Раздел I. УСЛОВИЯ ЭКСТРЕМУМА ФУНКЦИЙ

Глава 1. ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ ОПТИМИЗАЦИИ И ОСНОВНЫЕ ПОЛОЖЕНИЯ

Постановка задачи поиска минимума функций содержит:

- *целевую функцию* $f(x)$, где $x = (x_1, ..., x_n)^T$, определенную на n-мерном евклидовом пространстве R^n. Ее значения характеризуют степень достижения цели, во имя которой поставлена или решается задача;
- *множество допустимых решений* $X \subseteq R^n$, среди элементов которого осуществляется поиск.

Требуется найти такой вектор x^* из множества допустимых решений, которому соответствует минимальное значение целевой функции на этом множестве:

$$f(x^*) = \min_{x \in X} f(x). \tag{1.1}$$

З а м е ч а н и я 1.1.

1. Задача поиска максимума функции $f(x)$ сводится к задаче поиска минимума путем замены знака перед функцией на противоположный (рис. 1.1):

$$f(x^*) = \max_{x \in X} f(x) = -\min_{x \in X} \left[-f(x)\right].$$

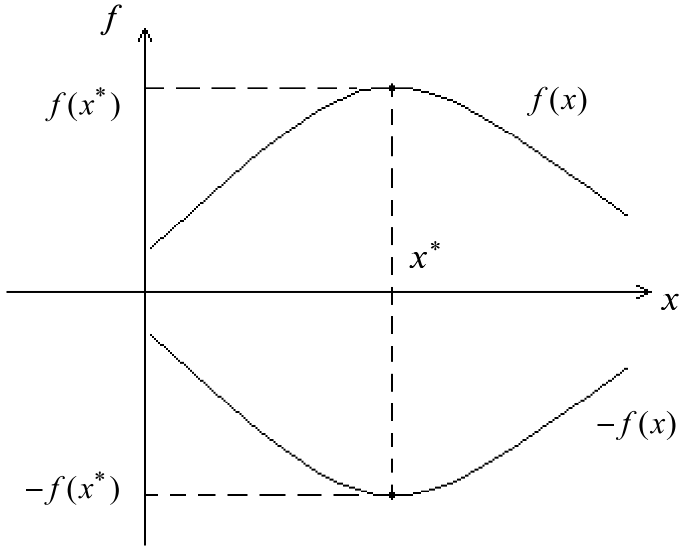


Рис. 1.1

2. Задача поиска минимума и максимума целевой функции $f(x)$ называется задачей поиска *экстремума*: $f(x^*) = \operatorname{extr}_{x \in X} f(x)$.

3. Если множество допустимых решений X задается ограничениями (условиями), накладываемыми на вектор x, то решается задача поиска *условного экстремума*. Если $X = R^n$, т.е. ограничения (условия) на вектор x отсутствуют, решается задача поиска *безусловного экстремума*.

4. Решением задачи поиска экстремума является пара $(x^*, f(x^*))$, включающая точку x^* и значение целевой функции в ней.

5. Множество точек минимума (максимума) целевой функции $f(x)$ на множестве X обозначим X^*. Оно может содержать конечное число точек (в том числе одну), бесконечное число точек или быть пустым.

Определение 1.1. Точка $x^* \in X$ называется точкой *глобального* (*абсолютного*) *минимума* функции $f(x)$ на множестве X, если функция достигает в этой точке своего наименьшего значения, т.е.

$$f(x^*) \le f(x) \quad \forall x \in X .$$

Определение 1.2. Точка $x^* \in X$ называется точкой *локального* (*относительного*) *минимума* функции $f(x)$ на множестве допустимых решений X, если существует $\varepsilon > 0$, такое, что если $x \in X$ и $\left\| x - x^* \right\| < \varepsilon$, то $f(x^*) \le f(x)$. Здесь $\|x\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$ – евклидова норма вектора x.

З а м е ч а н и я 1.2.

1. В определении 1.1 точка x^* сравнивается по величине функции со всеми точками из множества допустимых решений X, а в определении 1.2 – только с принадлежащими ее ε-окрестности (рис. 1.2).

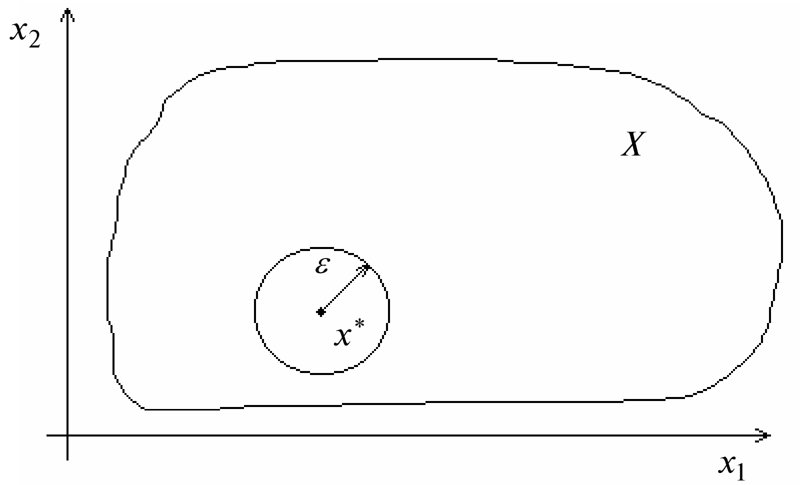


Рис. 1.2

2. Если в определениях 1.1 и 1.2 знак неравенства $\le$ заменить на $\ge$, то получим определения *глобального* (*абсолютного*) и *локального* (*относительного*) *максимумов.*

3. Глобальный экстремум всегда является одновременно локальным, но не наоборот.

Определение 1.3. *Поверхностью уровня* функции $f(x)$ называется множество точек, в которых функция принимает постоянное значение, т.е. $f(x) = \text{const}$. Если $n = 2$, поверхность уровня изображается *линией уровня* на плоскости R^2.

Пример 1.1. Построить линии уровня функций:

а) $f(x) = x_1^2 + x_2^2$; б) $f(x) = \dfrac{x_1^2}{4} + x_2^2$;

в) $f(x) = x_2^2 - x_1^2$; г) $f(x_1, x_2) = x_2^2$.

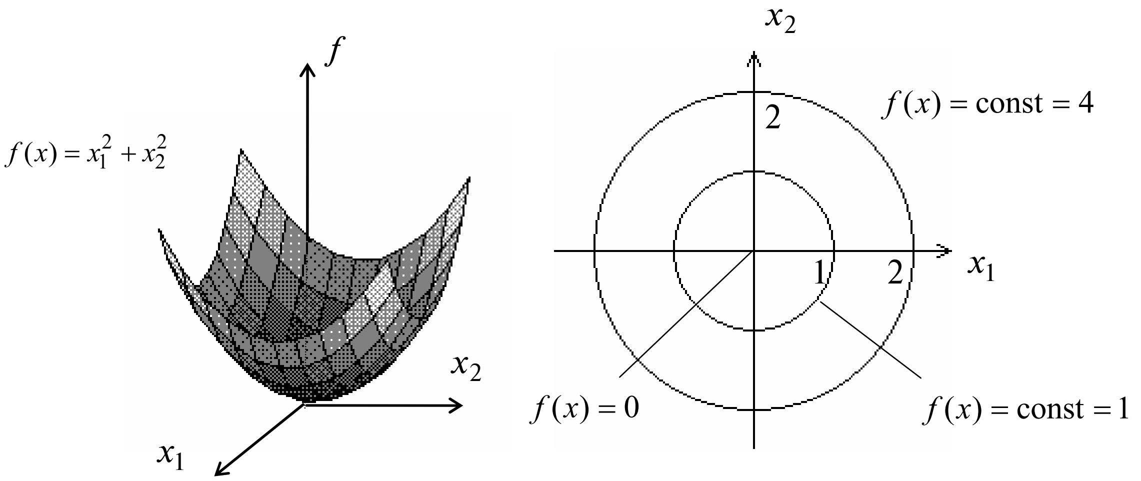
☐ Уравнения линий уровня имеют следующий вид [10]:

а) $f(x) = x_1^2 + x_2^2 = \text{const} = r^2$ – уравнение окружностей с центром в точке $(0,0)^T$ и радиусом, равным r (рис. 1.3, *а*);

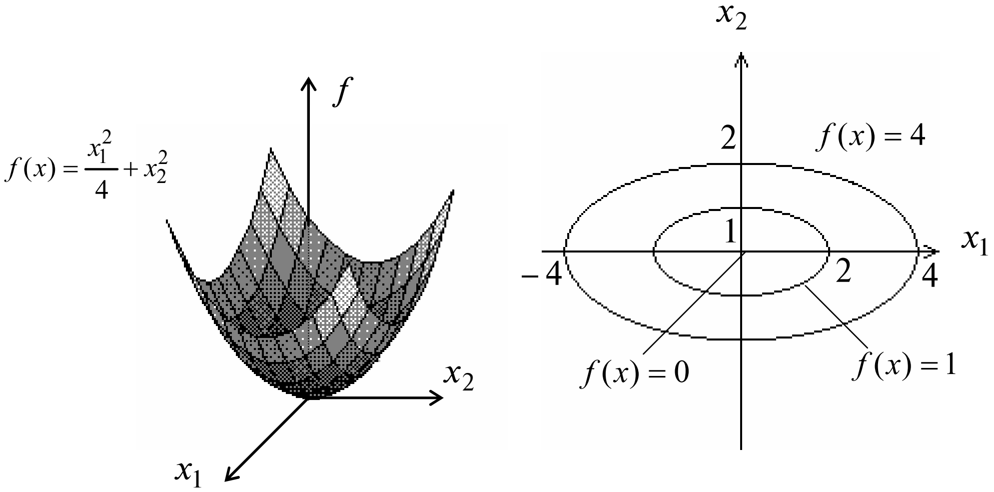
б) $f(x) = \dfrac{x_1^2}{4} + x_2^2 = \text{const}$ – уравнение эллипса. Если $\text{const} = 1$, то $a = 2$ и $b = 1$ – большая и малая полуоси (рис. 1.3, *б*);

в) $f(x) = x_2^2 - x_1^2 = \text{const}$ – уравнение гиперболы (рис. 1.3, *в*);

г) $f(x_1, x_2) = x_2^2 = \text{const}$ – уравнение двух параллельных оси Ox_1 прямых (рис. 1.3, *г*). ■

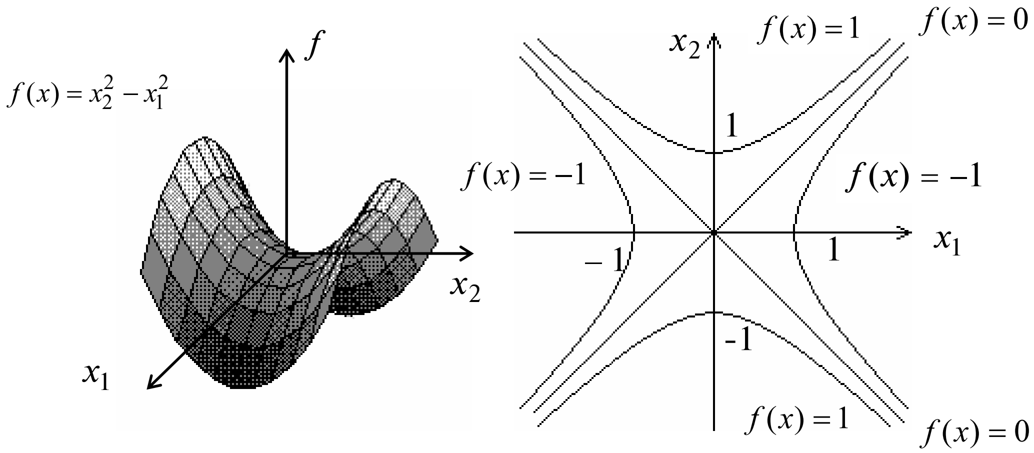


а)

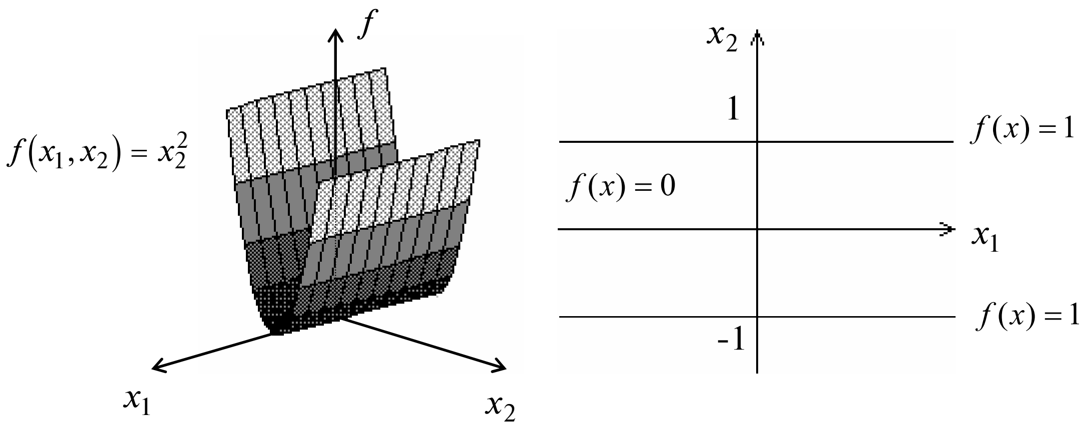


б)

Рис. 1.3



в)



г)

Рис. 1.3 (Окончание)

Пример 1.2. На рис. 1.4 изображены линии уровня функции $f(x)$. Цифры указывают ее значение на соответствующей линии. Точкам A и B соответствуют значения функции $f(A) = 10$ и $f(B) = 5$. Требуется классифицировать точки экстремума.

