Прикладные методы оптимизации

Постановка задачи оптимизации

Рассмотрим n-мерное евклидово пространство $\mathbb{R}^n$ с евклидовой нормой $\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ для вектора (точки) $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$.

Для постановки задачи оптимизации требуется:

– *целевая функция* $f(x)$, где $x = (x_1, \ldots, x_n)^T$, определенную на пространстве $\mathbb{R}^n$. Значения этой функции характеризуют степень достижения цели, ради которой решается задача;

– *множество допустимых точек* $X \subseteq \mathbb{R}^n$, среди элементов которого осуществляется поиск.

Задача оптимизации ставится следующим образом: требуется найти такой вектор x^* из множества X допустимых решений, которому соответствует минимальное (или максимальное) значение целевой функции на этом множестве, то есть

$$f(x^*) = \min_{x \in X} f(x) \quad \left(f(x^*) = \max_{x \in X} f(x)\right) \tag{1.1}$$

Задача поиска *минимума* и *максимума* целевой функции $f(x)$ называется задачей поиска *экстремума*:

$$f(x^*) = \operatorname*{extr}_{x \in X} f(x).$$

Задача поиска максимума функции $f(x)$ сводится к задаче поиска минимума путём замены знака перед функцией на противоположный:

$$f(x^*) = \max_{x \in X} f(x) = -\min_{x \in X} f(x) \qquad (1.2)$$

Поэтому, если не оговорено противного, все выкладки будут проводиться для задачи на минимум. Соответствующие результаты для задачи на максимум тогда без труда могут быть получены из соотношений (1.2).

Если множество допустимых решений X задается ограничениями (условиями), накладываемыми на вектор x, то решается задача поиска *условного экстремума*. В этом случае X является собственным подмножеством $\mathbb{R}^n$, то есть $X \subset \mathbb{R}^n$ и $X \neq \mathbb{R}^n$. Если же $X = \mathbb{R}^n$, то есть ограничения (условия) на вектор x отсутствуют, решается задача поиска *безусловного экстремума.*

Решить задачу (1.1) означает:

– либо найти пару $(x^*, f(x^*))$, включающую точку x^* и значение целевой функции в ней;

– либо показать, что целевая функция $f(x)$ не ограничена на множестве допустимых значений, то есть $\inf\limits_{x \in X} f(x) = -\infty$ в случае задачи на минимум или $\sup\limits_{x \in X} f(x) = +\infty$ в случае задачи на максимум;

– либо показать, что множество допустимых значений пусто, то есть $X = \emptyset$.

Множество точек минимума (максимума) целевой функции $f(x)$ на множестве X обозначим X^*. Оно может содержать конечное число точек (в частности, 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, то есть

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

Определение 1.2. Точка $x^* \in X$ называется точкой *локального (относительного) минимума* функции $f(x)$ на множестве X, если существует $\varepsilon > 0$, такое, что если $x \in X$ и $\|x - x^*\| < \varepsilon$, то $f(x^*) \leq f(x)$.

Отличие глобального минимума от локального состоит в том, что в случае глобального минимума точка x^* сравнивается со всеми точками из множества допустимых точек X, а в случае локального минимума – только с принадлежащими ε-окрестности.

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

Отметим, что глобальный экстремум всегда является одновременно локальным. Обратное неверно.

Определение 1.3. *Поверхностью уровня* S_α, $\alpha \in \mathbb{R}$, функции $f(x)$ называется множество точек $x \in \mathbb{R}^n$, в которых функция принимает постоянное значение α, то есть $f(x) = \alpha$. Если $n = 2$, поверхность уровня называется *линией уровня* на плоскости $\mathbb{R}^2$.

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

$$f'(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)^T.$$

Важное свойство градиента состоит в том, что градиент функции в точке x направлен по нормали к поверхности уровня, проходящей через точку x, то есть перпендикулярно к касательной плоскости, проведенной в точке x, в сторону наибольшего возрастания функции в данной точке.

