ равны искомым коэффициентам, а значение $x[0]$ равно минимуму целевой функции задачи (2), т.е. величине (1).

Если в результате выполнения оператора процедуры

distkm(z,n,m,s,eps,x,it)

переменная *it* получила значение -1, следует заново обратиться к процедуре, предварительно увеличив eps (например, в 10 раз).

4°. Описание процедуры.

```
procedure distkm(z,n,m,s,eps,x,it);
value n,m,s,eps; array z,x;
integer n,m,s,it; real eps;
   begin array b[0:m,0:m], d[0:m], y[1:m];
   integer array num[0:m]; real t,t1,w;
   integer i,j,k,l,p,q;
   it:=0; num[0]:=1;
   w:=x[0]:=t1:=0;
   for j:=1 step 1 until m do
   begin x[j]:=b[0,j]:=0;
      for i:=1 step 1 until m do b[i,j]:=0;
      b[j,j]:=-1; t:=0;
      for i:=1 step 1 until n do t:=t+z[i,j] × z[i,1];
      t:=b[j,0]:=2×t;
      if j le s then t:=t-b[1,0]else b[j,1]:=-1;
      if j le 2 or t lt w then begin w:=t; k:=j end;
      num[j]:=-j
   end; b[0,0]:=1; t1:=b[1,0];
   x[1]:=1; num[k]:=0; num[1]:=-(m+1);
   q:=if k gt s then 1 else 0; j:=m;
mk0: if j=s then w:=w+t1;
   b[j,0]:=b[j,0]-w; b[j,k]:=1;
   b[j,1]:=b[j,1]+q;
   j:=j-1; if j ge 1 then go to mk0;
   x[0]:=b[k,0]:=-w; b[k,1]:=b[k,1]+q;
   if w-t1 ge -eps then go to fin;
iterx: for j:=1 step 1 until m do
   begin y[j]:=0;
      for i:=1 step 1 until n do
         y[j]:=y[j]+z[i,k]×z[i,j];
      y[j]:=2×y[j]
   end;
iterv: p:=-1; w:=0; it:=it+1;
   for j:=0 step 1 until m do
```



```
   begin
      if k lt 0 then d[j]:=-b[j,-k] else
         begin d[j]:=if k gt s then 0 else b[j,0];
         for i:=1 step 1 until m do d[j]:=d[j]+y[i]×b[j,i]
         end;
      if num[j]=0 or d[j] le 10-10 then go to mk;
      t:=b[j,0]/d[j];
      if p lt 0 or t lt w then begin w:=t; p:=j end;
mk: end;
   if p lt 0 then begin it:=-1; go to fin end;
   l:=k; k:=-num[p];
   if k le m then x[abs(k)]:=0;
   num[p]:=l;
   for j:=0 step 1 until m do
   begin b[p,j]:=b[p,j]/d[p];
      for i:=0 step 1 until m do
         if i ne p then b[i,j]:=b[i,j]-b[p,j]×d[i];
      if num[j] ge 0 then x[num[j]]:=b[j,0]
   end;
   if p ne 1 and b[1,0] gt eps then go to
      if k lt 0 then iterv else iterx;
fin: x[0]:= abs(x[0]/2)end distkm
```

5°. Контрольный пример. Рассмотрим контрольный пример из § 2.1. Заменим в нем промежуток минимизации [0, 0.8] промежутком [0.4, 0.8]. Решение X^* задачи аппроксимации на [0, 0.8] не является решением для [0.4, 0.8], но удовлетворяет всем ограничениям задачи, поэтому X^* можно использовать как некоторую допустимую точку в градиентном методе решения задачи для промежутка [0.4, 0.8]. Воспользуемся процедурой distkm для определения направления наискорейшего спуска из точки X^*.

Входным параметрам присваиваем значения $n = 5$, $s = 2$, $m = 5$, $eps = 10^{-8}$, $Z[1:5, 1:5] =$

$$= \begin{pmatrix} 1.00000 & 1.00000 & 0.00000 & 1.28000 & 0.00000 \\ 5.62500_{10}-1 & 6.40000_{10}-1 & -1.28000_{10}2 & 0.00000 & 2.00000 \\ 3.16406_{10}-1 & 4.09600_{10}-1 & -1.28000_{10}2 & -1.60000_{10}1 & 7.68000 \\ 1.77978_{10}-1 & 2.62144_{10}-1 & -1.20000_{10}1 & -2.00000_{10}1 & 1.22880 \\ 1.00113_{10}-1 & 1.67772_{10}-1 & -1.12000_{10}2 & -2.10000_{10}1 & 1.46801 \end{pmatrix}.$$

Результаты счета: $it = 4$, $x[0:m] = (1.002, 0, 1, 0.005293,$

0, 0.03016). Отсюда следует, что квадрат расстояния от L до K равен 1.002 и g = (1, 0.02284, -0.0362, -0.00258, 0.01779).

6°. Значения параметров n и m не могут быть большими, так как в теле процедуры используются массивы $z[1:n,1:m]$ и $b[0:m, 0:m]$, которые должны храниться в оперативной памяти машины.

§ 2.8. Нахождение направления наискорейшего спуска в минимаксных задачах

1°. Постановка задачи. Пусть на E_n заданы непрерывно дифференцируемые функции $f_i(X)$, $i\in I=1:N$, $h_j(X)$, $j\in J=N+1:N_1$. Рассмотрим задачу минимизации функции $\varphi(X)=\max\limits_{i\in I} f_i(X)$ на множестве

$$\Omega:=\{\, X\in E_n \mid h_j(X)\leqslant 0,\ j\in J \,\}.$$

Предположим, что множество Ω выпукло и выполнено условие Слейтера.

Пусть $X\in\Omega$. Направление наискорейшего спуска функции φ на множестве Ω в точке X можно находить, решая задачу квадратичного программирования следующего вида [9]:

$$(V,V)\to\min \qquad (1)$$

при ограничениях

$$(A_i,V)\geqslant 1,\quad i\in R(X), \qquad (2)$$

$$(A_i,V)\geqslant 0,\quad i\in Q(X), \qquad (3)$$

где $A_i=(1;\ f'_{iX}(X))\quad \forall\, i\in R(X)$,

$A_i=(0;\ h'_{iX}(X))\quad \forall\, i\in Q(X)$.

Предположим, что в рассматриваемой точке X система векторов $\{A_i\}$, $i\in R(X)\cup Q(X)$, линейно-независима.

Пусть $V(X)=(v_0(X), v_1(X),\dots, v_n(X))$ - решение задачи (1), $g^*(X)=(1/v_0(X))(v_1(X), v_2(X),\dots, v_n(X))$. Тогда вектор $g(X)=-g^*(X)/\|g^*(X)\|$ является направлением наискорейшего спуска функции φ на Ω в точке X, причем

$$\frac{\partial\varphi(X)}{\partial g(X)}=-\|g^*(X)\|=-\sqrt{\frac{1}{v_0(X)}-1}.$$

Кроме того, вектор $g^*(X)$ допускает представление

$$g^*(X)=\Sigma_{i\in R(X)}\beta_i^* f'_{iX}(X)+\Sigma_{i\in Q(X)}\beta_i^* h'_{iX}(X), \qquad (4)$$

$$\Sigma_{i\in R(X)}\beta_i^*=1,\quad \beta_i^*\geqslant 0\quad \forall i\in R(X)\cup Q(X).$$

2°. Метод решения. Для решения задачи (1) можно применять любые методы квадратичного программирования (см., например, §~2.7). Но в задаче (1) минимизируется не квадратичная функция общего вида, а квадрат нормы вектора. Поэтому для ее решения удобно применить метод наискорейшего спуска. В приводимой ниже процедуре down реализуется метод наискорейшего спуска для решения задачи (1).

3°. Смысл формальных параметров процедуры down. Входные параметры: p - число индексов в $R(X)$; m - число индексов в $R(X)\cup Q(X)$; n - размерность пространства E_n; $a[1:m,1:n]$ - матрица, в первых p строках которой расположены векторы $f'_{iX}(X)$, $i\in R(X)$, а в $m-p$ остальных - векторы $h'_i(X)$, $i\in Q(X)$.

Выходными параметрами процедуры являются $mp[1:n]$ - вектор, равный $-g(X)$; $mt[1:m]$ - вектор коэффициентов β_i^* в представлении (4); u - число положительных коэффициентов β_i^*; q - число положительных коэффициентов β_i^* при $i\in R(X)$; r - число, равное $-\partial\varphi(X)/\partial g(X)$.