□ Функция рассматривается на множестве R^2, т.е. решается задача поиска безусловного экстремума. В точке A с координатами $(1,3)^T$ достигается локальный минимум; в точке B с координатами $(3,1)^T$ достигаются локальный и глобальный минимум одновременно; в точке C нет ни минимума, ни максимума, так как по одним направлениям функция убывает, а по другим – возрастает. Заметим, что изображенная структура линий уровня типична для так называемых многоэкстремальных задач.■

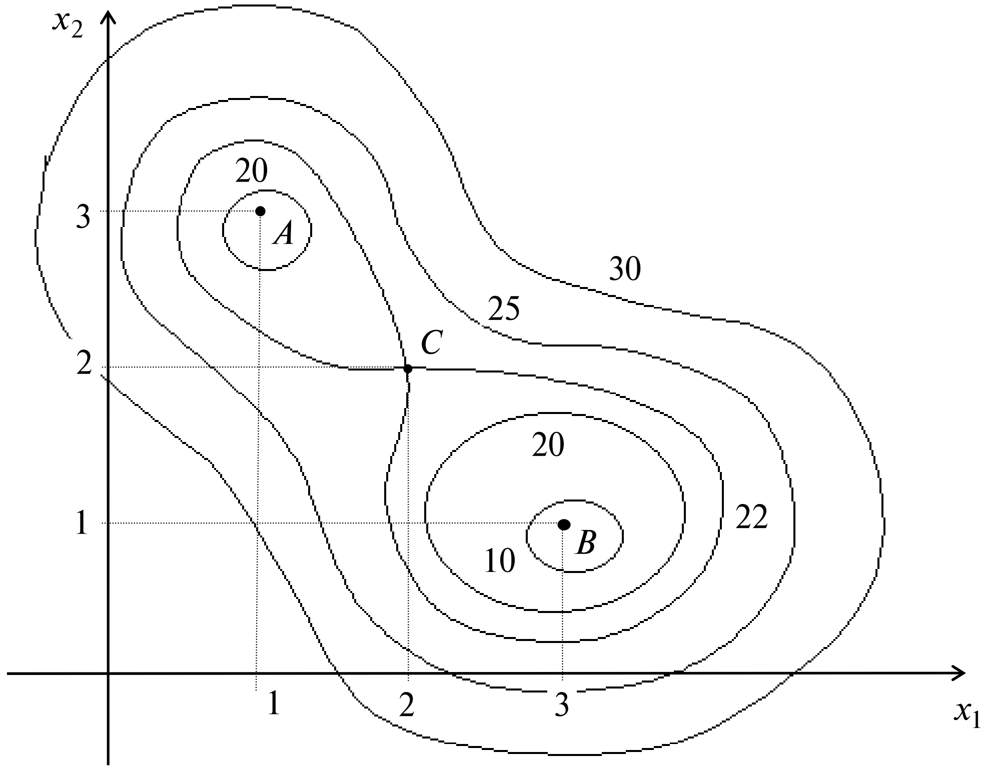


Рис. 1.4

Пример 1.3. На рис. 1.5 изображен график функции $f(x)$, определенной на множестве $X = R$. Требуется классифицировать точки экстремума.

☐ Решается задача поиска безусловного экстремума. На рис. 1.5 выделим ε-окрестности точек $A, B, \ldots, F$ и проверим выполнение определений 1.1 и 1.2 с учетом пп. 1, 2 замечаний 1.1, 1.2. В результате получаем: A – точка локального минимума; B, E – точки локального максимума; бесконечное множество точек из отрезка CD – точки локального минимума; F – точка локального и одновременно глобального минимума; глобальный максимум отсутствует. ■

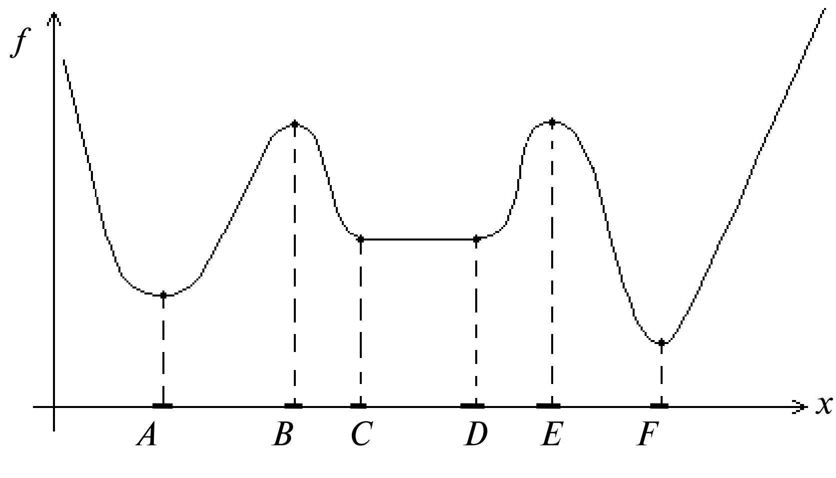


Рис. 1.5

Пример 1.4. Найти точки экстремума функций $f(x) = x_1^2 + x_2^2$ и $f(x) = \dfrac{x_1^2}{4} + x_2^2$ на множестве R^2.

□ Обе функции имеют в точке $x^* = (0,0)^T$ локальный и одновременно глобальный минимум, а локальных и глобальных максимумов не имеют (см. рис. 1.3 *а,б*). ■

Пример 1.5. Найти точки экстремума функции $f(x) = x_1^2 + x_2^2$ на множестве $X = \left\{ x \mid x_2^2 - x_1 + 3 = 0 \right\}$.

□ Решается задача поиска условного экстремума. Линии уровня функции $f(x)$ представляются окружностями (см. рис. 1.3, *а*), а множество допустимых решений X – параболой с уравнением $x_1 = x_2^2 + 3$. В точке $x^* = (3,0)^T$, $f(x^*) = 9$ достигается глобальный минимум (рис. 1.6). Заметим, что эта точка является точкой касания линии уровня и кривой, описывающей множество X. Глобальный максимум на данном множестве не достигается. Если поменять знак перед функцией на противоположный, то в точке x^* функция $f(x) = -x_1^2 - x_2^2$ будет достигать глобального максимума на множестве X, что соответствует п. 1 замечаний 1.1. ■

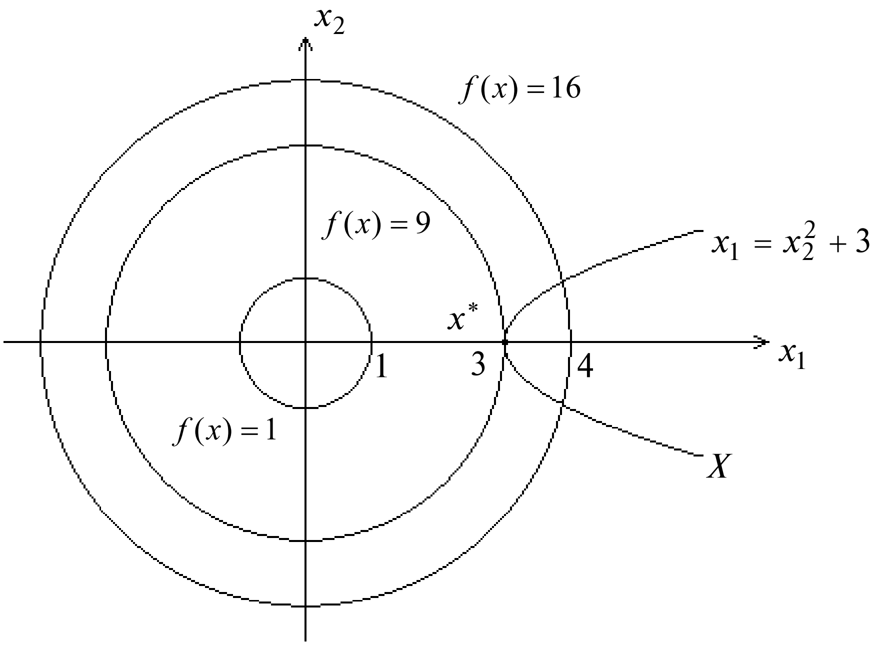


Рис. 1.6

Пример 1.6. Найти точки экстремума функции $f(x) = x_1^2 + x_2^2$ на множестве $X = \left\{ x \mid x_1 + x_2 - 2 = 0 \right\}$.

□ Решается задача поиска условного экстремума. Линии уровня функции $f(x)$ представляются окружностями (см. рис. 1.3, *а*), а множество допустимых решений X – графиком прямой. В точке $x^* = (1,1)^T$, $f(x^*) = 2$ достигается глобальный минимум (рис. 1.7). Глобальный максимум на данном множестве не существует. Заметим, что, как и в примере 1.5, в точке x^* линии уровня целевой функции касаются кривой, описывающей ограничения. ■

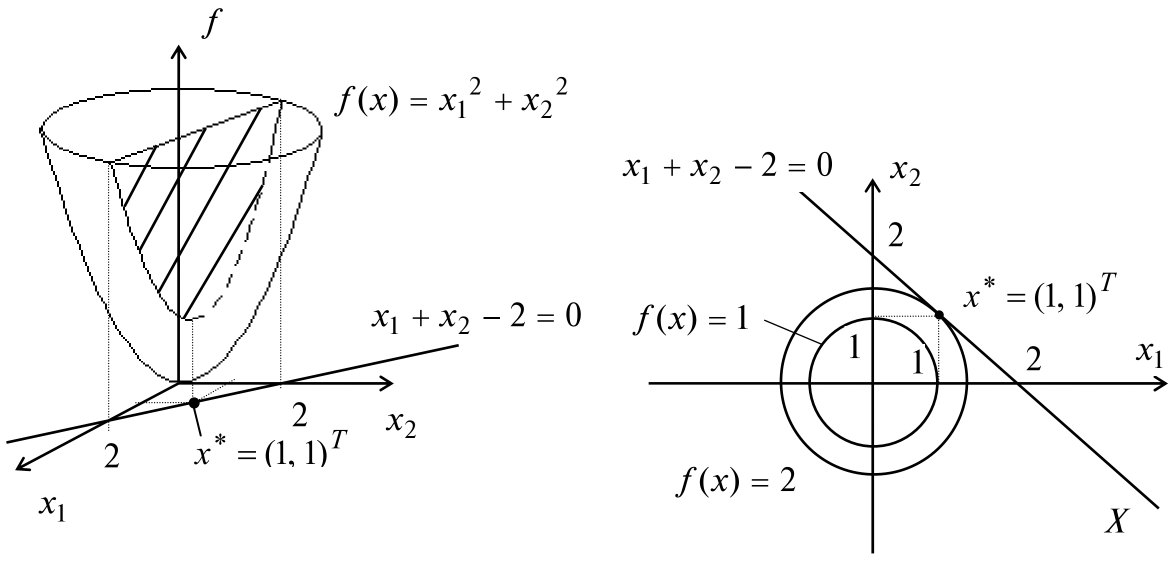


Рис. 1.7

Пример 1.7. Найти точки экстремума функции $f(x) = x_1^2 + x_2^2$ на множестве $X = \left\{ x \mid x_1^2 - x_2^2 - 1 = 0 \right\}$.