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

Используя понятие поверхности уровня целевой функции и градиента (антиградиента) этой функции, можно дать геометрическую интерпретацию решения задачи (1.1), если задача поставлена в пространстве $\mathbb{R}^2$. Эта интерпретация для задачи минимума состоит в следующем. Строим множество допустимых точек X. Через некоторую точку $x \in X$ проводим линию уровня S_α. Антиградиент $-f'(x)$ в этой точке x показывает направление наибольшего убывания функции $f(x)$ в данной точке. Сдвигая точку $x \in X$ в сторону антиградиента, находим крайнее положение линии уровня S_{α^*}, если оно существует, такое, что $S_{\alpha^*} \cap X \neq \emptyset$, а при $\alpha > \alpha^*$ будет $S_\alpha \cap X = \emptyset$. Тогда множество $X^* = \{x \in X \mid f(x) = \alpha^*\}$ и будет множеством решений задачи (1.1). Если же такого положения линии уровня не существует, то целевая функция не ограничена на множестве допустимых точек, то есть $\inf\limits_{x \in X} f(x) = -\infty$.

Пример 1.1. Найти точки экстремума функции $f(x) = x_1^2 + x_2^2$ на множестве $X = \{x \mid x_2^2 - x_1 + 3 = 0\}$.

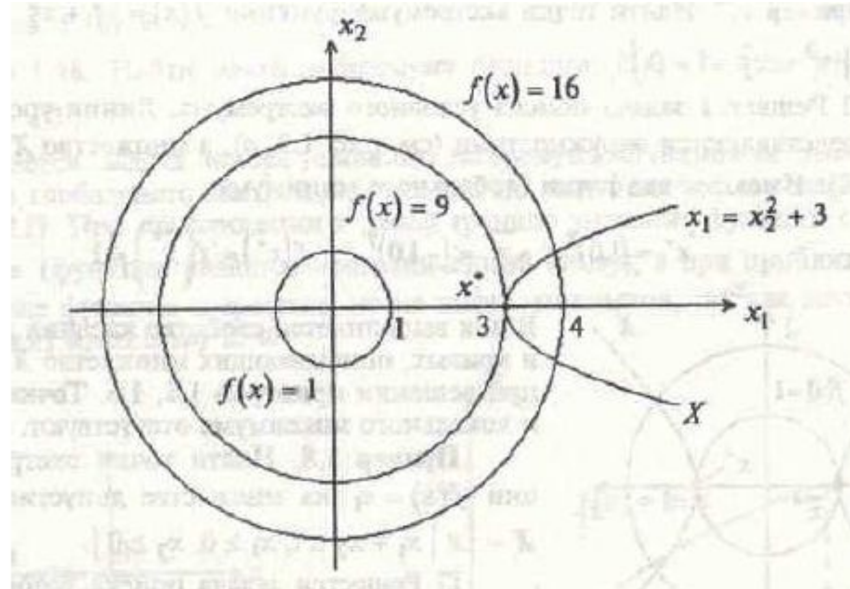

Рис. 1.1:

Решение. Решается задача поиска условного экстремума. Линии уровня S_α функции $f(x)$ представляют из себя окружности $x_1^2 + x_2^2 = \alpha$ радиуса $R = \sqrt{\alpha}$, а множество допустимых точек X есть парабола с уравнением $x_1 = x_2^2 + 3$. Градиент целевой функции $f'(x) = (2x_1, 2x_2)^T$, например, в точке $x^1 = (4;0)$ равен $f'(x^1) = (8;0)^T$, антиградиент есть $-f'(x^1) = (-8;0)^T$. Таким образом, в точке $x^* = (3;0)^T$ достигается глобальный минимум задачи и $f(x^*) = 9$ (рис. 1.1). Точка x^* является точкой касания линии уровня и множества допустимых точек X, то есть параболы $x_1 = x_2^2 + 3$. Глобальный максимум на данном множестве не достигается, так как, какое бы большое α ни взять, всегда будет $S_\alpha \cap X \neq \emptyset$. $\square$