З а м е ч а н и е. В теле процедуры down используется стандартная подпрограмма p 1052 решения системы линейных алгебраических уравнений.

4°. Описание процедуры.

```
procedure down(m,p,n,a,mp,mt,q,u,r);
integer m,p,n,q,u; real r; array a,mp,mt;
  begin integer i,j,k; real s,t; array mpt[1:m];
  boolean b; boolean array z[1:m];
  u:=m; s:=1/p; q:=p; if m>p then t:=1/(m-p);
  for i:=1 step 1 until m do
  begin mt[i]:=if i>p then t else s; z[i]:=true end;
  l1:begin integer i1,k1; array aa[1:u+1,1:  +1];
```

```
    i1:=0;
    for i:=1 step 1 until m do if z[i] then
       begin k1:=i1; i1:=i1+1;
          for k:=i step 1 until m do if z[k] then
          begin k1:=k1+1; s:=0;
          for j:=1 step 1 until n do
          s:=s+a[i,j]× a[k,j];
          if i ⩽ p∧ k ⩽ p then  s:=s+1;
          aa[k1,i1]:=aa[i1,k1]:=s
          end
       end;
    for i:=1 step 1 until u do
    aa[i,u+1]:=if i > q then 0 else 1;
    p1052(u+1,u,aa);
    s:=0;
    for i:=1 step 1 until q do s:=s+aa[i,u+1];
    if s > 1 then s:=1; r:=sqrt(1/s-1);
    k:=0;
    for i:=1 step 1 until m do if z[i]then
       begin k:=k+1; mpt[i]:=aa[k,u+1]/s
       end else mpt[i]:=0
    end;
 b:=true; t:=1;
 for i:=1 step 1 until m do if z[i] then
    begin if mpt[i]<0 then
       begin b:=false; s:=mt[i]/(mt[i]-mpt[i]);
          if s<t then  begin t:=s; j:=i  end
       end
    end;
 if b then go to 12;
 z[j]:=false;  u:=u-1; if j ⩽ p then q:=q-1;
 for i:=1 step 1 until m do if z[i].then
    mt[i]:=t×(mpt[i]-mt[i])+mt[i];
 go to 11;
12: u:=m; q:=p;
  for i:=1  step 1 until m do
  begin mt[i]:= mpt[i]; if mt[i]=0 then
     begin u:=u-1; if i ⩽ p then q:=q-1;
        z[i]:=false end
  end;
```



```
  for j:=1 step 1 until n do
  begin s:=0;
     for i:=1 step 1 until m do if z[i] then
        s:=s+mt[i] × a[i,j];
     mp[j]:=s
  end;
  s:=r×r;
  for i:=1 step 1 until m do if ¬z[i] then
     begin t:=0;
        for j:=1 step 1 until n do
           t:=t+mp[j]× a[i,j];
        if i > p then s:=0;
        if s > t then
           begin z[i]:=true; u:=u+1;
           if i ⩽ p then q:=q+1; go to 11  end
     end;
  if r > 0 then
  for j:=1 step 1 until n do mp[j]:=mp[j]/r
  end
```

5°. Контрольный пример. Положим $m = 6$, $p = 3$, $n = 5$. Элементы матрицы $a[1:m,1:n]$ расположены в первых шести строках таблицы. В последней строке этой таблицы записаны компоненты вектора $mp[1:n]$; $r = 22957982$.

-0.18559730	-0.02432001	0.80958187	-0.73956334	-0.05771911
-0.21087205	0.06900608	0.32231295	0.98034584	-0.91607034
-0.22012317	-0.25042378	-0.45567977	-0.40895783	0.51494801
-0.97291028	0.07023919	-0.16951501	-0.27021396	-0.75872075
-0.56054961	-0.76365983	-0.61223018	0.20513308	-0.39638602
0.28175222	-0.41490924	-0.41679799	0.20970261	0.60392820
-0.73702216	-0.65958123	0.13590930	0.04749682	-0.03199419

Вычисления проводились по следующей программе:

```
begin integer m,p,n; read(m,p,n);
  begin real r; integer q,u;
  array a[1:m,1:n], mp[1:n], mt[1:m];
  <описание процедуры down>;
  read (a);
  down(m,p,n,a,mp,mt,q,u,r);
  print(m,p,n,a,mp,r) end end
```

§ 2.9. Обобщенный градиентный спуск с растяжением пространства

1°. Постановка задачи. Рассматривается задача

$$\varphi(X) \stackrel{def}{=} \max_{i \in 0:m} f_i(X) \to \min_{X \in E_n},$$

где функции $f_i(X)$ непрерывно дифференцируемы и выпуклы. Тогда, как известно, функция $\varphi(X)$ также выпукла и у нее в каждой точке существует субградиент. Для минимизации $\varphi(X)$ может быть использован метод обобщенного градиентного спуска [116]. Метод, однако, носит более общий характер и может применяться для минимизации произвольных квазидифференцируемых функций. В этом параграфе рассматривается модификация метода - обобщенный градиентный спуск с растяжением пространства, - позволяющая ускорить сходимость метода в задачах с функциями "овражного типа".

2°. Метод решения. Процедура shor реализует вариант метода обобщенного градиентного спуска с растяжением пространства [117-119]. При этом наибольший объем памяти занимает рабочая матрица $d[1:s, 1:n]$. Ее строками являются векторы, в направлении которых производится растяжение пространства. Число s растяжений должен заказывать пользователь. В принципе возможна такая организация вычислений, при которой через каждые s растяжений происходит восстановление первоначального состояния процесса. Дальнейший поиск осуществляется с новым начальным приближением, которым является результат предыдущего цикла. Такая организация вычислений позволяет избежать (в некоторой степени) ошибки, которая может накапливаться с ростом числа итераций. Обобщенный градиентный спуск соответствует случаю, когда коэффициент растяжения пространства равен 1.

3°. Смысл формальных параметров процедуры shor: n - размерность вектора X; m - число функций, определяющих φ; s - число растяжений пространства (вычисления прекращаются после того, как произведено s растяжений); $alfa$ - коэффициент растяжения пространства; h - начальное значение шага; ch - коэффициент изменения шага (шаг меняется по правилу $h_{i+1} = h_i \times ch$); $barrier$ - параметр, характеризующий момент растяжения пространства (если обобщенный градиент изменился незначительно по сравнению с предыдущей итерацией, растяжения не происходит; при значительном же изменении, которое характеризуется параметром $barrier$, производится растяжение пространства); $x1[1:n]$ - начальное приближение (кроме того, $x1$ используется как рабочий массив); $x[1:n]$ - значение переменной, соответствующее наименьшему достигнутому значению (напомним, что метод не является релаксационным); it - счетчик общего числа итераций; $record$ - величина, равная $\min_{k \in 0:it} \varphi(X_k)$; f - идентификатор процедуры-функции с заголовком f(i,x); value i; integer i; array x; процедура вычисляет значение функции f_i в точке X; dif - идентификатор процедуры с заголовком dif(a,k,g); value k; integer k; array a, g; процедура вычисляет градиент k-й функции в точке a и засылает его в массив g.

4°. Описание процедуры.

```
procedure shor(n,m,s,alfa,h,ch,barrier,x1)
results:(it,record,x) procedure:(f,dif);
value n,m,s,alfa,h,ch,barrier;
integer n,m,s,it;
real alfa,h,ch,barrier,record;
array x1,x; real procedure f; procedure dif;
   begin integer i,p1; real u,u1,q,hx,max;
   array d[1:s,1:n], grad,r,g,dir[1:n];
   procedure geg(x,max,g); real max; array x,g;
      begin integer i,l; real a; max:=f(0,x); l:=0;
      for i:=1 step 1 until m do
      begin a:=f(i,x);
         if max lt a then begin max:=a; l:=i  end
      end; dif(x,l,g)
      end geg;
   procedure ext(p,g); value p; integer p; array g;
      begin integer i; real s; s:=0;
      for i:=1 step 1 until n do
      s:=s+g[i]×d[p,i]; s:=q×s;
      for i:=1 step 1 until n do
         g[i]:=g[i]+s×d[p,i]
      end ext;
   real procedure norm(g); array g;
      begin integer i; real s; s:=0;
      for i:=1 step 1 until n do
      s:=s+g[i]×g[i]; norm:=sqrt(s)
```

```
    end norm;
  it:=p1:=0; q:=1/alfa-1;
  geg(x1,record,grad);
  for i:=1 step 1 until n do
  begin x[i]:=x1[i]; dir[i]:=grad[i]end;
  u:=norm(grad); if u lt 10-4 then go to fin; hx:=h/u;
mk1:it:=it+1; geg(x1,max,g);
  if max lt record then
    begin record:=max;
      for i:=1 step 1 until n do x[i]:=x1[i] end;
  if norm(g) lt 10-4 then go to fin;
  for i:=1 step 1 until p1 do ext(i,g);
  for i:=1 step 1 until n do r[i]:=g[i]-grad[i];
  u1:=norm(r); if u1/u le barrier then go to mk2;
  p1:=p1+1; if p1 gt s then go to fin;
  for i:=1 step 1 until n do d[p1,i]:=r[i]/u1;
  ext(p1,g); u:=norm(g); h:=ch×h; hx:=h/u;
  for i:=1 step 1 until n do
    dir[i]:=grad[i]:=g[i];
  for i:=p1 step -1 until 1 do ext(i,dir);
mk2: for i:=1 step 1 until n do
       x1[i]:=x1[i]-hx×dir[i];
  go to mk1;
fin: end shor
```