□ Решается задача поиска условного экстремума. Линии уровня функции $f(x)$ представляются окружностями (см. рис. 1.3, *а*), а множество X – гиперболой (рис. 1.8). Имеются две точки глобального минимума:

$$x^* = (1,\ 0)^T, \qquad \overline{x}^* = (-1,\ 0)^T, \qquad f(x^*) = f\left(\overline{x}^*\right) = 1.$$

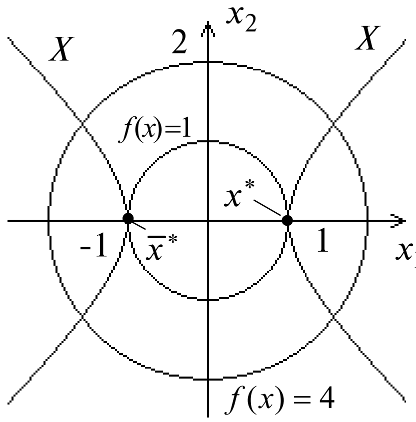


Рис. 1.8

В них выполняется свойство касания линий уровня и кривых, описывающих множество X , отмеченное при решении примеров 1.5, 1.6. Точки глобального и локального максимумов отсутствуют. ■

Пример 1.8. Найти точки экстремума функции $f(x) = x_1$ на множестве допустимых решений $X = \left\{ x \mid x_1 + x_2 \le 1,\ x_1 \ge 0,\ x_2 \ge 0 \right\}$.

□ Решается задача поиска условного экстремума. Линии уровня целевой функции описываются уравнением $f(x) = x_1 = \text{const}$ и представляются прямыми, параллельными оси Ox_2. Множество допустимых решений, где все ограничения выполняются одновременно, на рис. 1.9 заштриховано. В точке $C = (1, 0)^T$, $f(C) = 1$ достигается глобальный максимум, на множестве точек отрезка *AB* достигается глобальный минимум со значением целевой функции $f(A) = f(B) = 0$. ■

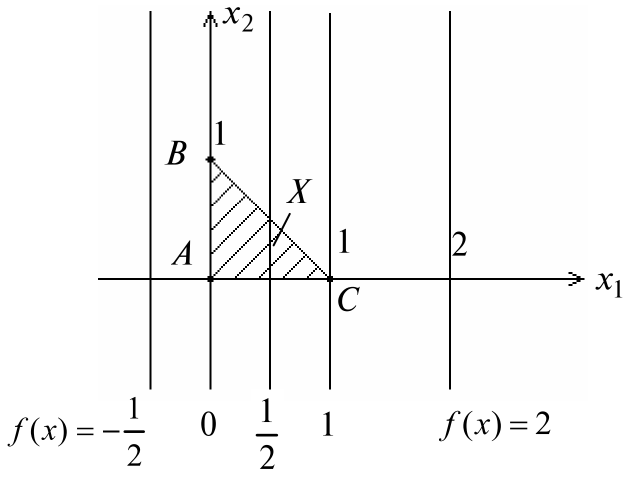


Рис. 1.9

Пример 1.9. Найти точки экстремума функции $f(x) = -x^2$ на множестве $X = \{ x \mid -2 \le x \le 1 \}$.

☐ Решается задача поиска условного экстремума. Глобальный максимум достигается в точке B при $x = 0$, $f(B) = 0$. Локальный минимум достигается в точке C при $x = 1$, $f(C) = -1$, а глобальный минимум – в точке A при $x = -2$, $f(A) = -4$ (рис. 1.10). ■

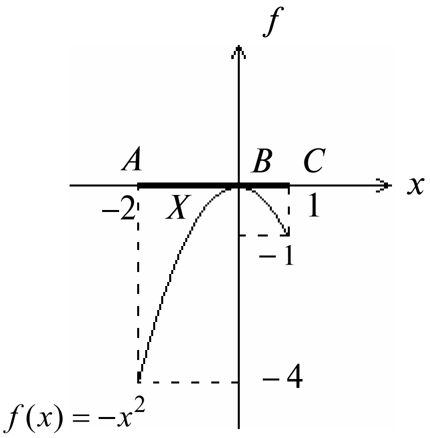


Рис. 1.10

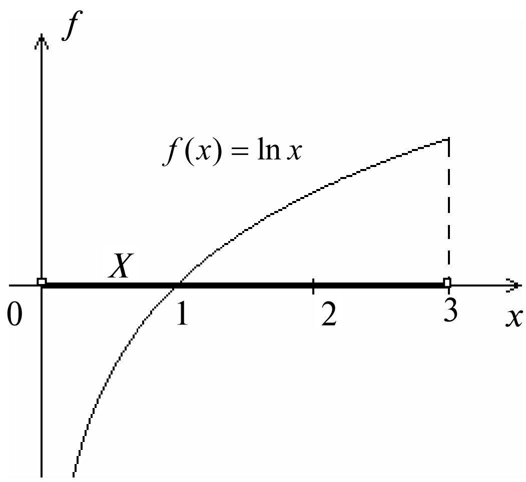


Рис. 1.11

Пример 1.10. Найти точки экстремума функции $f(x) = \ln x$ на множестве $X = \{ x \mid 0 < x < 3 \}$.

☐ Решается задача поиска условного экстремума. Функция не имеет точек локального и глобального экстремумов, так как не выполняются определения 1.1 и 1.2 (рис. 1.11). При приближении к левой границе значение функции стремится к $f^* = -\infty$ (функция не ограничена снизу), а при приближении к правой границе функция возрастает, но не имеет максимума, так как точка $x = 3$ не принадлежит множеству X. ■

Определение 1.4. *Градиентом* $\nabla f(x)$ непрерывно дифференцируемой функции $f(x)$ в точке x называется вектор-столбец, элементами которого являются частные производные первого порядка, вычисленные в данной точке:

$$\nabla f(x) = \begin{pmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \dfrac{\partial f(x)}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$

Градиент функции направлен по нормали к поверхности уровня (см. определение 1.3), т.е. перпендикулярно к касательной плоскости, проведенной в точке x, в сторону наибольшего возрастания функции в данной точке.

Определение 1.5. *Матрицей Гессе* $H(x)$ дважды непрерывно дифференцируемой в точке x функции $f(x)$ называется матрица частных производных второго порядка, вычисленных в данной точке:

$$H(x) = \begin{pmatrix} \dfrac{\partial^2 f(x)}{\partial x_1^2} & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{pmatrix},$$

где $h_{ij} = \dfrac{\partial^2 f(x)}{\partial x_i \, \partial x_j}$, $i, j = 1, \ldots, n$.

З а м е ч а н и я 1.3.

1. Матрица Гессе является симметрической размеров $(n \times n)$.

2. Вместе с градиентом можно определить вектор *антиградиента*, равный по модулю вектору градиента, но противоположный по направлению. Он указывает в сторону наибольшего убывания функции в данной точке.

3. С помощью градиента и матрицы Гессе, используя разложение в ряд Тейлора, приращение функции $f(x)$ в точке x может быть записано в форме

$$\Delta f(x) = f(x + \Delta x) - f(x) = \nabla f(x)^T \Delta x + \frac{1}{2} \Delta x^T H(x) \Delta x + o(\|\Delta x\|^2), \qquad (1.2)$$

где $o(\|\Delta x\|^2)$ – сумма всех членов разложения, имеющих порядок выше второго; $\Delta x^T H(x) \Delta x$ – квадратичная форма.

Пример 1.11. Для функции $f(x)=x_1^2+x_2^2$ необходимо: а) вычислить и построить градиент в точках $x^0=(0,1)^T$, $x^1=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^T$, $x^2=(1,0)^T$, $x^3=(0,-1)^T$; б) найти матрицу Гессе.

□ Согласно определениям 1.4 и 1.5 имеем:

$$\nabla f(x)=(2x_1,2x_2)^T,\quad \nabla f(x^0)=(0,2)^T,\quad \nabla f(x^1)=\left(\sqrt{2},\sqrt{2}\right)^T,\quad \nabla f(x^2)=(2,0)^T,$$

$$\nabla f(x^3)=(0,-2)^T,\qquad H(x)=\begin{pmatrix}2&0\\0&2\end{pmatrix}.$$

Заметим, что матрица Гессе квадратичной функции не зависит от x. На рис. 1.12 изображены найденные градиенты. ■

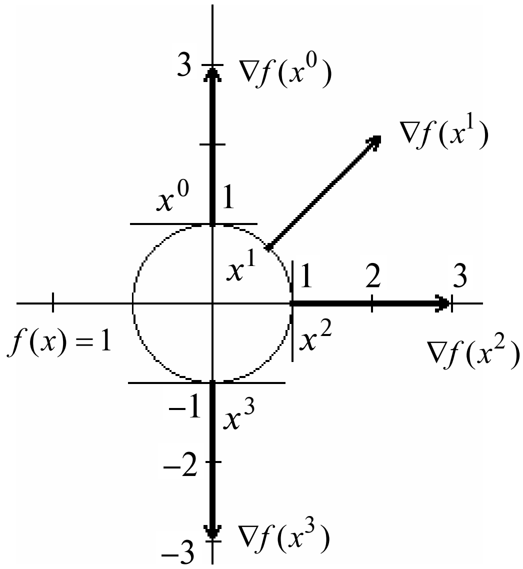


Рис. 1.12

Пример 1.12. Для функции $f(x)=x_1^2+x_2^4$ вычислить градиент и найти матрицу Гессе в точках $x^0=(0,0)^T$, $x^1=(1,1)^T$.

□ Согласно определениям 1.4 и 1.5 имеем:

$$\nabla f(x)=\left(2x_1,4x_2^3\right)^T, H(x)=\begin{pmatrix}2&0\\0&12x_2^2\end{pmatrix};\quad \nabla f(x^0)=(0,0)^T, H(x^0)=\begin{pmatrix}2&0\\0&0\end{pmatrix};$$

$$\nabla f(x^1)=(2,4)^T, H(x^1)=\begin{pmatrix}2&0\\0&12\end{pmatrix}. \blacksquare$$

Определение 1.6. Квадратичная форма $\Delta x^T H(x)\Delta x$ (а также соответствующая матрица Гессе $H(x)$) называется:

- *положительно определенной* $(H(x)>0)$, если для любого ненулевого Δx выполняется неравенство $\Delta x^T H(x)\Delta x>0$;

- *отрицательно определенной* $\left(H(x) < 0\right)$, если для любого ненулевого Δx выполняется неравенство $\Delta x^T H(x) \Delta x < 0$;
- *положительно полуопределенной* $\left(H(x) \ge 0\right)$, если для любого Δx выполняется неравенство $\Delta x^T H(x) \Delta x \ge 0$ и имеется отличный от нуля вектор Δx, для которого $\Delta x^T H(x) \Delta x = 0$;
- *отрицательно полуопределенной* $\left(H(x) \le 0\right)$, если для любого Δx выполняется неравенство $\Delta x^T H(x) \Delta x \le 0$ и имеется отличный от нуля вектор Δx, для которого $\Delta x^T H(x) \Delta x = 0$;
- *неопределенной* ($H(x) \gtrless 0$) , если существуют такие векторы Δx, $\Delta \tilde{x}$, что выполняются неравенства $\Delta x^T H(x) \Delta x > 0$, $\Delta \tilde{x}^T H(x) \Delta \tilde{x} < 0$;
- *тождественно равной нулю* $\left(H(x) \equiv 0\right)$, если для любого Δx выполняется равенство $\Delta x^T H(x) \Delta x = 0$.

Пример 1.13. Классифицировать квадратичную форму и матрицу Гессе $H = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, полученную в примере 1.11.

☐ Выпишем квадратичную форму:

$$\Delta x^T H \, \Delta x = \begin{pmatrix} \Delta x_1 & \Delta x_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 2\Delta x_1^2 + 2\Delta x_2^2 .$$

Очевидно, что $\Delta x^T H \, \Delta x > 0$ для любого ненулевого Δx. Согласно определению 1.6 квадратичная форма (матрица Гессе) положительно определенная. ■

Пример 1.14. Классифицировать квадратичную форму и матрицу Гессе $H = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, полученную в примере 1.12.