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

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

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

С помощью градиента и матрицы Гессе, используя разложение в ряд Тейлора, приращение функции $f(x)$ в точке x можно записать в форме

$$\Delta f(x) = f(x + \Delta x) - f(x) = f'(x)^T \Delta x + \frac{1}{2} \Delta x^T H(x) \Delta x + o(\|\Delta x\|^2),$$

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

Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

есть симметричная $n \times n$ матрица, то есть матрица, у которой $a_{ij} = a_{ji}$, $i, j = \overline{1, n}$.

Квадратичная форма $x^T A x$ (а также соответствующая $n \times n$ симметричная матрица A) называется:

– *положительно определенной* ($A > 0$), если для любого ненулевого x выполняется неравенство $x^T A x > 0$;

– *отрицательно определенной* ($A < 0$), если для любого ненулевого x выполняется неравенство $x^T A x < 0$;

– *положительно полуопределенной* ($A \geq 0$), если для любого x выполняется неравенство $x^T A x \geq 0$ и существует вектор $x \neq 0$, для которого $x^T A x = 0$;

– *отрицательно полуопределенной* ($A \leq 0$), если для любого x выполняется неравенство $x^T A x \leq 0$ и существует вектор $x \neq 0$, для которого $x^T A x = 0$;

– *неопределенной* ($A \gtrless 0$), если существуют такие векторы $\tilde{x}$ и $\tilde{\tilde{x}}$, что выполняются неравенства $\tilde{x}^T A \tilde{x} > 0$ и $\tilde{\tilde{x}}^T A \tilde{\tilde{x}} < 0$;

– *тождественно равной нулю* ($A \equiv 0$), если для любого x будет $x^T A x = 0$.

Определение 1.6. Для симметричной матрицы A определители

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ldots, \Delta_n = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

называются *угловыми минорами*. Определители m-ого порядка $(m \leq n)$, получающиеся из определителя матрицы A вычеркиванием каких-либо $(n-m)$ строк и $(n-m)$ столбцов с одними и теми же номерами, называются *главными минорами*.

Теорема 1.1 (критерий Сильвестра). *Для симметричной матрицы A справедливы утверждения:*

1. *Для того чтобы матрица A была положительно определенной, необходимо и достаточно, чтобы знаки угловых миноров были строго положительны:*

$$\Delta_1 > 0,\ \Delta_2 > 0, \ldots, \Delta_n > 0;$$

2. *Для того чтобы матрица A была отрицательно определенной, необходимо и достаточно, чтобы знаки угловых миноров чередовались, начиная с отрицательного:*

$$\Delta_1 < 0,\ \Delta_2 > 0, \ldots, (-1)^n \Delta_n > 0;$$

3. *Для того чтобы матрица A была положительно полуопределенной, необходимо и достаточно, чтобы все главные миноры матрицы A были неотрицательны;*

4. *Для того чтобы матрица A была отрицательно полуопределенной, необходимо и достаточно, чтобы все главные миноры матрицы A четного порядка были неотрицательны, а все главные миноры нечетного порядка – неположительны.*

Пример 1.2. Исследовать матрицы:

$$A = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} -2 & -2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -8 \end{pmatrix}$$

на знакоопределенность.

Решение. Рассмотрим матрицу A. Так как

$$\Delta_1 = -2 < 0, \quad \Delta_2 = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = 4 - 1 = 3 > 0,$$

$$\Delta_3 = (-2)\begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = (-2) \cdot 3 = -6 < 0,$$

то есть знаки угловых миноров чередуются, начиная с отрицательного, то матрица A отрицательно определенная.