5°. Контрольный пример. Решалась чебышевская задача аппроксимации функции $1/(1+t)$ экспоненциальными полиномами второй степени на сетке, состоящей из точек $t_i = i/30$, $i = 0, 1, 2, \ldots, 30$. Задача имеет вид

$$\max_{i\in 0:30} \left| EP(X,t_i) - 1/(1+t_i) \right| \to \min_{X\in E_4},$$

где

$$EP(X,t) = x_1 \exp(x_3 t) + x_2 \exp(x_4 t), \quad X = (x_1, x_2, x_3, x_4).$$

В качестве исходных брались следующие значения параметров: $s = 10$, $m = 30$, $alfa = 2$, $h = 0.5$, $ch = 0.78$, $barrier = 0.9$, $x1 = (0.75, 0, 0, 0)$. Максимальное уклонение при этом равнялось 0.25.

Рабочая программа.

```
begin integer n,m,s,it;
real alfa,h,ch,barrier,record;
read(n,m,s,alfa,h,ch,barrier);
  begin array x,x1[1:2×n];
  real procedure f(i,x); value i; integer i; array x;
    begin integer k; real u,z; u:=0; z:=i/m;
    for k:=1 step 1 until n do
      u:=u+x[k]×exp(x[n+k]×z);
    f:=abs(u-1/(1+z)) end f;
  procedure dif(x,l,g); value l; integer l; array x,g;
    begin integer i, sigma; real u,z; u:=0; z:=1/m;
    for i:=1 step 1 until n do
    begin g[i]:=exp(x[n+i]×z);
      g[n+i]:=x[i]×g[i]; u:=u+g[n+i] end;
    sigma:=sign(u-1/(1+z)); u:=z×sigma;
    for i:=1 step 1 until n do
    begin g[i]:=g[i]×sigma;
      g[n+i]:=g[n+i]×u end
    end dif;
  <описание процедуры shor>;
  read(x1);
  shor(2×n,m,s,alfa,h,ch,barrier,x1,
       it,record,x,f,dif);
  print(it,record,x)
  end end program
```

Результаты вычислений: $it = 19$, $record = 0.01896$, $x = (0.7287, 0.2523, -0.8815, -0.3410)$.

6°. В задачах небольшой размерности s рекомендуется брать в пределах от $2n$ до $5n$. Опыт показывает, что наилучшие значения коэффициента растяжения пространства $alfa$ находятся в пределах от 2 до 4. Такой выбор подтверждается и теоретически, ибо при $alfa > 4$ норма оператора, обратного к оператору растяжения пространства, становится малой, что замедляет сходимость метода. Если eps – требуемая точность вычисления X, то начальное значение h, параметры ch, s и eps должны быть связаны соотношением $eps = (ch)^s h$. При $ch = 1$ получаем метод обобщенного градиентного спуска с растяжением пространства и постоянным шагом. Выбор параметра $barrier$ можно ограничить пределами от 0.9 до 1. Данный метод рекомендуется применять, когда не нужна слишком большая точ-

ность либо при нахождении начального приближения для других методов, имеющих бо́льшую скорость сходимости.

§ 2.10. Процедура выравнивания максимумов

1°. Постановка задачи. Рассмотрим задачу

$$\varphi(X) \overset{def}{=} \max_{t\in[a,b]} f(X,t) \to \min_{X\in E_n} .$$

После шага в направлении наискорейшего спуска часто оказывается полезным возвратиться на поверхность равных максимумов, т.е. решить следующую систему уравнений:

$$\begin{aligned} &f(X,t_i) - f(X,t_m) = 0 , \quad i\in 1:m-1 , \\ &f'_t(X,t_i) = 0 , \quad i\in 1:m , \quad t_i \neq a , \quad t_i \neq b . \end{aligned} \tag{1}$$

Обратимся к задаче нелинейного программирования с континуумом ограничений

$$F(X) \to \min_{X\in\Omega} ,$$

где $\Omega = \{ X\in E_n \mid \varphi(X) \overset{def}{=} \max\limits_{t\in[a,b]} f(X,t) \leqslant 0 \}$.

Если $\varphi(X) = 0$, то, сделав шаг в направлении наискорейшего спуска, можно выйти из Ω, поэтому для возвращения на множество Ω необходимо решить систему уравнений вида

$$\begin{aligned} &f(X,t_i) = 0 , \quad i\in 1:m , \\ &f'_t(X,t_i) = 0 , \quad i\in 1:m , \quad t_i \neq a , \quad t_i \neq b . \end{aligned} \tag{2}$$

Рассмотрим, наконец, задачу о нахождении максимального отрезка аппроксимации: требуется найти $\max(b-a)$ на множестве

$$\Omega = \{ (X,a,b)\in E_{n+2} \mid \varphi(X,a,b) \overset{def}{=} \max_{t\in[a,b]} f(X,t) \leqslant 0 \} .$$

В этой задаче после шага в направлении наискорейшего спуска обычно приходится решать систему

$$\begin{aligned} &f(X,t_i) = 0 , \quad i\in 1:m , \\ &f(X,a) = 0 , \\ &f(X,b) = 0 , \\ &f'_t(X,t_i) = 0 , \quad i\in 1:m . \end{aligned} \tag{3}$$

2°. Метод решения. Системы (1)–(3) естественно решать методом выравнивания максимумов. Для этого из верхней группы уравнений систем (1), (2) или (3) определяют поправку ΔX по методу Ньютона с псевдообратной матрицей (либо обратной, если $m = n+1$ для (1) и $m = n$ для (2), (3)). Затем при $X = X + \Delta X$ уточняют точки внутренних локальных максимумов функции f по t и т.д. Поправку ΔX находят процедурой rest, а максимумы уточняют процедурой mmax.

Замечание 1. В системе (3) из верхней группы уравнений находят поправку не только для X, но и для a и b.

3°. Смысл формальных параметров. Процедура rest. Входными параметрами являются: m – число выравниваемых максимумов; n – размерность пространства, в котором производится минимизация; $f[1:m]$ – массив, компоненты которого имеют значения $f(X,t_i)$, $i\in 1:m$; $a[1:m,1:n]$ – матрица, строки которой – векторы $f'_X(X,t_i)$, $i\in 1:m$; e – булевская переменная, которой следует присвоить значение false, если нужно решить систему (1), или true, если нужно решить системы (2) или (3).

Выходным параметром является вектор $g[1:n]$, совпадающий с поправкой ΔX.

Замечание 2. В теле процедуры rest используется стандартная программа p1052 решения системы линейных алгебраических уравнений.

Процедура mmax (уточняет по конечно-разностному аналогу метода Ньютона точки локального максимума функции f по t).

Входными параметрами являются: f – идентификатор процедуры-функции с заголовком f(x,t); value t; real t; array x; m – число точек максимума, которые требуется уточнить; d – точность, с которой вычисляются точки локального максимума функции f по t ; a – значение левого конца отрезка $[a,b]$, b – правого; h – длина шага конечных разностей (частные производные функции f вычисляют приближенно); $x[1:n]$ – вектор, совпадающий с X.

Выходными параметрами являются массивы $mt, mf[1:m]$, i-е компоненты которых равны соответственно t_i и $f(X,t_i)$, $i\in 1:m$.