☐ Выпишем квадратичную форму:

$$\Delta x^T H \, \Delta x = \begin{pmatrix} \Delta x_1 & \Delta x_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 2 \, \Delta x_1^2 .$$

Очевидно, что $\Delta x^T H\left(x\right) \Delta x \ge 0$ для любого вектора Δx и $\Delta x^T H(x) \Delta x = 0$ для $\Delta x_1 = 0$ и любых $\Delta x_2 \ne 0$. Согласно определению 1.6 квадратичная форма (матрица Гессе) положительно полуопределенная. ■

Пример 1.15. Найти матрицу Гессе целевой функции $f(x) = x_1^2 - x_2^2$ и классифицировать ее.

☐ Следуя определению 1.5, получаем $H(x) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. Выпишем соответствующую квадратичную форму:

$$\Delta x^T H \, \Delta x = \begin{pmatrix} \Delta x_1 & \Delta x_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 2\Delta x_1^2 - 2\Delta x_2^2 .$$

При $\Delta x_1 = 0$ и любых $\Delta x_2 \neq 0$ квадратичная форма отрицательна, а при $\Delta x_1 \neq 0$ и $\Delta x_2 = 0$ положительна. Согласно определению 1.6 квадратичная форма (матрица Гессе) неопределенная. ■

Определение 1.7. Множество $X \subseteq R^n$ называется *выпуклым*, если оно содержит всякий отрезок, концы которого принадлежат X, т.е. если для любых $x^1, x^2 \in X$ и $0 \le \lambda \le 1$ справедливо $\lambda x^1 + (1-\lambda)x^2 \in X$.

Пример 1.16. Требуется выбрать выпуклые множества среди изображенных на рис. 1.13.

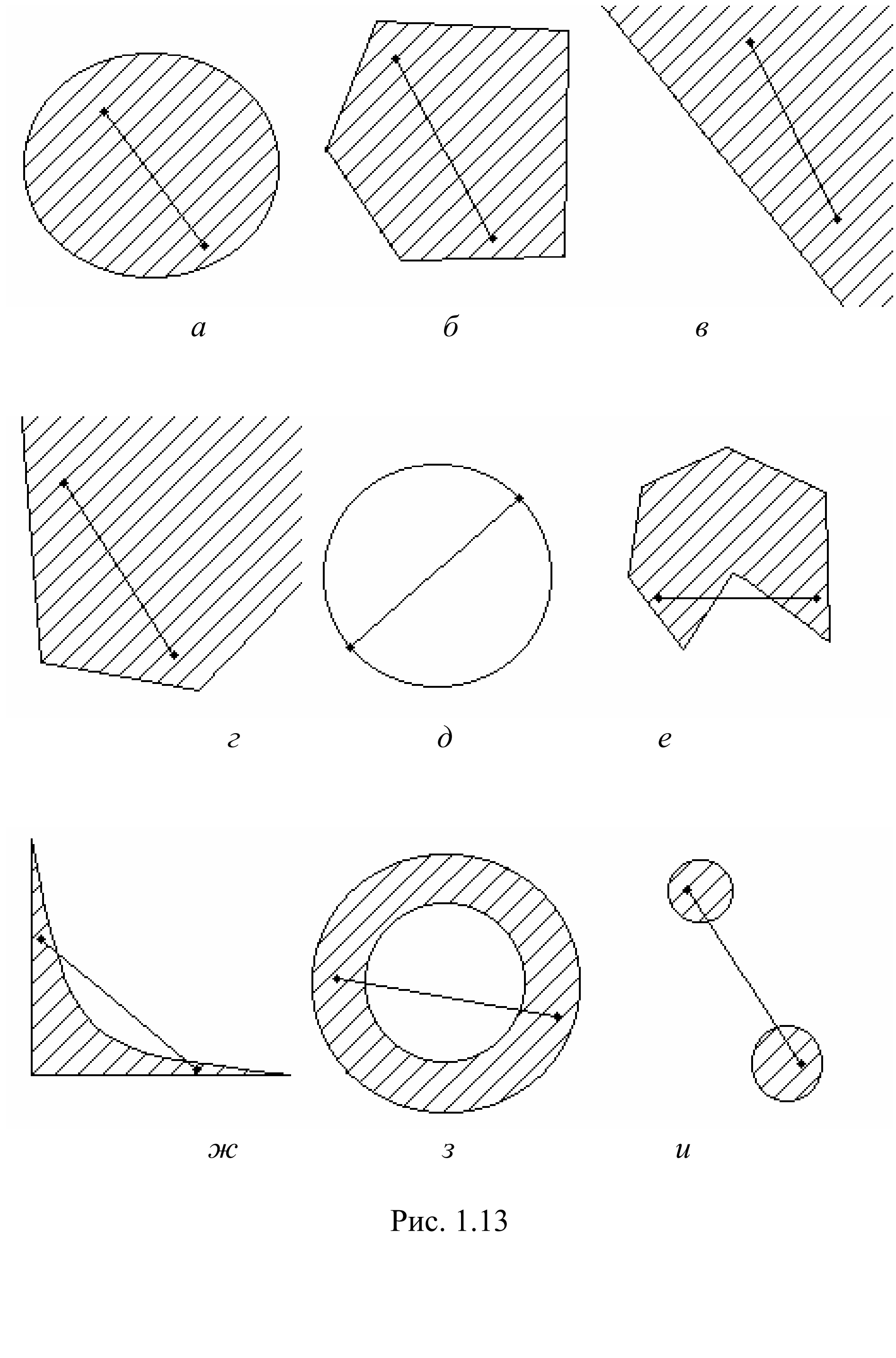


Рис. 1.13

☐ На рис. 1.13, *а–г* множества выпуклые, так как удовлетворяют определению 1.7, а остальные не являются выпуклыми. ■

Образно говоря, выпуклыми являются множества, которые не содержат «вмятин», «дырок» и состоят из одного «куска». Примерами выпуклых множеств служат также само пространство R^n, отрезок, прямая, шар.

Определение 1.8. Функция $f(x)$, определенная на выпуклом множестве X, называется *выпуклой*, если $f\left(\lambda x^1 + (1-\lambda) x^2\right) \le \lambda f(x^1) + (1-\lambda) f(x^2) \quad \forall x^1, x^2 \in X$, $0 \le \lambda \le 1$.

Определение 1.9. Функция $f(x)$, определенная на выпуклом множестве X, называется *строго выпуклой*, если $f\left(\lambda x^1 + (1-\lambda) x^2\right) < \lambda f(x^1) + (1-\lambda) f(x^2) \quad \forall x^1, x^2 \in X$, $x^1 \ne x^2$, $0 < \lambda < 1$.

Определение 1.10. Функция $f(x)$, определенная на выпуклом множестве X, называется *сильно выпуклой* с константой $l > 0$, если

$$f\left(\lambda x^1 + (1-\lambda) x^2\right) \le \lambda f(x^1) + (1-\lambda) f(x^2) - \frac{l}{2} \lambda (1-\lambda) \left\| x^1 - x^2 \right\|^2$$

$$\forall x^1, x^2 \in X, \;\; x^1 \ne x^2, \;\; 0 \le \lambda \le 1.$$

З а м е ч а н и я 1.4.

1. Функцию $f(x)$ называют *выпуклой*, если ее график целиком лежит не выше отрезка, соединяющего две произвольные точки графика. Функцию называют *строго выпуклой*, если ее график целиком лежит ниже отрезка, соединяющего две произвольные, но не совпадающие точки графика.

2. Если функция сильно выпуклая, то она одновременно строго выпуклая и выпуклая. Если функция строго выпуклая, то она одновременно выпуклая.

3. Выпуклость функции можно определить по матрице Гессе:

- если $H(x) \ge 0 \;\; \forall x \in R^n$, то функция выпуклая;
- если $H(x) > 0 \;\; \forall x \in R^n$, то функция строго выпуклая;
- если $H(x) \ge lE \;\; \forall x \in R^n$, где E – единичная матрица, то функция сильно выпуклая.

Пример 1.17. Дана функция $f(x) = x^2$, определенная на множестве $X = \left\{ x \,\middle|\, -1 \le x \le 1 \right\}$ (рис. 1.14). Требуется исследовать ее на выпуклость.

☐ Функция является строго выпуклой согласно п. 1 замечаний 1.4, так как ее график целиком лежит ниже отрезка, соединяющего две произвольные, но не совпадающие точки графика (см. рис. 1.14). Более того, функция одновременно является сильно выпуклой, так как согласно п. 3 замечаний 1.4 выполняется условие $H(x) = f''(x) = 2 \ge l$ при $0 < l \le 2$. Очевидно, условия выпуклости и строгой выпуклости также выполняются (см. п.3 замечаний 1.4), что иллюстрирует справедливость п.2 замечаний 1.4. ■

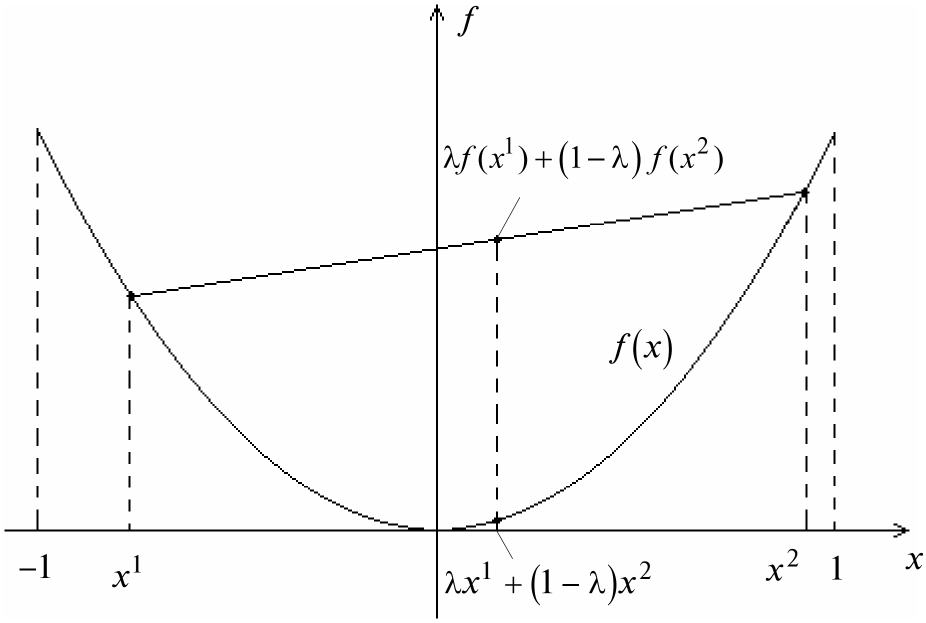


Рис. 1.14

Пример 1.18. Дана функция $f(x)=x$, определенная на множестве $X=\{x \mid 0 \le x \le 1\}$ (рис. 1.15). Требуется исследовать ее на выпуклость.

□ Согласно п. 1 замечаний 1.4 функция является выпуклой, так как ее график целиком лежит не выше отрезка, соединяющего две произвольные точки графика, но не является строго выпуклой и тем более сильно выпуклой. ■

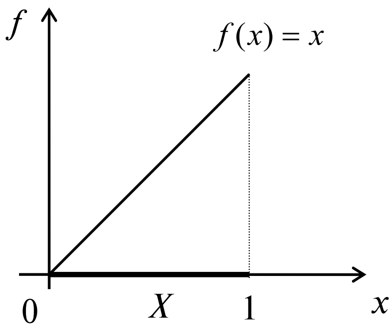


Рис. 1.15

Пример 1.19. Исследовать выпуклость функции $f(x)=x_1^2+x_2^2$ на множестве R^2.

□ Согласно результату примера 1.11 матрица Гессе удовлетворяет условию $H(x)=\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix} \ge l\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ при $0<l\le 2$. Следуя п. 3 замечаний 1.4, можно сделать вывод о сильной выпуклости функции. Одновременно она является строго выпуклой и выпуклой (см. п. 2 замечаний 1.4). ■

Далее при решении примеров используются следующие свойства выпуклых функций.

Утверждение 1.1.

1. *Если* $f(x)$ *выпуклая функция на выпуклом множестве X, то всякая точка локального минимума является точкой ее глобального минимума на X* (см. пример 1.6).

2. *Если выпуклая функция достигает своего минимума в двух различных точках, то она достигает минимума во всех точках отрезка, соединяющего эти две точки* (см. пример 1.8).

3. *Если* $f(x)$ *строго выпуклая функция на выпуклом множестве X, то она может достигать своего глобального минимума на X не более чем в одной точке* (см. пример 1.6).