Рассмотрим теперь матрицу B. Так как

$$\Delta_1 = -2 < 0, \quad \Delta_2 = \begin{vmatrix} -2 & 2 \\ 2 & 2 \end{vmatrix} = 0, \quad \Delta_3 = \begin{vmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -8 \end{vmatrix} = 0,$$

то эта матрица не является ни отрицательно, ни положительно определенной. Проверим знаки главных миноров. Главные миноры первого порядка получаются из Δ_3 в результате вычеркивания двух строк и столбцов с одинаковыми номерами. Имеем три главных минора первого порядка: $-2, -2, -8$. Главные миноры второго порядка получаются из Δ_3 в результате вычеркивания по одной строке и столбцу с одинаковым номером. Имеем три главных минора второго порядка: $16, 16, 0$. Главный минор третьего порядка совпадает с $\Delta_3 = 0$. Так как все главные миноры четного порядка неотрицательны, а все миноры нечетного порядка неположительны, то на основании п.4 теоремы 1.1 можно сделать заключение о том, что матрица B является отрицательно полуопределенной. $\square$

Элементы выпуклого анализа

Выпуклый анализ – раздел математики, в котором изучаются выпуклые объекты: множества, функции и экстремальные задачи.

Определение 1.7. Множество $X \subseteq \mathbb{R}^n$ называется *выпуклым*, если наряду с любыми двумя точками $x^1, x^2 \in X$ это множество содержит отрезок с концами x^1 и x^2, то есть, для любых точек $x^1, x^2 \in X$ и для любого числа $\lambda \in [0, 1]$ справедливо включение $\lambda x^1 + (1 - \lambda)x^2 \in X$.

Для отрезка, соединяющего точки x^1 и x^2, используется обозначение $[x^1, x^2]$, то есть

$$[x^1, x^2] = \{x \mid x = \lambda x^1 + (1 - \lambda)x^2,\ \lambda \in [0, 1]\}.$$

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

Определение 1.8. Функция $f(x)$, определенная на выпуклом множестве X, называется *выпуклой*, если для любых точек $x^1, x^2 \in X$ и для любого $\lambda \in [0, 1]$ выполняется неравенство

$$f(\lambda x^1 + (1 - \lambda)x^2) \leq \lambda f(x^1) + (1 - \lambda) f(x^2).$$

Определение 1.9. Функция $f(x)$, определенная на выпуклом множестве X, называется *строго выпуклой*, если для любых точек $x^1, x^2 \in X$ таких, что $x^1 \neq x^2$, и для любого $\lambda \in (0,1)$ выполняется неравенство

$$f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda) f(x^2).$$

Определение 1.10. Функция $f(x)$, определенная на выпуклом множестве X, называется *сильно выпуклой* с константой $l > 0$, если для любых точек $x^1, x^2 \in X$ таких, что $x^1 \neq x^2$, и для любого $\lambda \in (0,1)$ выполняется неравенство

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

Другими словами, функцию $f(x)$ называют выпуклой, если она целиком лежит не выше отрезка, соединяющего две произвольные точки ее графика. Функцию называют строго выпуклой, если она целиком лежит ниже отрезка, соединяющего две произвольные, но не совпадающие точки ее графика.

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

Выпуклость функции можно определить, исследуя ее матрицу Гессе.

Теорема 1.2. *Пусть $f(x)$ есть дважды непрерывно-дифференцируемая функция и $H(x)$ – ее матрица Гессе. Тогда*

а) если $H(x) \geq 0 \ \forall x \in \mathbb{R}^n$, то функция $f(x)$ выпуклая на $\mathbb{R}^n$;

б) если $H(x) > 0 \ \forall x \in \mathbb{R}^n$, то функция $f(x)$ строго выпуклая на $\mathbb{R}^n$;

в) если $H(x) \geq lE \ \forall x \in \mathbb{R}^n$, где $l > 0$ есть некоторая константа, а E – единичная матрица, то функция $f(x)$ сильно выпуклая на $\mathbb{R}^n$ с константой l.