4°. Описание процедур.

```
procedure rest(m,n,f,a,g,e);
integer m,n; array f,a,g; boolean e;
   begin integer i,j,k,q,u;
   q:=if e then m else m-1; u:=q+1;
      begin real s; array aa[1:u, 1:u];
      for i:=1 step 1 until q do
      if e then aa[i,u]:=abs(f[i])
         else begin
         for j:=1 step 1 until n do a[i,j]:=a[i,j]-a[m,j];
         aa[i,u]:=abs(f[m])-abs(f[i]) end;
      for i:=1 step 1 until q do
      for k:=i step 1 until q do
      begin s:=0; for j:=1 step 1 until n do
         s:=s+a[i,j]× a[k,j];
      aa[i,k]:=aa[k,i]:=s end;
      p1052 (u,q,aa);
      for j:=1 step 1 until n do
      begin s:=0;
         for i:=1 step 1 until q do
         s:=s+aa[i,u]× a[i,j];
      g[j]:=s end end end rest.
procedure mmax(f,m,d,a,b,h,x,mt,mf);
integer m; real d,a,b,h; array x,mt,mf;
real procedure f;
   begin integer i; real f1,f2,f3,g,z,t;
   for i:=1 step 1 until m do
   begin z:=mt[i]; f2:=f(x,z);
      if z<b∧z>a then
      begin
      l1: f3:=f(x,z+h); f1:=f(x,z-h);
         g:=h×(f3-f1)/(f1-2×f2+f3)/2;
         z:=z-g; f2:=f(x,z);
         if abs(g)>d then go to l1
      end
      mt[i]:=z; mf[i]:=f2
   end end mmax
```


5°. Контрольный пример приведен в § 2.12.

6°. Выравнивание максимумов следует производить с точностью, пропорциональной квадрату текущего расстояния от начала координат до многогранника.

§ 2.11. Вычисление ограничителя шага спуска

1°. Постановка задачи и метод решения. Рассматривается задача

$$\varphi(X)\overset{def}{=}\max_{t\in[a,b]} f(X,t)\to\min_{X\in E_n},$$

где $f(X,t)$ – непрерывно дифференцируемая на $E_n\times[a,b]$ функция.

Пусть в точке X найдено направление наискорейшего спуска функции φ – вектор $g(X)$, $\|g(X)\|=1$. Выберем произвольные положительные числа δ и α_0.

Найдем множество

$$R_\delta(X)=\{\,t\in[a,b]\mid f(X,t)\geq\varphi(X)-\delta,$$

t – точка локального максимума функции $f\,\}$.

Предположим, что в рассматриваемой точке множество $R_\delta(X)$ содержит конечное число элементов:

$$R_\delta(X)=\{\,t_i\mid i\in 1:p\,\},$$

причем $f(X,t_i)=\varphi(X)\quad\forall i\in 1:m\quad(m<p)$,

$$f(X,t_i)<\varphi(X)\quad\forall i\in m+1:p.$$

Положим $r(X)=-\partial\varphi(X)/\partial g(X)$. Найдем числа α_j такие, что

$$\varphi(X)-\alpha_j r(X)=f(X,t_j)+\alpha_j\left(f'_X(X,t_j),g(X)\right)\quad\forall j\in m+1:p.$$

Естественно ограничить длину шага спуска числом

$$s=\min_{j\in m+1:p}\alpha'_j,$$

где

$$\alpha'_j=\begin{cases}\alpha_j, & \text{если}\quad\alpha_j>0,\\ \alpha_0, & \text{если}\quad\alpha_j\leq 0.\end{cases}$$

Такой выбор ограничителя шага спуска осуществляется процедурой lp. При использовании процедуры необходимо знать число элементов множества

$$R'(X)=\{\, t\in R(X) \mid (f'_X(X,t), g(X))=-r(X)\,\}.$$

З а м е ч а н и е. После вычисления ограничителя шага можно перейти к отысканию числа $\alpha(X)$ такого, что

$$\varphi(X+\alpha(X)g(X))=\min_{\alpha\in(0,s]}\varphi(X+\alpha g(X)).$$

2°. Смысл формальных параметров процедуры lp. Входные параметры: m,p,n,r,x,g (имеют тот же смысл, что и при постановке задачи); ah – число, равное α_0; q – число точек в множестве $R'(X)$; f – идентификатор процедуры-функции с заголовком f(x,t); value t;real t;array x; h – длина шага конечных разностей (частные производные вычисляются приближенно); $mt,mf[1:p]$ – массивы, i-е компоненты которых равны соответственно t_i и $f(X,t_i)$ $(i\in 1:p)$.

Выходными параметрами являются: s – ограничитель шага; m – число, равное q, если $s=\alpha h$, или $q+1$, если $s<\alpha h$.

3°. Описание процедуры.

```
procedure lp(m,q,n,p,f,s,ah,h,r,x,mt,mf,g);
integer m,q,n,p; real s,ah,h,r;
array x,mt,mf,g; real procedure f;
   begin integer i,j; real w,z,t,f1,f2; s:=ah;
   for i:=m+1 step 1 until p do
   begin w:=0; t:=mt[i]; f2:=mf[i];
      for j:=1 step 1 until n do
      begin z:=x[j]; x[j]:=z+h; f1:=f(x,t); x[j]:=z;
         w:=w+g[j] x(f1-f2)/h end:
      z:=r+w;
      if z>0 then
      begin t:=(mf[1]-f2)/z; if s>t then s:=t end
   end;
   m:=if s<ah then q+1 else q
   end lp
```

4°. Контрольный пример приведен в следующем параграфе.

5°. Числа δ и α_0 необходимо уменьшать при приближении к стационарной точке функции φ.

§ 2.12. Контрольный пример для процедур rest, mmax, lp, down

1°. Была решена следующая задача построения электрической цепи со структурой фильтра нижних частот:

$$(\alpha-b)\to\min \qquad (1)$$

на множестве $\Omega=\{(X,b)\in E_5 \mid \varphi(X,b)\overset{\text{def}}{=}\max\limits_{t\in[\alpha,b]} f(X,t)\leqslant 0{,}0363\}$,

где $f(X,t)=|\Gamma(X,t)|$, $\Gamma(X,t)=\dfrac{Z_{вх}-1}{Z_{вх}+1}$, $Z_{вх}=\dfrac{A_{11}Z_н+A_{12}}{A_{21}Z_н+A_{22}}$,

$$\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}=\begin{bmatrix}\dfrac{(Z_1+Z_2)(Z_3+Z_4)+Z_3Z_4}{Z_1Z_3} & \dfrac{Z_2(Z_3+Z_4)+Z_3Z_4}{Z_3}\\ \dfrac{Z_3(Z_4+Z_5)+(Z_1+Z_2)(Z_3+Z_4+Z_5)}{Z_1Z_3Z_5} & \dfrac{(Z_2+Z_3)(Z_4+Z_5)+Z_2Z_3}{Z_3Z_5}\end{bmatrix},$$

$\alpha=0.7$, $Z_н=0.5$, $Z_1=\sqrt{-1}\,/\,t$, $Z_{k+1}=(\sqrt{-1}\,x_k t)^{(-1)^k}$, $k\in 1:4$.

Таким образом, требуется выбрать параметр X так, чтобы функция f принимала значения, не превосходящие 0.0363 на отрезке $[\alpha,b]$ максимальной длины (α фиксировано). Частные случаи задачи такого типа рассмотрены в [8, 14].

2°. Для решения задачи (1) был применен метод, в котором после шага в направлении наискорейшего спуска производилось, если это было необходимо, возвращение на множество Ω. Одна итерация такого метода использована в качестве контрольного примера для процедур rest, mmax, lp, down.

Значения функции f вычислялись процедурой-функцией с описанием

```
real procedure f(x,t); value t; real t; array x;
   begin integer i; real u,v,s1,s2,s3,s4; array z[2:5];
      for i:=2 step 1 until 5 do z[i]:=t×x[i-1];
      u:=1-t×z[2]; v:=1-z[3]×z[4];
      s1:=u×v -t×z[4]; s3:=(v-s1)/t;
      s2:=z[5]×s1+t+z[3]×u;
      s4:=(s2-z[3]-z[5]×v)/t;
      u:=.5×s1; z[2]:=s4-u; z[4]:=s4+u;
      u:=.5×s2; z[3]:=u-s3; z[5]:=u+s3;
      f:=sqrt((z[2]×z[4]+z[3]× z[5])↑2+
      (z[3]× z[4] - z[2]× z[5])↑2)/
      (z[4]↑2 + z[5]↑2) end
```

Кроме того, использовалась процедура вычисления матрицы частных производных для работы процедур down и rest, имеющая следующее описание:

```
procedure cmg(m,n,f,x,mt,mf,mg,e,h);
integer m,n; real procedure f; real h;
array x,mt,mf,mg; boolean e;
   begin integer i,j,k,q; real z; j:=0; q:=n-1;
   if ¬e then begin j:=j+1;
      for k:=1 step 1 until q do mg[1,k]:=0;
      mg[1,n]:=-1 end;
   for i:=j+1 step 1 until m do
   begin for k:=1 step 1 until q do
      begin z:=x[k]; x[k]:=z+h;
         mg[i,k]:=(f(x,mt[i])-mf[i])/h; x[k]:=z end;
      mg[i,n]:=0 end;
   q:=j+1; z:=mt[q]+h;
   mg[q,n]:=(f(x,z)-mf[1])/h  end
```

Заметим, что если матрица mg используется при обращении к процедуре down, то переменной e следует присвоить значение false, а если к rest, то true.