Определение 1.11. Функция $f(x)$ удовлетворяет *условию Липшица* на отрезке $[a, b]$, если существует такое число $L > 0$ (константа Липшица), что

$$\left| f(x') - f(x'') \right| \le L \left| x' - x'' \right| \tag{1.3}$$

для всех x' и x'', принадлежащих $[a, b]$.

З а м е ч а н и я 1.5.

1. Если неравенство (1.3) выполняется с константой L, то оно справедливо для бесконечного множества констант, больших L. Как правило, представляет интерес минимальная из констант Липшица.

2. Из условия (1.3) следует непрерывность функции $f(x)$ на отрезке $[a, b]$. Если, кроме того, функция имеет на $[a, b]$ непрерывную производную, то константа Липшица $L = \max\limits_{x \in [a,b]} \left| f'(x) \right|$.

3. Условие (1.3) означает, что модуль углового коэффициента любой хорды графика функции $f(x)$ не превосходит L.

Пример 1.20. Проверить, удовлетворяют ли условию Липшица следующие функции: а) $f(x) = 2x$ на отрезке $[0, 1]$; б) $f(x) = \sin x$ на отрезке $[0, \pi]$; в) $f(x) = \sqrt{x}$ на отрезке $[0, 1]$.

□ Воспользуемся определением 1.11 и п.2 замечаний 1.5:

а) функция $f(x) = 2x$ удовлетворяет условию Липшица на отрезке $[0, 1]$ с константой $L = 2$;

б) функция $f(x) = \sin x$ удовлетворяет условию Липшица на отрезке $[0, \pi]$ с константой $L = \max\limits_{x \in [0,\pi]} \left| \cos x \right| = 1$;

в) функция $f(x) = \sqrt{x}$ не удовлетворяет условию Липшица на отрезке $[0, 1]$, так как при $x \to +0$ угловой коэффициент касательной к графику неограниченно возрастает, а переходя в (1.3) к пределу при $\left| x' - x'' \right| \to 0$, можно заключить, что если в некоторой точке существует касательная к графику функции $f(x)$, то модуль ее углового коэффициента не может превышать L. ■

Задачи для самостоятельного решения

1. Найти точки экстремума функции $f(x) = \dfrac{(x_1 - 3)^2}{4} + \dfrac{(x_2 + 2)^2}{9}$ на множестве R^2.

Ответ: в точке $x^* = (3, -2)^T$ – локальный и одновременно глобальный безусловный минимум.

2. Найти точки экстремума функции $f(x) = -(x_1 + 1)^2 - \dfrac{(x_2 + 4)^2}{16}$ на множестве R^2.

Ответ: в точке $x^* = (-1, -4)^T$ – локальный и одновременно глобальный безусловный максимум.

3. Найти точки экстремума функции $f(x) = \dfrac{1}{x}$ на множестве $X = \{x \mid 0 < x < 2\}$.

Ответ: функция не имеет точек локального и глобального экстремумов.

4. Найти точки экстремума функции $f(x) = (x_1 - 1)^2 + (x_2 - 1)^2$ на множестве $X = \{x \mid x_1 + x_2 = 0\}$.

Ответ: в точке $x^* = (0,0)^T$ – локальный и одновременно глобальный условный минимум.

5. Найти точки экстремума функции $f(x) = x_1^2 - 4x_1x_2 + 4x_2^2$ на множестве R^2.

Ответ: так как $f(x) = (x_1 - 2x_2)^2$, то во всех точках прямой с уравнением $x_1 = 2x_2$ достигается локальный и одновременно глобальный минимум.

6. Проверить знакоопределенность матрицы Гессе целевой функции $f(x) = x_1^3 + x_2^3 - 3x_1x_2$ в точке $(0,0)^T$.

Ответ: матрица Гессе и соответствующая квадратичная форма неопределенные.

7. Проверить знакоопределенность матрицы Гессе целевой функции $f(x) = \dfrac{x_1^2}{4} + x_2^2$. Исследовать целевую функцию на выпуклость.

Ответ: матрица Гессе положительно определенная. Функция является сильно выпуклой, так как $H(x) \geq l\,E$ при $0 < l \leq \dfrac{1}{2}$.

8. Проверить знакоопределенность матрицы Гессе целевой функции $f(x) = x_1^2 - 4x_1x_2 + 4x_2^2$. Исследовать целевую функцию на выпуклость.

Ответ: матрица Гессе положительно полуопределенная, функция выпуклая.

Глава 2. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ БЕЗУСЛОВНОГО ЭКСТРЕМУМА

Постановка задачи

Дана дважды непрерывно дифференцируемая функция $f(x)$, определенная на множестве $X = R^n$.

Требуется исследовать функцию $f(x)$ на экстремум, т.е. определить точки $x^* \in R^n$ ее локальных минимумов и максимумов на R^n:

$$f(x^*) = \min_{x \in R^n} f(x); \qquad f(x^*) = \max_{x \in R^n} f(x). \tag{2.1}$$

Стратегия решения задачи

Находятся точки x^* локальных экстремумов с помощью необходимых условий первого и второго порядка (порядок условий определяется порядком используемых производных), а также достаточных условий безусловного локального экстремума. Вычисляются значения $f(x^*)$ функции в найденных точках локальных экстремумов.

Утверждение 2.1 (необходимые условия экстремума первого порядка).

Пусть $x^* \in R^n$ есть точка локального минимума (максимума) функции $f(x)$ на множестве R^n и $f(x)$ дифференцируема в точке x^*. Тогда градиент функции $f(x)$ в точке x^* равен нулю, т.е.

$$\nabla f(x^*) = 0 \tag{2.2}$$

или

$$\frac{\partial f(x^*)}{\partial x_i} = 0, \qquad i = 1, \dots, n. \tag{2.3}$$

Определение 2.1. Точки x^*, удовлетворяющие условию (2.2) или (2.3), называются *стационарными.*

Утверждение 2.2 (необходимые условия экстремума второго порядка).

Пусть точка x^* есть точка локального минимума (максимума) функции $f(x)$ на множестве R^n и функция $f(x)$ дважды дифференцируема в этой точке. Тогда матрица Гессе $H(x^*)$ функции $f(x)$, вычисленная в точке x^*, является положительно полуопределенной (отрицательно полуопределенной), т.е.

$$H(x^*) \ge 0, \tag{2.4}$$

$$(\ H(x^*) \le 0\ ). \tag{2.5}$$

Утверждение 2.3 (достаточные условия экстремума).

Пусть функция $f(x)$ *в точке* $x^* \in R^n$ *дважды дифференцируема, ее градиент равен нулю, а матрица Гессе является положительно определенной (отрицательно определенной), т.е.*

$$\nabla f(x^*) = 0 \quad \text{и} \quad H(x^*) > 0, \tag{2.6}$$

$$(\, H(x^*) < 0 \,). \tag{2.7}$$

Тогда точка x^* *есть точка локального минимума (максимума) функции* $f(x)$ *на множестве* R^n.

Определение 2.2. Рассмотрим определитель матрицы Гессе $H(x^*)$, вычисленной в стационарной точке

$$\det H(x^*) = \begin{vmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{vmatrix}.$$

1. Определители $\Delta_1 = h_{11}$, $\Delta_2 = \begin{vmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{vmatrix}, \ldots,$ $\Delta_n = \begin{vmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & \ddots & \vdots \\ h_{n1} & \cdots & h_{nn} \end{vmatrix}$ называются *угловыми минорами*.

2. Определители m-го порядка ($m \le n$), получающиеся из определителя матрицы $H(x^*)$ вычеркиванием каких-либо ($n - m$) строк и ($n - m$) столбцов с одними и теми же номерами, называются *главными минорами.*

Для проверки выполнения достаточных условий экстремума и необходимых условий второго порядка используются два способа.

Первый способ (с помощью угловых и главных миноров).

- **Критерий проверки достаточных условий экстремума** (критерий Сильвестра).

Для того чтобы матрица Гессе $H(x^*)$ *была положительно определенной* ($H(x^*) > 0$) *и точка* x^* *являлась точкой локального минимума, необходимо и достаточно, чтобы знаки угловых миноров были строго положительны:*

$$\Delta_1 > 0, \quad \Delta_2 > 0, \ldots, \quad \Delta_n > 0. \tag{2.8}$$

Для того чтобы матрица Гессе $H(x^*)$ *была отрицательно определенной* ($H(x^*) < 0$) *и точка* x^* *являлась точкой локального максимума, необходимо и достаточно, чтобы знаки угловых миноров чередовались, начиная с отрицательного:*

$$\Delta_1 < 0, \; \Delta_2 > 0, \; \Delta_3 < 0, \ldots, \; (-1)^n \Delta_n > 0. \tag{2.9}$$

- **Критерий проверки необходимых условий экстремума второго порядка.**

1. *Для того чтобы матрица Гессе* $H(x^*)$ *была положительно полуопределенной* ($H(x^*) \geq 0$) *и точка* x^* *может быть являлась точкой локального минимума, необходимо и достаточно, чтобы все главные миноры определителя матрицы Гессе были неотрицательны.*

2. *Для того чтобы матрица Гессе* $H(x^*)$ *была отрицательно полуопределенной* ($H(x^*) \leq 0$) *и точка* x^* *может быть являлась точкой локального максимума, необходимо и достаточно, чтобы все главные миноры четного порядка были неотрицательны, а все главные миноры нечетного порядка – неположительны.*

Второй способ (с помощью собственных значений матрицы Гессе).

Определение 2.3. *Собственные значения* λ_i, $i = 1,\ldots,n$, матрицы $H(x^*)$ размеров $(n \times n)$ находятся как корни характеристического уравнения (алгебраического уравнения n-й степени):

$$\left| H(x^*) - \lambda E \right| = \begin{vmatrix} h_{11} - \lambda & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} - \lambda & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \ldots & h_{nn} - \lambda \end{vmatrix} = 0. \quad (2.10)$$

З а м е ч а н и е 2.1. Собственные значения вещественной симметрической матрицы $H(x^*)$ вещественные.

Оба способа проверки достаточных и необходимых условий экстремума второго порядка приведены в табл. 2.1.

Алгоритм решения задачи

Шаг 1. Записать необходимые условия экстремума первого порядка в виде (2.3) и найти стационарные точки x^* в результате решения системы n в общем случае нелинейных алгебраических уравнений с n неизвестными. Для численного решения системы могут использоваться методы простой итерации, Зейделя, Ньютона.

Шаг 2. В найденных стационарных точках x^* проверить выполнение достаточных, а если они не выполняются, то необходимых условий второго порядка с помощью одного из двух способов (см. табл. 2.1).

Шаг 3. Вычислить значения $f(x^*)$ в точках экстремума.

Описанный алгоритм отображен на рис. 2.1, где показана последовательность действий в случаях выполнения и невыполнения соответствующих условий экстремума при применении первого способа.

З а м е ч а н и я 2.2.

1. Продолжение исследований, которое требуется в ряде случаев, разобранных в табл. 2.1, при решении практических задач, как правило, не проводится, за исключением небольшого числа модельных примеров.

2. Если требуется определить глобальные экстремумы, то они находятся в результате сравнения значений функции в точках локальных минимумов и максимумов с учетом ограниченности функции на R^n.

3. Для случая функции $f(x)$ одной переменной ($n=1$) можно сформулировать следующее правило, заменяющее п. 2 алгоритма.

Если функция $f(x)$ и ее производные непрерывны, то точка x^* является точкой экстремума тогда и только тогда, когда число m – четное, где m – порядок первой не обращающейся в нуль в точке x^* производной. Если $f^{(m)}(x^*)>0$, то в точке x^* – локальный минимум, а если $f^{(m)}(x^*)<0$, то в точке x^* – локальный максимум. Если число m – нечетное, в точке x^* нет экстремума.

4. Часто на практике, особенно при применении численных методов поиска экстремума, рассматриваемых в последующих разделах, требуется проверить, выполняются ли необходимые и достаточные условия экстремума в некоторой точке. Такой анализ необходим, так как многие численные методы позволяют найти лишь стационарную точку, тип которой требует уточнения.

Рис. 2.1

Пример 2.1. Найти экстремум функции $f(x)=x_1^2+x_2^2$ на множестве R^2.

☐ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1}=2x_1=0; \qquad \frac{\partial f(x)}{\partial x_2}=2x_2=0.$$

В результате решения системы получим стационарную точку $x^*=(0,0)^T$.