Далее приводится программа проверки работы процедур rest, mmax, lp, down.

```
begin integer n,p,q; n:=5; p:=3; q:=4;
   begin integer m,i,k,u,l; boolean e;
   array x[1:q], g[1:n], mt,mf,mq[1:p], mg[1:p,1:n];
   <описание процедур f,cmg,rest,mmax,lp,down>;
   real h,ah,s,d,r,a,b;
   ah:=.1; d:= 10-5; m:=2; e:=false;
   read(x,mt,mf); a:=mt[2]; b:=mt[1];
11: cmg(p,n,f,x,mt,mf,mg,e,h);
   if mf[1]-mf[m]>d then go to 12;
   down(p,1,n,mg,g,mq,u,l,r); i:=l-u;
   lp(m,i,q,f,s,ah,h,r,x,mt,mf,g);
   print(k,x,mt,mf,s); k:=k+1;
   for i:=1 step 1 until q do  x[i]:=x[i]-s×g[i];
   b:=b-s×g[n]; e:=true; go to 13;
12: for i:=1 step 1 until m do mq[i]:=mf[i]-.0363;
   rest(m,n,mq,mg,g,e);
   for i:=1 step 1 until q do x[i]:=x[i]-g[i];
   b:=b-g[n]; k:=k+1;
13: mt[1]:=b;
   mmax(f,m,d,a,b,h,x,mt,mf);
   print(k,x,mt,mf);
   if k<3 then go to 11
   end end
```


Результаты помещены в табл. 1.

З а м е ч а н и е. В качестве начальных значений для массивов mt, mf и x следует взять соответствующие массивы из табл. 1 при $k = 0$.

Т а б л и ц а 1

		$k = 0$		
x	.72852	2.21918	1.03062	.97081
mt	1.01059	.7	.87878	
mf	.03630	.03630	.03452	
s	.01044			
		$k = 1$		
x	.72650	2.21582	1.03045	.97599
mt	1.01877	.7	.87715	
mf	.03672	.03631	.03642	
		$k = 2$		
x	.72628	2.21615	1.03016	.97609
mt	1.01825	.7	.87739	
mf	.03630	.03630	.03630	

В заключение отметим, что после 12 итераций наискорейшего спуска с возвращением на множество Ω были получены следующие окончательные результаты (табл. 2).

Т а б л и ц а 2

x	68808	2.25395	.99200	1.03581
mt	1.04976	.7	.89915	
mf	.03630	.03630	.03630	

УКАЗАТЕЛЬ ЛИТЕРАТУРЫ

1. Алиев М.С. Условия совместности одной системы линейных неравенств. – В кн.: Функциональный анализ, теория функций и их приложения. Вып. 1. Махачкала, 1974, с. 26–29.

2. Алявдин П.В. Оптимизация предварительного напряжения и пригруза геометрически нелинейных систем как минимаксная задача. – В кн.: У1 Всесоюз. конф. по экстремальным задачам. Ч.I. Таллин, 1973, с. 21–22.(Тезисы докл.)

3. Аппроксимация и интерполяция функций (алгоритмы и программы машинного проектирования самолетов). – В кн.: Труды Центрального аэрогидродинамич. ин-та им. Н.Е.Жуковского. Вып. 1646. М., 1975. 99 с.

4. Береснев В.В. Необходимые условия экстремума в выпуклой задаче максимина на связанных множествах. – "Кибернетика", 1972, № 2, с. 87–91.

5. Болтянский В.Г. Метод шатров в теории экстремальных задач. – "Успехи мат. наук", 1975, т. XXX, вып. 3, с. 3–55.

6. Болтянский В.Г., Чеботару И.С. Необходимые условия в задаче о минимаксе. – "Докл. АН СССР", 1973, т. 213, № 2, с. 257–260.

7. Брокан Л.Н. Ускорение сходимости в непрерывных минимаксных задачах. – В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 73–84.

8. Бунимович Б.Ф., Васильев Л.В., Мещанов В.П. Определение максимальной области аппроксимации в задачах синтеза линейных электрических цепей. – "Радиотехника", 1976, № 1, с. 78–80.

9. Васильев Л.В., Митчелл Б.Ф. Один конечный метод нахождения направления наискорейшего спуска. – В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 59–82.

10. Васильев Л.В. О нахождении направления наискорейшего спуска в минимаксных задачах. – В кн.: VI Всесоюз. конф. по экстремальным задачам. Ч. I. Таллин, 1973, с. 61–62.(Тезисы докл.)

11. Вершик А.М., Малоземов В.Н., Певный А.Б. Наилучшая кусочно-полиномиальная аппроксимация. – "Сиб. мат. журн.", 1975, т. 16, № 5, с. 925–938.

12. Виноградова Т.К., Демьянов В.Ф. Принцип минимакса в задачах оптимального управления. – "Докл. АН СССР", 1973, т. 213, № 3, с. 512–514.

13. Войтон Е.Ф., Демьянов В.Ф. Разыскание направления наискорейшего спуска в минимаксных задачах при связанных ограничениях. – В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 30–40.

14. Войтон Е.Ф., Полякова Л.Н. Об одном классе экстремальных задач при связанных ограничениях. – В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 41–46.

15. Войтон Е.Ф., Шушков В.И. О поиске направления наискорейшего спуска в некоторых минимаксных задачах. – "Журн. вычисл. мат. и мат. физики", 1972, т. 12, № 4, с. 1013–1021.

16. Габасов Р., Кириллова Ф.М. Качественная теория оптимальных процессов. М., "Наука", 1971. 507 с.

17. Гавурин М.К. Нелинейные функциональные уравнения и непрерывные аналоги итеративных методов. – "Изв. высш. учеб. заведений. Математика", 1958, № 5, с. 18–31.

18. Гасс С. Линейное программирование. М., Физматгиз, 1961. 303 с.

19. Гермейер Ю.Б. Введение в теорию исследования операций. М., "Наука", 1971. 383 с.

20. Гехер К. Теория чувствительности и допусков электронных цепей. М., "Советское радио", 1973. 200 с.

21. Гирсанов И.В. Лекции по математической теории экстремальных задач. М., Изд-во Моск. ун-та, 1970. 118 с.

22. Гольштейн Е.Г. Теория двойственности в математическом программировании и ее приложения. М., "Наука", 1971. 351 с.

23. Гольштейн Е.Г. Обобщенный метод отыскания седловых точек. – "Экономика и мат. методы", 1972, т. 8, вып. 4, с. 569–579.

24. Горелик В.А. Приближенное нахождение максимина с ограничениями, связывающими переменные. – "Журн. вычисл. мат. и мат. физики", 1972, т. 12, № 2, с. 510–517.

25. Горелик В.А., Федоров В.В. Метод внешней точки в задаче определения кратного максимина с ограничениями. – "Журн. вычисл. мат. и мат. физики", 1975, т. 15, № 3, с. 599–607.

26. Грачев Н.И., Евтушенко Ю.Г. Численный метод решения минимаксных задач. – "Журн. вычисл. мат. и мат. физики", 1971, т. 11, № 2, с. 375–384.

27. Гурин Л.Г., Столярова Е.М. Принцип максимума в одной минимаксной задаче. – "Журн. вычисл. мат. и мат. физики", 1973, т. 13, № 5, с. 1175–1185.

28. Данскин Дж.М. Теория максимина и ее приложение к задачам распределения вооружения. М., "Советское радио", 1970. 200 с.

29. Дарховский Б.С., Левитин Е.С. О дифференциальных свойствах минимакса в одной задаче распределения ресурсов. – "Журн. вычисл. мат. и мат. физики", 1973, т. 13, № 5, с. 1219-1227.

30. Даугавет В.А. Альтернансные свойства решений нелинейных минимаксных задач с нелинейными ограничениями. – "Журн. вычисл. мат. и мат. физики", 1976, т. 16, № 3, с. 784-788.

31. Даугавет В.А. Одна нелинейная задача аппроксимации функций двух переменных. – В кн.: VI Всесоюз. конф. по экстремальным задачам. Ч. I. Таллин, 1973, с. 112.(Тезисы докл.)

32. Даугавет В.А., Малоземов В.Н. Альтернансные свойства решений нелинейных минимаксных задач с невыпуклыми ограничениями. – "Докл. АН СССР", 1975, т. 225, № 2, с. 253-255.

33. Демьянов В.Ф. Ускорение сходимости при решении минимаксных задач. – "Журн. вычисл. мат. и мат. физики", 1972, т. 12, № 1, с. 51-60.

34. Демьянов В.Ф. Минимакс: дифференцируемость по направлениям. Л., Изд-во Ленингр. ун-та, 1974. 112 с.

35. Демьянов В.Ф. О непрерывном методе решения минимаксных задач. – "Журн. вычисл. мат. и мат. физики", 1975, т. 15, № 3, с. 592-598.

36. Демьянов В.Ф. Пакетный принцип минимакса. – "Вестн. Ленингр. ун-та", 1976, № 19, с. 35-39.

37. Демьянов В.Ф. К методу экстремального базиса. – "Докл. АН СССР", 1976, т. 229, № 2, с. 272-275.

38. Демьянов В.Ф., Малоземов В.Н. К теории нелинейных минимаксных задач. – "Успехи мат. наук", 1971, т. XXVI, вып. 3, с. 53-104.

39. Демьянов В.Ф., Малоземов В.Н. Введение в минимакс. М., "Наука", 1972. 368 с.

40. Демьянов В.Ф., Певный А.Б. Численные методы разыскания седловых точек. – "Журн. вычисл. мат. и мат. физики", 1972, т. 12, № 5, с. 1099-1127.

41. Демьянов В.Ф., Певный А.Б. Первые и вторые маргинальные значения задач математического программирования. – "Докл. АН СССР", 1972, т. 207, № 2, с. 277-280.

42. Демьянов В.Ф., Певный А.Б. Некоторые оценки в минимаксных задачах. – "Кибернетика", 1972, № 1, с. 107-112.

43. Демьянов В.Ф., Рубинов А.М. Приближенные методы решения экстремальных задач. Л., Изд-во Ленингр. ун-та, 1968. 180 с.

44. Дубовицкий А.Я., Милютин А.А. Задачи на экстремум при наличии ограничений. – "Журн. вычисл. мат. и мат. физики", 1965, т. 5, № 3, с. 395-453.

45. Евтушенко Е.Г., Жадан В.Г. Численные методы решения некоторых задач исследования операций. – "Журн. вычисл. мат. и мат. физики", 1973, т. 13, № 3, с. 583-598.

46. Ермольев Ю.М. Методы решения нелинейных экстремальных задач. – "Кибернетика", 1966, № 4, с. 1-17.

47. Ермольев Ю.М., Нурминский Е.А. Предельные экстремальные задачи. – "Кибернетика", 1973, № 4, с. 130-132.

48. Ермольев Ю.М., Шор Н.З. О минимизации недифференцируемых функций. – "Кибернетика", 1967, № 1, с. 101-102.

49. Зангвилл У.И. Нелинейное программирование. М., "Советское радио", 1973. 311 с.

50. Зойтендейк Г. Методы возможных направлений. М., ИЛ, 1963. 176 с.

51. Зуховицкий С.И., Поляк Р.А., Примак М.Е. Алгоритм для решения задачи выпуклого программирования. – "Докл. АН СССР", 1963, т. 153, № 5, с. 991-994.

52. Зуховицкий С.И., Поляк Р.А., Примак М.Е. Вогнутые игры многих лиц (численные методы). – "Экономика и мат. методы", 1971, т. 7, № 6, с. 889-900.

53. Иоффе А.Д., Тихомиров В.М. Теория экстремальных задач. М., "Наука", 1974. 479 с.

54. Итеративные методы в теории игр и программировании. Под ред. В.З.Беленького и В.А.Волконского. М., "Наука", 1974. 239 с.

55. Канторович Л.В. О методе Ньютона для функциональных уравнений. – "Докл. АН СССР", 1949, т. 59, с. 1237-1240.

56. Канторович Л.В. Функциональный анализ и прикладная математика. – "Успехи мат. наук", 1948, т. 3, вып. 3 (28), с. 89-185.

57. Карлин С. Математические методы в теории игр, программировании и экономике. М., "Мир", 1964. 836 с.

58. Коллатц Л. Функциональный анализ и вычислительная математика. М., "Мир", 1968. 447 с.

59. Красовский Н.Н., Субботин А.И. Позиционные дифференциальные игры. М., "Наука", 1974. 456 с.

60. Крейн М.Г., Нудельман А.А. Проблема моментов Маркова и экстремальные задачи. М., "Наука", 1973. 551 с.

61. Куратовский К. Топология. Т. I. М., "Мир", 1966. 594 с.

62. Кутателадзе С.С., Рубинов А.М. Двойственность Минковского и ее приложения. – "Успехи мат. наук", 1972, т. 27, вып. 3 (165),

с. 127-176.

63. Кюнци Г., Крелле В. Нелинейное программирование. М., "Советское радио", 1965. 300 с.

64. Ланне А.А. Оптимальный синтез линейных электрических цепей. М., "Связь", 1969. 293 с.

65. Левитин Е.С. Дифференцируемость по параметру оптимального значения в параметрических задачах нелинейного программирования. - В кн.: VI Всесоюз. конф. по экстремальным задачам. Ч. II. Таллин, 1973, с. 13-14.(Тезисы докл.)

66. Лоран П.-Ж. Аппроксимация и оптимизация. М., "Мир", 1975. 496 с.

67. Макаров В.Л., Рубинов А.М. Математическая теория экономической динамики и равновесия. М., "Наука", 1973. 335 с.

68. Малоземов В.Н. Относительная ε-стационарность в дискретных минимаксных задачах. - В кн.: Оптимизация. Вып. 4 (21). Новосибирск, 1971, с. 28-40.

69. Малоземов В.Н. К нахождению алгебраического полинома наилучшего приближения. - "Кибернетика", 1969, № 5, с. 125-131.

70. Малоземов В.Н. О сходимости сеточного метода в нелинейных минимаксных задачах. -"Вестн. Ленингр. ун-та", 1971, № 19, с. 35-37.

71. Малоземов В.Н. Необходимые и достаточные условия минимакса. - В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 85-97.

72. Малоземов В.Н. Минимаксиминная задача с вырожденной целевой функцией. - В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 47-53.

73. Малоземов В.Н. Таблицы коэффициентов полиномов наилучшего приближения. - В кн.: Методы вычислений. Вып. 9. Л., Изд-во Ленингр. ун-та, 1974, с. 7-15.

74. Малоземов В.Н. О выравнивании максимумов. - "Журн. вычисл. мат. и мат. физики", 1976, т. 16, № 3, с. 781-784.

75. Малоземов В.Н. О достаточных условиях локального минимакса. - "Вестн. Ленингр. ун-та", 1976, № 7, с. 55-59.

76. Малоземов В.Н., Певный А.Б. Альтернансные свойства решений нелинейных минимаксных задач. - "Докл. АН СССР", 1973, т. 212, № 1, с. 37-39.

77. Малоземов В.Н., Певный А.Б. О наилучшей кусочно-полиномиальной аппроксимации. - "Вестн. Ленингр. ун-та", 1976, № 19, с. 90-96.

78. Малоземов В.Н., Певный А.Б. Рекуррентные вычисления. Л., Изд-во Ленингр. ун-та, 1976. 56 с.

79. Математическая теория оптимальных процессов. М., Физматгиз, 1961. 391 с. Авт.: Л.С.Понтрягин, В.Г.Болтянский, Р.В.Гамкрелидзе, Е.Ф.Мищенко.

80. Моисеев Н.Н. Численные методы в теории оптимальных систем. М., "Наука", 1971. 424 с.

81. Наилучшее приближение табличных функций. Ч. 1. Киев, 1973. 212 с.

82. Никольский С.М. Приближение функций тригонометрическими полиномами в среднем. - "Изв. АН СССР. Сер. мат.", 1946, т. 10, № 3, с. 207-256.