Критерии проверки достаточных и необходимых условий второго порядка в задаче поиска безусловного экстремума

Таблица 2.1

№ п/п	$\nabla f(x^*)$	$H(x^*)$	Вид условий	Первый способ	Второй способ	Тип стационарной точки x^*
1	0	> 0	Достаточные условия экстремума	$\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_n > 0$	$\lambda_1 > 0, \ldots, \lambda_n > 0$	Локальный минимум
2	0	< 0	Достаточные условия экстремума	$\Delta_1 < 0, \Delta_2 > 0, \ldots, (-1)^n \Delta_n > 0$	$\lambda_1 < 0, \ldots, \lambda_n < 0$	Локальный максимум
3	0	≥ 0	Необходимые условия экстремума второго порядка	Все главные миноры определителя матрицы $H(x^*)$ неотрицательны	$\lambda_1 \geq 0, \ldots, \lambda_n \geq 0$	Может быть локальный минимум, требуется дополнительное исследование
4	0	≤ 0	Необходимые условия экстремума второго порядка	Все главные миноры четного порядка неотрицательны, а нечетного порядка неположительны	$\lambda_1 \leq 0, \ldots, \lambda_n \leq 0$	Может быть локальный максимум, требуется дополнительное исследование
5	0	$= 0$	Необходимые условия экстремума второго порядка	Матрица Гессе состоит из нулевых элементов	$\lambda_1 = 0, \ldots, \lambda_n = 0$	Требуется дополнительное исследование
6	0	$\gtrless 0$	Необходимые условия экстремума второго порядка	Не выполняются условия п. 1–5	λ_i имеют разные знаки	Нет экстремума

2. Проверим выполнение достаточных условий экстремума.

Первый способ. Матрица Гессе имеет вид $H(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. Так как $\Delta_1 = h_{11} = 2 > 0$, $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0$, то в точке x^* локальный минимум (строка 1 в табл. 2.1).

Второй способ. Найдем собственные значения матрицы Гессе, используя (2.10): $\begin{vmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = 0$. Отсюда $(2-\lambda)^2 = 0$ и $\lambda_1 = \lambda_2 = 2 > 0$. Так как все собственные значения положительны, то в точке x^* локальный минимум (строка 1 в табл. 2.1). Из примера 1.19 следует, что функция является строго выпуклой на множестве R^2. Поэтому точка локального минимума является и точкой глобального минимума (см. п. 3 утверждения 1.1).

3. Вычислим значение функции в точке глобального минимума: $f(x^*) = 0$. ■

Пример 2.2. Найти экстремум функции $f(x) = x_1^2 - x_2^2$ на множестве R^2.

□ 1. Запишем необходимые условия первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 = 0; \qquad \frac{\partial f(x)}{\partial x_2} = -2x_2 = 0.$$

В результате решения системы получим стационарную точку $x^* = (0,0)^T$.

2. Проверим выполнение достаточных условий экстремума.

Первый способ. Матрица Гессе имеет вид $H(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. Так как $\Delta_1 = h_{11} = 2 > 0$, $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4 < 0$, то достаточные условия экстремума не выполняются (строки 1 и 2 в табл. 2.1). Согласно схеме (см. рис. 2.1) проверим выполнение необходимых условий второго порядка. Главные миноры первого порядка ($m = 1$) получаются из Δ_2 в результате вычеркивания $n - m = 2 - 1 = 1$ строк и столбцов с одинаковыми номерами: –2, 2. Главный минор второго порядка ($m = 2$) получается из Δ_2 в результате вычеркивания $n - m = 0$ строк и столбцов, т.е. совпадает с Δ_2: -4. Отсюда следует, что необходимые условия экстремума второго порядка не выполняются (строки 3 и 4 в табл. 2.1). Так как матрица Гессе не является нулевой, то можно сделать вывод о том, что в точке x^* нет экстремума (строка 6 в табл. 2.1).

Второй способ. Найдем собственные значения матрицы Гессе, используя (2.10):

$$\begin{vmatrix} 2-\lambda & 0 \\ 0 & -2-\lambda \end{vmatrix} = (2-\lambda)(-2-\lambda) = 0.$$

Отсюда $\lambda_1 = 2 > 0$, $\lambda_2 = -2 < 0$, т.е. собственные значения имеют разные знаки. Поэтому точка x^* не является точкой минимума или максимума (строка 6 в табл. 2.1), а является *седловой* точкой (аналогична изображенной на рис. 1.3, *в*).

3. Функция не имеет экстремума на множестве R^2. ■

Пример 2.3. Найти экстремум функции $f(x) = x_1^2 + x_2^4$ на множестве R^2.

☐ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 = 0; \qquad \frac{\partial f(x)}{\partial x_2} = 4x_2^3 = 0.$$

В результате решения системы получим стационарную точку $x^* = (0,0)^T$.

2. Проверим выполнение достаточных условий экстремума. Матрица Гессе имеет вид $H(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^{*2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Так как $\Delta_1 = h_{11} = 2 > 0$, $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & 0 \end{vmatrix} = 0$, то достаточные условия экстремума не выполняются (строки 1 и 2 в табл. 2.1). Согласно схеме (см. рис. 2.1) проверим выполнение необходимых условий экстремума второго порядка. Аналогично решению примера 2.2 получаем главные миноры первого порядка: 2, 0 и главный минор второго порядка: 0. Так как все главные миноры неотрицательные, то в точке x^* может быть минимум и требуется дополнительное исследование (строка 3 в табл. 2.1).

3. Вычислим значение целевой функции в точке x^*: $f(x^*) = 0$ и рассмотрим ε-окрестность точки x^*, а также поведение функции $f(x)$ на множестве R^2. При любых $x \in R^2$ имеем $f(x) \geq f(x^*) = 0$ (см. рис. 1.2), что соответствует не только определению 1.2, но и определению 1.1. Поэтому точка x^* является точкой глобального минимума. ■

Пример 2.4. Найти экстремум целевой функции $f(x) = 2x_1^2 + x_1 x_2 + x_2^2$ на множестве R^2.

☐ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = 4x_1 + x_2 = 0; \qquad \frac{\partial f(x)}{\partial x_2} = x_1 + 2x_2 = 0.$$

В результате решения системы получим стационарную точку $x^* = (0,\ 0)^T$.

2. Проверим выполнение достаточных условий экстремума первым способом. Матрица Гессе имеет вид $H(x^*) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$. Так как $\Delta_1 = 4 > 0$, $\Delta_2 = 8 - 1 = 7 > 0$, то точка x^* является точкой локального минимума (строка 1 в табл. 2.1) и $H(x^*) > 0$. Согласно п. 3 замечаний 1.4 функция является строго выпуклой. Поэтому точка x^* – точка глобального минимума (см. п. 3 утверждения 1.1).

3. Вычислим значение функции в точке глобального минимума: $f(x^*) = 0$. ■

Пример 2.5. Найти экстремум функции $f(x) = (1 - x_1)^2 + 10\left(x_2 - x_1^2\right)^2$ на множестве R^2.

☐ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = -2(1 - x_1) - 40x_1\left(x_2 - x_1^2\right) = 0; \qquad \frac{\partial f(x)}{\partial x_2} = 20\left(x_2 - x_1^2\right) = 0.$$

В результате решения системы получим стационарную точку $x^* = (1,1)^T$.

2. Проверим выполнение достаточных условий экстремума первым способом.

Матрица Гессе имеет вид $H(x^*) = \begin{pmatrix} 82 & -40 \\ -40 & 20 \end{pmatrix}$. Так как $\Delta_1 = 82 > 0$, $\Delta_2 = 82 \cdot 20 - 40^2 = 1640 - 1600 = 40 > 0$, то в точке x^* локальный минимум (строка 1 в табл. 2.1) и $H(x^*) > 0$. Согласно п. 3 замечаний 1.4 функция является строго выпуклой. Поэтому точка x^* – точка глобального минимума (см. п. 3 утверждения 1.1).

3. Вычислим значение функции в точке глобального минимума: $f(x^*) = 0$. ■

Пример 2.6. Найти экстремум функции $f(x) = -x_1^2 - x_2^2 - x_3^2 - x_1 + x_1 x_2 + 2x_3$ на множестве R^3.

□ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = -2x_1 - 1 + x_2 = 0, \qquad \frac{\partial f(x)}{\partial x_2} = -2x_2 + x_1 = 0, \qquad \frac{\partial f(x)}{\partial x_3} = -2x_3 + 2 = 0.$$

В результате решения системы получим стационарную точку $x^* = \left(-\frac{2}{3}, -\frac{1}{3}, 1\right)^T$.

2. Проверим выполнение достаточных условий экстремума.

Первый способ. Матрица Гессе имеет вид $H\left(x^*\right) = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$. Так как $\Delta_1 = -2 < 0$, $\Delta_2 = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = 4 - 1 = 3 > 0$, $\Delta_3 = (-2) \cdot 3 = -6 < 0$, т.е. знаки угловых миноров чередуются, начиная с отрицательного, то точка x^* – точка локального максимума (строка 2 в табл. 2.1).

Второй способ. Найдем собственные значения матрицы Гессе, используя (2.10):

$$\det\left(H - \lambda E\right) = \begin{vmatrix} -2-\lambda & 1 & 0 \\ 1 & -2-\lambda & 0 \\ 0 & 0 & -2-\lambda \end{vmatrix} = 0.$$

Отсюда $\left(-2-\lambda\right)[(-2-\lambda)^2 - 1] = 0$ и $\lambda_1 = -2 < 0, \lambda_2 = -1 < 0, \lambda_3 = -3 < 0$. Так как все собственные значения матрицы Гессе отрицательны, то в точке x^* – локальный максимум (строка 2 в табл. 2.1).

3. Вычислим значение функции в точке локального максимума: $f(x^*) = \frac{4}{3}$.■

Пример 2.7. Найти экстремум функции $f(x) = x_1^3 + x_2^2 + x_3^2 + x_2 x_3 - 3x_1 + 6x_2 + 2$ на множестве R^3.

□ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = 3x_1^2 - 3 = 0, \qquad \frac{\partial f(x)}{\partial x_2} = 2x_2 + x_3 + 6 = 0, \qquad \frac{\partial f(x)}{\partial x_3} = 2x_3 + x_2 = 0.$$

В результате решения системы получим две стационарные точки:

$$x^{1*} = (1,-4,2)^T, \quad x^{2*} = (-1,-4,2)^T.$$

2. Проверим выполнение достаточных условий в каждой стационарной точке двумя способами. Составим матрицу Гессе $H(x) = \begin{pmatrix} 6x_1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

Исследуем точку $x^{1*} = (1,-4,2)^T$.

Первый способ. Матрица Гессе имеет вид $H(x^{1*}) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$. Так как $\Delta_1 = 6 > 0$, $\Delta_2 = \begin{vmatrix} 6 & 0 \\ 0 & 2 \end{vmatrix} = 12 > 0$, $\Delta_3 = 18 > 0$, то точка x^{1*} является точкой локального минимума (строка 1 в табл. 2.1).

Второй способ. Найдем собственные значения матрицы Гессе, используя (2.10):

$$\begin{vmatrix} 6-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix} = (6-\lambda)\left[(2-\lambda)^2 - 1\right] = 0.$$

Отсюда $\lambda_1 = 6 > 0$, $\lambda_2 = 3 > 0$, $\lambda_3 = 1 > 0$ и точка x^{1*} является точкой локального минимума (строка 1 в табл. 2.1).

Исследуем точку $x^{2*} = (-1,-4,2)^T$.

Первый способ. Матрица Гессе имеет вид $H(x^{2*}) = \begin{pmatrix} -6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$. Так как $\Delta_1 = -6 < 0$, $\Delta_2 = \begin{vmatrix} -6 & 0 \\ 0 & 2 \end{vmatrix} = -12 < 0$, $\Delta_3 = -18 < 0$, то достаточные условия экстремума не выполняются (строки 1 и 2 в табл. 2.1). Согласно схеме (см. рис. 2.1) проверим необходимые условия экстремума второго порядка. Главные миноры первого порядка ($m = 1$) получаются из $\Delta_3 = \begin{vmatrix} -6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}$ в результате вычеркивания $n - m = 3 - 1 = 2$ строк и 2 столбцов с одинаковыми номерами: -6, 2, 2. Главные миноры второго порядка ($m = 2$) получаются из Δ_3 в результате вычеркивания $n - m = 3 - 2 = 1$ строк и столбцов с одинаковыми номерами: 3, –12, –12. Главный минор третьего порядка ($m = 3$) получается из Δ_3 в результате вычеркивания $n - m = 3 - 3 = 0$ строк и столбцов, т.е. совпадает с $\Delta_3 = -18$. Отсюда следует, что необходимые условия экстремума второго порядка не выполняются (строки 3 и 4 в табл. 2.1). Так как матрица Гессе не является нулевой, то можно сделать вывод о том, что в точке x^{2*} нет экстремума (строка 6 в табл. 2.1).

Второй способ. Найдем собственные значения матрицы Гессе:

$$\begin{vmatrix} -6-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix} = (-6-\lambda)\left[(2-\lambda)^2 - 1\right] = 0 .$$

Отсюда $\lambda_1 = -6 < 0,\ \lambda_2 = 3 > 0,\ \lambda_3 = 1 > 0$, т.е. собственные значения имеют разные знаки. Поэтому в точке x^{2*} нет экстремума (строка 6 в табл. 2.1).

3. Вычислим значение целевой функции в точке x^{1*} локального минимума: $f(x^{1*}) = -12$. ■

Пример 2.8. Найти экстремум функции $f(x) = -x_1^2 + 2x_1x_2 - x_2^2 - 4x_3^2$ на множестве R^3.

□ 1. Выпишем необходимые условия экстремума первого порядка:

$$\frac{\partial f(x)}{\partial x_1} = -2x_1 + 2x_2 = 0, \qquad \frac{\partial f(x)}{\partial x_2} = 2x_1 - 2x_2 = 0, \qquad \frac{\partial f(x)}{\partial x_3} = -8x_3 = 0 .$$

В результате решения системы получим бесконечное множество стационарных точек, удовлетворяющих соотношениям $x_1^* = x_2^*,\ x_3^* = 0$.

2. Проверим выполнение достаточных условий экстремума.

Первый способ. Матрица Гессе имеет вид $H(x^*) = \begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -8 \end{pmatrix}$. Так как $\Delta_1 = -2 < 0,\ \Delta_2 = \begin{vmatrix} -2 & 2 \\ 2 & -2 \end{vmatrix} = 0,\ \Delta_3 = \begin{vmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -8 \end{vmatrix} = 0$, то достаточные условия экстремума не выполняются (строки 1 и 2 в табл. 2.1). Согласно схеме (см. рис.2.1) проверим необходимые условия второго порядка. Поступим аналогично решению примера 2.7.

Главные миноры первого порядка получаются из Δ_3 в результате вычеркивания двух строк и столбцов с одинаковыми номерами: –2, –2, –8. Главные миноры второго порядка получаются из Δ_3 в результате вычеркивания по одной строке и столбцу с одинаковым номером: 16, 16, 0. Главный минор третьего порядка совпадает с $\Delta_3 = 0$. Так как все главные миноры четного порядка неотрицательны, а все миноры нечетного порядка неположительны, то можно сделать вывод о том, что в исследуемых стационарных точках может быть максимум и требуется продолжение исследования (строка 4 в табл. 2.1).

Второй способ. Найдем собственные значения матрицы Гессе:

$$\begin{vmatrix} -2-\lambda & 2 & 0 \\ 2 & -2-\lambda & 0 \\ 0 & 0 & -8-\lambda \end{vmatrix} = (-8-\lambda)\left[(-2-\lambda)^2 - 4\right] = 0 .$$

Отсюда $\lambda_1 = -8 < 0$, $\lambda_2 = 0$, $\lambda_3 = -4 < 0$, т.е. собственные значения неположительны. Поэтому в стационарных точках может быть максимум (строка 4 в табл. 2.1).

3. Функция $f(x)$ может быть записана в виде $f(x) = -\left(x_1 - x_2\right)^2 - 4x_3^2$. В каждой из найденной в п. 1 стационарной точке $f(x^*) = 0$. Исходя из структуры функции $f(x)$ можно сделать вывод о том, что для любых $x \in R^3$ справедливо: $f(x) \le f(x^*) = 0$. На основании определения 1.1 функция на множестве точек, удовлетворяющих условию $x_1^* = x_2^*, x_3^* = 0$, достигает глобального максимума. ■

Пример 2.9. Найти экстремум функции $f(x) = x^3 - 2x^2 + x + 1$ на множестве R.

□ 1. Выпишем необходимые условия экстремума первого порядка:

$$\frac{d\,f(x)}{dx} = 3x^2 - 4x + 1 = 0.$$

Решив уравнение, получим стационарные точки $x^{1*} = \frac{1}{3}$, $x^{2*} = 1$.

2. Проверим выполнение достаточных условий экстремума. Так как $n = 1$, то матрица Гессе состоит из одного элемента: $h_{11} = \frac{d^2 f(x)}{dx^2} = 6x - 4$. В точке $x^{1*} = \frac{1}{3}$ имеем $\Delta_1 = -2 < 0$, а в точке $x^{2*} = 1$: $\Delta_1 = 2 > 0$. Поэтому в точке x^{1*} – локальный максимум, а в точке x^{2*} – локальный минимум (строки 1 и 2 в табл. 2.1).

3. Вычислим значения функции в точках экстремума:

$$f(x^{1*}) = \frac{31}{27}, \qquad f(x^{2*}) = 1. \ ■$$

Пример 2.10. Найти экстремум функции $f(x) = \left(x-1\right)^6$ на множестве R.

□ 1. Запишем необходимые условия экстремума первого порядка:

$$\frac{d\,f(x)}{dx} = 6\left(x-1\right)^5 = 0.$$

Отсюда стационарная точка $x^* = 1$.

2. Проверим выполнение достаточных условий экстремума с учетом п. 3 замечаний 2.2. Первая не равная нулю производная имеет порядок $m = 6$: $f^{(6)}(x) = 6! > 0$. Так как число m – четное, то функция достигает в точке x^* локального минимума.

3. Вычислим значение функции в точке минимума: $f(x^*) = 0$. ■

Пример 2.11. Найти экстремум функции

$$f(x) = 5x^6 - 36x^5 + \frac{165}{2}x^4 - 60x^3 + 36$$

на множестве R.

□ 1. Выпишем необходимое условие экстремума первого порядка:

$$\frac{df(x)}{dx} = 30x^5 - 180x^4 + 330x^3 - 180x^2 = 30x^2(x-1)(x-2)(x-3) = 0.$$

Отсюда получим стационарные точки: $x^{1*} = 0, x^{2*} = 1, x^{3*} = 2, x^{4*} = 3$.

2. Проверим выполнение достаточных условий экстремума:

$$\frac{d^2 f(x)}{dx^2} = 150x^4 - 720x^3 + 990x^2 - 360x;$$

$$f''(x^{1*}) = 0,\ f''(x^{2*}) = 60 > 0,\ f''(x^{3*}) = -120 < 0,\ f''(x^{4*}) = 540 > 0.$$

Поэтому в точках x^{2*}, x^{4*} – локальный минимум, а в точке x^{3*} – локальный максимум (см. строки 1 и 2 в табл. 2.1 или п. 3 замечаний 2.2). В точке x^{1*} достаточные условия не выполняются (строка 5 в табл. 2.1), поэтому вычислим третью производную:

$$f^{(3)}(x^{1*}) = 600x^3 - 2160x^2 + 1980x - 360\ \big|_{x^{1*}=0} = -360.$$

Так как эта производная отлична от нуля и имеет нечетный порядок, то в точке x^{1*} нет экстремума (см. п. 3 замечаний 2.2). ■

Пример 2.12 (задача об обработке результатов экспериментов).

Пусть в результате компьютерных вычислений или экспериментальных исследований получены значения сеточной функции $f(x)$ в $(n+1)$ точках $x_0, x_1, \dots, x_n$, называемых узлами:

$$f_0 = f(x_0),\ f_1 = f(x_1), \dots,\ f_n = f(x_n).$$

Требуется наилучшим образом обработать (сгладить) результаты, т.е. заменить сеточную функцию непрерывной функцией так, чтобы можно было находить значения функции при произвольном значении аргумента на отрезке $[x_0, x_n]$.

Для решения задачи можно использовать алгебраические *сглаживающие многочлены*

$$f_m(x,a) = \sum_{j=0}^{m} a_j x^j = a_0 + a_1 x + \dots + a_m x^m, \tag{2.11}$$

где $a = (a_0, a_1, \dots, a_m)^T$ – вектор неизвестных параметров, m – степень многочлена (обычно $m \le n$). В методе наименьших квадратов [24] предлагается находить искомые коэффициенты многочлена $a_0, a_1, \dots, a_m$ из условия минимума среднеквадратической погрешности:

$$\delta_m(a) = \sqrt{\frac{1}{n+1} \sum_{i=0}^{n} \left[f_m(x_i, a) - f_i\right]^2} \to \min_{a_0, a_1, \dots a_m}. \tag{2.12}$$

Очевидно, минимум в (2.12) с учетом (2.11) достигается, если коэффициенты многочлена обеспечивают безусловный минимум (поскольку на значения коэффициентов не наложено никаких ограничений) целевой функции вида

$$\Delta = \sum_{i=0}^{n} \left[a_0 + a_1 x_i + \ldots + a_m x_i^m - f_i \right]^2 \to \min_{a_0, a_1, \ldots a_m} . \qquad (2.13)$$

Для решения задачи применим необходимые условия безусловного экстремума:

$$\frac{\partial \Delta}{\partial a_j} = 0, \quad j = 0,1,\ldots,m .$$

В результате получим систему

$$\frac{\partial \Delta}{\partial a_0} = 2\sum_{i=0}^{n} \left[a_0 + a_1 x_i + \ldots + a_m x_i^m - f_i \right] \cdot 1 = 0 ,$$

$$\frac{\partial \Delta}{\partial a_1} = 2\sum_{i=0}^{n} \left[a_0 + a_1 x_i + \ldots + a_m x_i^m - f_i \right] \cdot x_i = 0 ,$$

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

$$\frac{\partial \Delta}{\partial a_m} = 2\sum_{i=0}^{n} \left[a_0 + a_1 x_i + \ldots + a_m x_i^m - f_i \right] \cdot x_i^m = 0$$

или, поделив на 2 и сгруппировав слагаемые,

$$\underbrace{\left(\sum_{i=0}^{n} 1 \right)}_{n+1} a_0 + \left(\sum_{i=0}^{n} x_i \right) a_1 + \left(\sum_{i=0}^{n} x_i^2 \right) a_2 + \ldots + \left(\sum_{i=0}^{n} x_i^m \right) a_m = \sum_{i=0}^{n} f_i ,$$

$$\left(\sum_{i=0}^{n} x_i \right) a_0 + \left(\sum_{i=0}^{n} x_i^2 \right) a_1 + \left(\sum_{i=0}^{n} x_i^3 \right) a_2 + \ldots + \left(\sum_{i=0}^{n} x_i^{m+1} \right) a_m = \sum_{i=0}^{n} x_i f_i , \qquad (2.14)$$

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

$$\left(\sum_{i=0}^{n} x_i^m \right) a_0 + \left(\sum_{i=0}^{n} x_i^{m+1} \right) a_1 + \left(\sum_{i=0}^{n} x_i^{m+2} \right) a_2 + \ldots + \left(\sum_{i=0}^{n} x_i^{2m} \right) a_m = \sum_{i=0}^{n} x_i^m f_i .$$

Обозначим: $s_0 = n+1 , \quad t_0 = f_0 + f_1 + \ldots + f_n ,$

$$s_k = x_0^k + x_1^k + \ldots + x_n^k , \; k = 1,\ldots,2m ;$$

$$t_k = x_0^k f_0 + x_1^k f_1 + \ldots + x_n^k f_n , \; k = 1,\ldots,m .$$

Тогда система (2.14) преобразуется к виду

$$\begin{aligned} & s_0 a_0 + s_1 a_1 + \ldots + s_m a_m = t_0 , \\ & s_1 a_0 + s_2 a_1 + \ldots + s_{m+1} a_m = t_1 , \\ & \quad \vdots \\ & s_m a_0 + s_{m+1} a_1 + \ldots + s_{2m} a_m = t_m . \end{aligned} \qquad (2.15)$$

Решив систему линейных алгебраических уравнений, можно найти неизвестные коэффициенты $a_0, a_1, ..., a_m$. Подставив решение в (2.11), получим искомую формулу, которая сглаживает экспериментальные данные.