83. Нурминский Е.А. Квазиградиентный метод решения задачи нелинейного программирования. - "Кибернетика", 1973, № 1, с. 122-125.

84. Певный А.Б. Минимизация функции максимина. - "Журн. вычисл. мат. и мат. физики", 1972, т. 12, № 1, с. 227-230.

85. Певный А.Б. Скорость сходимости некоторых методов нахождения минимакса и седловых точек. - "Кибернетика", 1972, № 4, с. 95-98.

86. Певный А.Б. Маргинальные значения в невыпуклом программировании. - В кн.: VI Всесоюз. конф. по экстремальным задачам. Ч. II. Таллин, 1973, с. 58-59.(Тезисы докл.)

87. Певный А.Б. Нахождение минимакса при связанных ограничениях. - В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 22-29.

88. Певный А.Б. О нахождении точки многогранника, ближайшей к началу координат. - В кн.: Оптимизация. Вып. 10 (27). Новосибирск, 1973, с. 54-58.

89. Полак Э. Численные методы оптимизации (единый подход). М., "Мир", 1974. 376 с.

90. Поляк Б.Т. Один общий метод решения экстремальных задач.- "Докл. АН СССР", 1967, т. 174, № 1, с. 33-36.

91. Поляк Б.Т. Минимизация негладких функционалов. - "Журн. вычисл. мат. и мат. физики", 1969, т. 9, № 3, с. 509-521.

92. Поляк Б.Т. Метод сопряженных градиентов в задачах на экстремум. - "Журн. вычисл. мат. и мат. физики", 1969, т. 9, № 4, с. 807-821.

93. Приближенное решение операторных уравнений. М., "Наука", 1969. 456 с. Авт.: М.А.Красносельский, Г.М.Вайникко, Н.П.Забрейко, Я.Б.Рутицкий, В.Я.Стеценко.

94. Программы оптимизации (приближение функций). Вып. 3. Свердловск, 1972. 91 с.

95. Программы оптимизации (приближение функций). Вып. 6. Свердловск, 1975. 170 с.

96. Пропой А.И. К теории максиминных задач. – "Журн. вычисл. мат. и мат. физики", 1971, т. 11, № 1, с. 65-78.

97. Пшеничный Б.Н. Необходимые условия экстремума. М., "Наука", 1969. 151 с.

98. Пшеничный Б.Н., Данилин Ю.М. Численные методы в экстремальных задачах. М., "Наука", 1975. 319 с.

99. Ремез Е.Я. О методе последовательных чебышевских интерполяций и о различных вариантах его реализации. – "Укр. мат. журн.", 1960, т. 12, № 2, с. 170-180.

100. Ремез Е.Я. Основы численных методов чебышевского приближения. Киев, "Наукова думка", 1969. 623 с.

101. Рокафеллар Р. Выпуклый анализ. М., "Мир", 1973. 469 с.

102. Рубинштейн Г.Ш. Конечномерные модели оптимизации. Новосибирск, Изд-во Новосибирск. ун-та, 1970. 228 с.

103. Рубинштейн Г.Ш. Двойственные экстремальные задачи."Докл. АН СССР", 1963, т.152, № 2, с. 288-291.

104. Сотсков А.И. Необходимые условия минимума для одного типа негладких задач. – "Докл. АН СССР", 1969, т. 189, № 2, с. 261-264.

105. Тихомиров В.М. Теория экстремальных задач и теория приближений. Киев, 1974. 36 с.

106. Толстоногов А.А. О свойствах функции максимума при ограничениях. – "Журн. вычисл. мат. и мат. физики", 1970, т. 10, № 3, с. 597-606.

107. Тюрин Ю.Н. Математическая формулировка упрощенной модели производственного планирования. – "Экономика и мат. методы", 1965, т. 1, № 3, с. 391-410.

108. Федоров В.В. Необходимые условия оптимальности с передачей информации. – "Журн. вычисл. мат. и мат. физики", 1975, т. 15, № 2, с. 505-508.

109. Фиакко А., Мак-Кормик Г. Нелинейное программирование. М., "Мир", 1972. 240 с.

110. Филиппов А.Ф. Дифференциальные уравнения с разрывной правой частью. – "Мат. сборник", 1960, т. 51 (93), № 1, с. 99-128.

111. Фрейман М.И. Решение задачи выпуклого программирования методом Келли с фиксированным количеством гиперплоскостей. – "Вестн.

Ленингр. ун-та", 1976, № 1, с. 155-157.

112. Хартман Ф. Обыкновенные дифференциальные уравнения. М., "Мир", 1970. 720 с.

113. Хаусдорф Ф. Теория множеств. М.-Л., ОНТИ, 1937. 304 с.

114. Черников С.Н. Линейные неравенства. М., "Наука", 1968. 488 с.

115. Шилов Г.Е. Математический анализ. Функции нескольких переменных. М., "Наука", 1972. 622 с.

116. Шор Н.З. Обобщенный градиентный спуск. – В кн.: Труды I зимней школы по математическому программированию в г. Дрогобыче. Вып. 3. М., 1969, с. 578-585.

117. Шор Н.З. Использование операции растяжения пространства в задачах минимизации выпуклых функций. – "Кибернетика", 1970, № 1, с. 6-12.

118. Шор Н.З., Журбенко Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов. – "Кибернетика", 1971, № 3, с. 51-59.

119. Шор Н.З., Шабашова Л.П. О решении минимаксных задач методом обобщенного градиентного спуска с растяжением пространства. – "Кибернетика", 1972, № 1, с. 82-88.

120. Amer R.A.-R. The approximation problem of electrical filters. Basel-Stuttgart, Birkhäuser, 1964. 65 p.

121. Approximation of functions. Ed.H.L.Garabedian. Amsterdam, Elsevier Publishing Co., 1965. 220 p.

122. Approximation theory. Ed.A.Talbot. London-New York, Acad. Press, 1970. 356 p.

123. Auslender A. Prlblèmes de minimax via l'analyse convexe et les inégalités variationelles: théorie et algoritmes. Lecture notes in economics and mathematical systems. Berlin-Heidenberg-New York, Springer-Verlag, 1972. 132 p.

124. Barrar R., Loeb H. Best non-linear uniform approximation with interpolation. – "Arch.Rational Mech. Anal.", 1969, vol.33, N 3, p.231-237.

125. Barrar R., Loeb H. On the continuity of the non-linear Tshebysheff operator. – "Pacific J. Math.", 1970, vol.32, N 3, p.593-601.

126. Barrar R., Loeb H. On the Remez algorithm for non-linear families. – "Numer. Math.", 1970, Bd 15, H.5, S.382-391.

127. Bertsekas D.P., Mitter S.K. A descent numerical method for optimisation problems with non-differentiable cost functionals.– "SIAM J. Control", 1973, vol.11, N 4, p.637-652.

128. Braess D. Geometrical characterizations for non-linear uniform approximation. - "J. Approxim. Theory", 1974, vol.11,N 3, p.260-274.

129. Brosowski B., Hoffmann K.-H. Eine Variationsungleichung und Anvendungen. - "Numer. Math.", 1974, Bd 22, H.2, S.137-147.

130. Carasso C. Convergence de l'algorithme de Rémès. - "J. Approxim. Theory", 1974, vol.11, N 2, p.149-158.

131. Chalmers B.L. Uniqueness of best approximation of a function and its derivatives. - "J. Approxim. Theory", 1973,vol.7, N 3, p.213-225.

132. Cody W.J. A survey of practical rational and polynomial approximation of functions. - "SIAM Rev.", 1970, vol.12,N 3,p.400-423.

133. Cody W.J., Stoer J. Rational Chebyshev approximation using interpolation. - "Numer. Math.", 1966, Bd 9, H.3,S.177-188.

134. Davidon W.C. Variance algorithm for minimization. - "Computer J.", 1968, vol.10, p.406-410.

135. Demyanov V.F. Algorithms for some minimax problems. - "J. Comp. and System Sci.", 1968, vol.2, N 4, p.342-380.

136. Deutch F. On uniform approximation with interpolatory constraints. - "J. Math. Anal. Appl.", 1968, vol.24, N 1,p.62-79.

137. Dunham C.B. Alternating minimax approximation with unequal restraints. - "J. Approxim. Theory", 1974, vol.10,N 3,p.199-205.

138. Fletcher R., Powell M.J.D. A rapidly convergent descent method for minimization. - "Comput. J.", 1963, vol.6, N 2, p.163-168.