Известно, что при $m = n$ изложенным методом находится многочлен, проходящий через все заданные точки (рис. 2.2,*а*). Этот многочлен называется *интерполяционным*, при этом погрешность (2.12) равна нулю. При $0 \le m < n$ функция (2.11) в общем случае не проходит через все заданные точки (рис. 2.2,*б*). Как правило, с увеличением степени многочлена значение погрешности уменьшается.

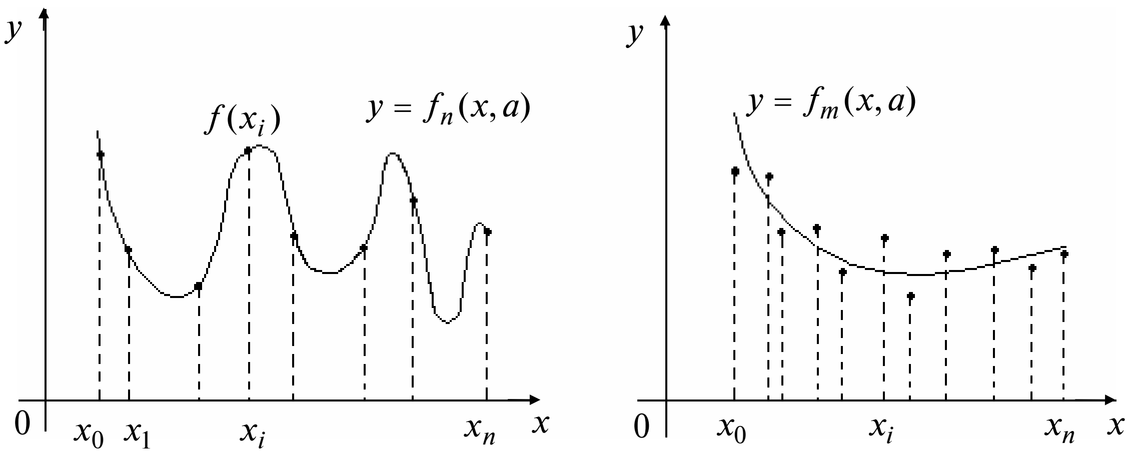


Рис. 2.2

Рассмотрим пример, когда результаты экспериментов представлены сеточной функцией, заданной табл. 2.2. Требуется найти многочлен, сглаживающий результаты эксперимента, при $m = 0, 1, 2, 3$.

Таблица 2.2

i	0	1	2	3
x_i	2	3	4	5
$f(x_i) = f_i$	7	5	8	7

□ 1. Пусть степень многочлена $m = 0$, тогда решение ищется в виде

$$f_0(x, a) = a_0 .$$

Для составления уравнения $s_0 a_0 = t_0$, следующего из (2.15), найдем коэффициенты s_0, t_0 (табл. 2.3). В результате получим $4a_0 = 27$, или $a_0 = \frac{27}{4} = 6{,}75$. Искомая сглаживающая функция имеет вид $f_0(x, a) = 6{,}75$, а среднеквадратическая погрешность $\delta_0(a) = 1{,}0897$.

Таблица 2.3

x_i	f_i	1	x_i^2	x_i^3	x_i^4	x_i^5	x_i^6	$x_i f_i$	$x_i^2 f_i$	$x_i^3 f_i$
2	7	1	4	8	16	32	64	14	28	56
3	5	1	9	27	81	243	729	15	45	135
4	8	1	16	64	256	1024	4096	32	128	512
5	7	1	25	125	625	3125	15625	35	175	875
14	27	4	54	224	978	4424	20514	96	376	1578
s_1	t_0	s_0	s_2	s_3	s_4	s_5	s_6	t_1	t_2	t_3

2. Пусть степень многочлена $m = 1$, тогда решение ищется в виде

$$f_1(x, a) = a_0 + a_1 x .$$

Составим систему (2.15), пользуясь табл. 2.3:

$$\begin{aligned} s_0 a_0 + s_1 a_1 &= t_0 , \\ s_1 a_0 + s_2 a_1 &= t_1 , \end{aligned} \quad \text{или} \quad \begin{aligned} 4a_0 + 14a_1 &= 27 , \\ 14a_0 + 54a_1 &= 96 . \end{aligned}$$

Решение системы: $a_0 = 5{,}7$; $a_1 = 0{,}3$. Искомая сглаживающая функция имеет вид $f_1(x, a) = 5{,}7 + 0{,}3x$, а среднеквадратичная погрешность $\delta_1(a) = 1{,}0368$.

3. Пусть $m = 2$, тогда решение ищется в виде

$$f_2(x, a) = a_0 + a_1 x + a_2 x^2 .$$

Составим систему (2.15), пользуясь табл. 2.3:

$$\begin{aligned} s_0 a_0 + s_1 a_1 + s_2 a_2 &= t_0 , \\ s_1 a_0 + s_2 a_1 + s_3 a_2 &= t_1 , \\ s_2 a_0 + s_3 a_1 + s_4 a_2 &= t_2 , \end{aligned} \quad \text{или} \quad \begin{aligned} 4a_0 + 14a_1 + 54a_2 &= 27 , \\ 14a_0 + 54a_1 + 224a_2 &= 96 , \\ 54a_0 + 224a_1 + 978a_2 &= 376 . \end{aligned}$$

Решение ее, полученное методом Гаусса [24], имеет вид $a_2 = \dfrac{1}{4}$, $a_1 = -\dfrac{29}{20}$, $a_0 = \dfrac{169}{20}$. Искомая сглаживающая функция $f_2(x, a) = \dfrac{169}{20} - \dfrac{29}{20} x + \dfrac{1}{4} x^2$, а среднеквадратичная погрешность $\delta_2(a) = 1{,}0062$.

4. Пусть $m = 3$, тогда решение ищется в виде

$$f_2(x, a) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 .$$

Составим систему (2.15), пользуясь табл. 2.3:

$$\begin{aligned} s_0a_0 + s_1a_1 + s_2a_2 + s_3a_3 &= t_0, \\ s_1a_0 + s_2a_1 + s_3a_2 + s_4a_3 &= t_1, \\ s_2a_0 + s_3a_1 + s_4a_2 + s_5a_3 &= t_2, \\ s_3a_0 + s_4a_1 + s_5a_2 + s_6a_3 &= t_3, \end{aligned} \quad \text{или} \quad \begin{aligned} 4a_0 + 14a_1 + 54a_2 + 224a_3 &= 27, \\ 14a_0 + 54a_1 + 224a_2 + 978a_3 &= 96, \\ 54a_0 + 224a_1 + 978a_2 + 4424a_3 &= 376, \\ 224a_0 + 978a_1 + 4424a_2 + 20514a_3 &= 1578. \end{aligned}$$

Решение ее, полученное методом Гаусса [24], имеет вид $a_3 = -\frac{3}{2}$, $a_2 = 16$, $a_1 = -\frac{107}{2}$, $a_0 = 62$, искомая функция $f_3(x,a) = 62 - \frac{107}{2}x + 16x^2 - \frac{3}{2}x^3$, а среднеквадратичная погрешность $\delta_3(a) = 0$.

На рис. 2.3 изображены заданная сеточная функция и найденные сглаживающие многочлены при $m = 0, 1, 2, 3$.

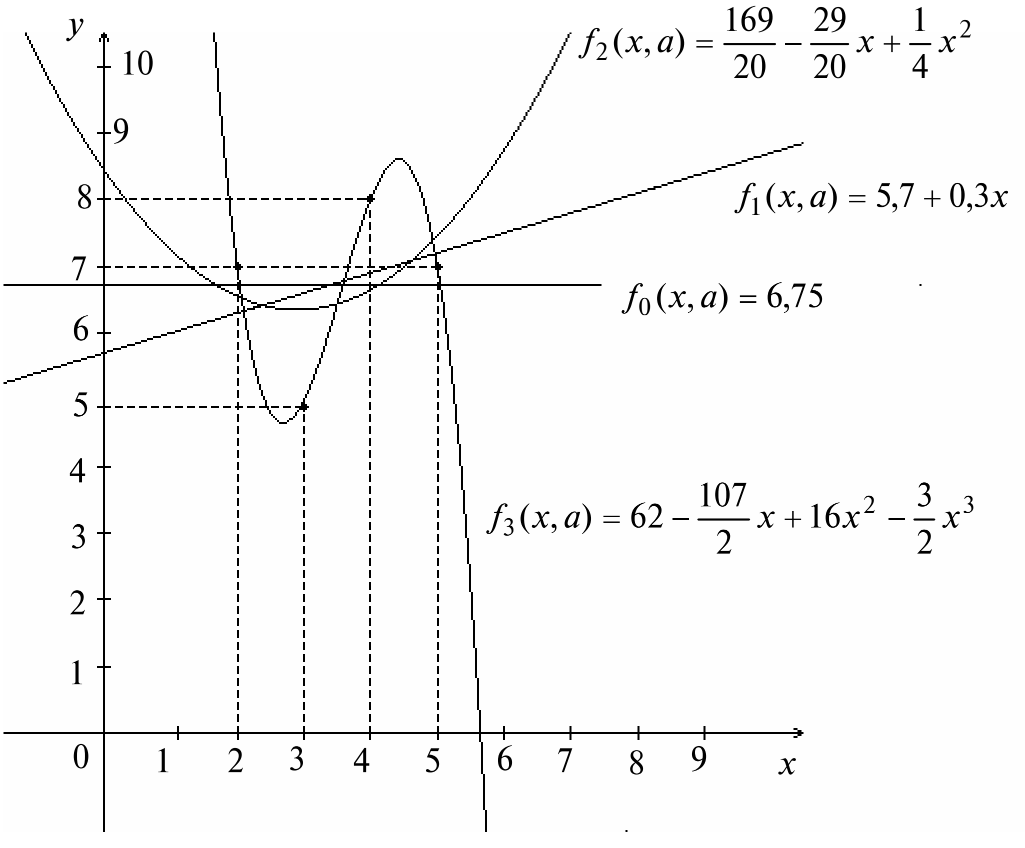


Рис. 2.3

Заметим, что график интерполяционного многочлена (при $m = 3$) проходит через все заданные точки, а величина среднеквадратической погрешности уменьшается с ростом степени m сглаживающего многочлена. ■

Задачи для самостоятельного решения

1. Найти безусловный экстремум функции $f(x) = 4x_1^2 + 3x_2^2 - 4x_1x_2 + x_1$.

Ответ: в точке $x^* = \left(-\frac{3}{16}, -\frac{1}{8}\right)^T$ – локальный и одновременно глобальный минимум.

2. Найти безусловный экстремум функции

$$f(x) = 2x_1^3 + 4x_1x_2^2 - 10x_1x_2 + x_2^2.$$

Ответ: в точке $x^* = (1, 1)^T$ – локальный минимум; в точке $x^* = (0, 0)^T$ нет экстремума, так как не выполняются необходимые условия экстремума второго порядка.

3. Найти безусловный экстремум функции

$$f(x) = x_1^3 - x_1x_2 + x_2^2 - 2x_1 + 3x_2 - 4.$$

Ответ: в точке $x^* = \left(\frac{1}{2}, -\frac{5}{4}\right)^T$ – локальный минимум; в точке $x^* = \left(-\frac{1}{3}, -\frac{5}{3}\right)^T$ нет экстремума, так как не выполняются необходимые условия экстремума второго порядка.

4. Найти безусловный экстремум функции $f(x) = (x_1 - 1)^4 + (x_2 - 3)^2$.

Ответ: в точке $x^* = (1, 3)^T$ выполняется необходимое условие экстремума второго порядка, т.е. $H(x^*) \geq 0$. Так как для любых $x \in R^2$ справедливо $f(x^*) = 0 \leq f(x)$, то по определению 1.1 точка x^* является точкой глобального минимума.

5. Найти безусловный экстремум функции $f(x) = (x_2 - x_1^2)^2 + (1 - x_1)^2$.

Ответ: в точке $x^* = (1, 1)^T$ – локальный минимум.

6. Найти безусловный экстремум функции

$$f(x) = 3x_1x_2 - x_1x_2^2 - x_1^2x_2.$$

Ответ: в точке $x^* = (1, 1)^T$ – локальный максимум; в точке $x^* = (0, 0)^T$ нет экстремума.

7. Найти безусловный экстремум функции

а) $f(x) = n^2 x_1^2 - x_1x_2 + \frac{(n+1)^2}{2} x_2^2 - n x_1$;

б) $f(x) = (n+10)\, x_1^2 + 2n x_1x_2 + (n+30) x_2^2 - 4(n+15) x_1 - 12n x_2$.