139. Gehner K.R. Characterization theorems for constrainted approximation problems via optimization theory. - "J. Approxim. Theory", 1975, vol.14, N 1, p.51-76.

140. Goldstein A.A. Constructive real analysis. New York, Harper and Row, 1967. 178 p.

141. Hoffmann K.-H. Nichtlineare Tschebyscheff - Approximation mit Nebenbedingungen: Anwendungen. - "Numer. Math.", 1970, Bd 14, H.4, S.383-393.

142. Hogan W.W. Directional derivatives for extremal values functions. - "West. Manag. Sci. Inst. Working paper.", Los Angeles, 1971, N 177. 85 p.

143. Johnson L.W. Unicity in approximation of a function and its derivatives. - "Math. Comput.", 1968, vol.22, N 104,p.873-875.

144. Johnson L.W. Uniform approximation of vector-valued functions. - "Numer. Math.", 1969, Bd 13, H.3, S.238-244.

145. Kaufman E.H., Taylor G.D. Uniform rational approximation of functions of several variables. - "Int. J. Numer. Math. Eng.", 1975, vol.9, N 2, p.297-323.

146. Kelley J.P. The cutting-plane method for solving convex problems. - "J. SIAM on Appl. Math.", 1960, vol.8, N 4, p.703-712.

147. Körner W. Tschebyscheff - Approximation durch Polynome.- "Wiss. Z. Hochschule Archit. und Bauwesen (Weimar)", 1964, Bd 11, N 1, S.88-94.

148. Lawson Ch.L. Characteristic properties of the segmented rational minmax approximation problem. - "Numer. Math.", 1964, Bd 6, H.4, S.293-301.

149. Lemerechal C. Note on an extension of "Davidon" methods to non-differentiable functions. - "Math. Progr.", 1974, vol.7, N 3, p.384-387.

150. Lemerechal C. An algorithm for minimizing convex functions. - In: Proceedings IFIP Congress (Stockholm, 1974). Amsterdam, North.-Holland Publ. Co., 1974, p.552-556.

151. Lemerechal C. Minimization of non-differentiable function with constraints. - In: Proceedings of 12th Conf. on Circuit Theory (Allerton, 1974). The Univ. of Illinois, 1974, p.16-24.

152. Lemerechal C. On extension of Davidon methods to non-differentiable problems. - In: Math. Progr. Study 3. Amsterdam, North-Holland Publ. Co., 1975, p.95-109.

153. Lewis J.T. Approximation with convex constraints. - "SIAM Rev.", 1973, vol.15, N 1, p.193-217.

154. Madsen K. An algorithm for minimax solution of overdetermined systems of non-linear equations. - "J. Inst. Math. Appl.", 1975, vol.16, p.321-328.

155. Madsen K. Minimax solution of non-linear equations without calculating derivatives. - In: Math. Progr. Study 3. Amsterdam, North-Holland Publ. Co., 1975, p.110-121.

156. Madsen K. Quadratic convergence of a minimax algorithm.- The Tech. Univ. of Denmark, 1975, Report. N 75-07. 6 p.

157. Moursund D.G. Chebyshev approximation of a function and its derivatives. - "Math. Comput.", 1964, vol.18, N 87,p.382-389.

158. Moursund D.G. Some computational aspects of the uniform approximation of a function and its derivatives. - "SIAM. J.Numer. Anal.", 1965, vol.2, N 3, p.464-472.

159. Meinardus G. Approximation von Funktionen und ihre numerische Behandlung. Berlin, Springer-Verlag, 1964. 180 S.

160. Meinardus G., Schwedt D. Nichtlineare Approximationen.- "Arch. Rational Mech. Anal.", 1964, vol.17, N 4, p.297-326.

161. Miele A., Huang H.Y., Heideman J.C. Sequential gradient - restoration algorithm for the minimization of constrained function. Ordinary and conjugate gradient versions. - "J. Optim. Theory and Appl.", 1969, vol.4, N 4, p.213-243.

162. Motzkin Th. Approximation by curves of a unisolvent family. - "Bull. Amer. Math. Soc.", 1949, vol.55, N 8, p.789-793.

163. Numerische Methoden der Approximationstheorie. Bd.1. Herausgegeben von L. Collatz und G.Meinardus. Basel-Stuttgart,Birkhäuser-Verlag, 1972. 246 S.

164. Optimization. Edited by R.Fletcher. London-New York, Acad. Press, 1969. 354 p.

165. Osborne M.R., Watson G.A. An algorithm for minimax approximation in the non-linear case. - "Comput. J.", 1969, vol.12, N 1, p.63-68.

166. Osborne M.R., Watson G.A. Note on singular minimax approximation problems. - "J. Math. Anal. Appl.", 1969, vol.25,N 3, p.692-700.

167. Paszkowski S. On approximation with nodes. - "Rozprawy Mathematyczne XIV". Warszawa, 1957. 63 s.

168. Powell M.J.D. The exchange algorithm on a discrete point set. - In: Methods of numerical approximation. Ed. D.C.Handcomb. Oxford, Pergamon Press, 1966, p.163.

169. Powell M.J.D. Theory of general non-linear minimax approximation. - In: Methods of numerical approximation. Ed.D.C.Handcomb. Oxford, Pergamon Press, 1966, p.155-162.

170. Powell M.J.D. The minimax solution of linear equations subject to bounds on the variables. - In: Proc. of the Fourth Matibota Conf. on Numer. Math. Winnipeg. Man., 1974, p.53-107.

171. Ralston A. Rational Chebyshev approximation by Remes' algorithms. - "Numer. Math.", 1965, Bd 7, H.4, S.322-330.

172. Rice J.R. Tchebysheff approximations by functions unisolvent of variable degree. - "Trans. Amer. Math. Soc.", 1961, vol.99, N 2, p.298-302.

173. Rice J.R. The approximation of functions. Vol.I. Linear theory. London-Ontario, Addison-Wesley. 1964. 203 p.

174. Rice J.R. The approximation of functions. Vol.II. Nonlinear and multivariable theory. London-Ontario, Addison-Wesley, 1969. 334 p.

175. Robinson S.M. A quadratically convergent algorithm for general non-linear programming problems. - "Math. Progr.", 1972, vol.3, N 2, p.145-156.

176. Taylor G.D. On approximation by polynomials having restricted ranges. - "SIAM J. Numer. Anal", 1968, vol.5, N 2, p.258-268.

177. Taylor G.D. Approximation by functions having restricted ranges. III. - "J. Math. Anal. Appl.", 1969, vol.27, N 2, p.241-248.

178. Tornheim L. On n-parameter families of functions and associated convex functions. - "Trans. Amer. Math. Soc.", 1950, vol.69, N 3, p.457-467.

179. Werner H., Stoer J., Bommass W. Rational Chebyshev approximation. - "Numer. Math.", 1967, Bd 10, H.4, S.289-306.

180. Wetterling W. Anwendung des newtonischen Iterationsverfahrens bei der Tschebyscheff - Approximation, insbesondere mit nichtlinear auftretenden Parametern. I, II. - "MTW", 1963, N 2, S.61-63; N 3, S.112-115.

181. Wolfe P.A. Technique for resolving degeneracy in linear programming. - "SIAM J.", 1963, vol.11, N 2, p.205-211.

182. Wolfe P. Note on a method of conjugate subgradients for minimizing non-differentiable functions. - "Math. Progr.", 1974, vol.7, N 3, p.380-393.

183. Wolfe P. A method of conjugate subgradients for minimizing non-differentiable functions. - In: Math. Progr. Study 3. Amsterdam, North-Holland Publ. Co., 1975, p.145-173.

ИБ № 418

ВОПРОСЫ ТЕОРИИ И ЭЛЕМЕНТЫ
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ МИНИМАКСНЫХ ЗАДАЧ

Редактор А.А.Гранаткина Техн.редактор Л.А.Топорина
Корректор И.А.Богданова

М-17135 Подписано к печати 5 Ш 1977 г.
Формат бумаги 60x90 1/16 Бум.тип. № 3
Печ.л. 12 Уч.-изд.л. 10,19 Бум.л. 6
Тираж 3000 экз. Заказ 356. Цена 1 р. 53 к.

Издательство ЛГУ им.А.А.Жданова
199164, Ленинград, Университетская наб., 7/9

Тульская типография Союзполиграфпрома при Государственном
комитете при Совете Министров СССР по делам издательств,
полиграфии и книжной торговли
г.Тула, пр.Ленина, 109