Выпуклые функции являются удобным средством задания выпуклых множеств.

Теорема 1.3. *Пусть $P \subseteq \mathbb{R}^n$ есть выпуклое множество, функции $g_j(x) = a_j^T x - b_j$, $j = \overline{1,k}$ – линейные функции, где $a_j \in \mathbb{R}^n$ – заданные векторы, а $b_j \in \mathbb{R}$ – заданные числа, функции $g_j(x)$, $j = \overline{k+1,m}$ – выпуклые на P. Тогда выпукло множество*

$$X = \{x \in P \mid g_j(x) \leq 0, \quad j = \overline{1,k}; \quad g_j(x) = 0, \quad j = \overline{k+1,m}\}.$$

При решении экстремальных задач используются следующие свойства выпуклых функций.

Теорема 1.4. *Если $f(x)$ – выпуклая функция на выпуклом множестве $X \subseteq \mathbb{R}^n$, то всякая точка локального минимума является точкой ее глобального минимума на X.*

Теорема 1.5. *Если вогнутая на выпуклом множестве X функция $f(x)$ достигает своего минимума в двух различных точках, то она достигает минимума во всех точках отрезка, соединяющего эти две точки.*

Теорема 1.6. *Если $f(x)$ – строго выпуклая функция на выпуклом множестве $X \subseteq \mathbb{R}^n$, то она может достигать своего глобального минимума на X не более, чем в одной точке.*

Безусловный экстремум

Дана функция $f(x)$, определенная и непрерывная на множестве $X = \mathbb{R}^n$. Требуется исследовать функцию $f(x)$ на экстремум, то есть определить точки $x^* \in \mathbb{R}^n$ ее локальных максимумов и максимумов на $\mathbb{R}^n$:

$$f(x^*) = \min_{x \in \mathbb{R}^n} f(x), \quad f(x^*) = \max_{x \in \mathbb{R}^n} f(x). \tag{1.3}$$

Точки x^* находятся с помощью необходимых условий локального экстремума первого и второго порядков (порядок условий определяется порядком используемых производных), а также достаточных условий безусловного локального экстремума.

Теорема 1.7 (необходимые условия экстремума первого порядка). *Пусть $x^* \in \mathbb{R}^n$ есть точка локального минимума (максимума) функции $f(x)$ на множестве $\mathbb{R}^n$ и $f(x)$ дифференцируема в точке x^*. Тогда градиент функции $f(x)$ в точке x^* равен нулю, то есть*

$$f'(x^*) = 0 \tag{1.4}$$

или, что то же самое,

$$\frac{\partial f(x^*)}{\partial x_i} = 0, \quad i = \overline{1, n}.$$

Определение 1.11. Точки x^*, удовлетворяющие условию (1.4), называются *стационарными.*

Теорема 1.8 (необходимые и достаточные условия минимума первого порядка для выпуклой функции). *Пусть $f(x)$ есть выпуклая непрерывно дифференцируемая функция на $\mathbb{R}^n$. Для того чтобы точка x^* была точкой глобального минимума для функции $f(x)$, необходимо и достаточно, чтобы градиент функции $f(x)$ в точке x^* равнялся нулю, то есть*

$$f'(x^*) = 0.$$

Теорема 1.9 (необходимые условия экстремума второго порядка). *Пусть точка x^* есть точка локального минимума (максимума) функции $f(x)$ на множестве $\mathbb{R}^n$ 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), то есть*

$$H(x^*) \geq 0 \quad (H(x^*) \leq 0). \tag{1.5}$$

Теорема 1.10 (достаточные условия экстремума). *Пусть функция $f(x)$ в точке $x^* \in \mathbb{R}^n$ дважды дифференцируема, ее градиент в точке x^* равен нулю, а матрица Гессе является положительно определенной (отрицательно определенной) в точке x^*, то есть*

$$f'(x^*) = 0 \text{ и } H(x^*) > 0 \quad (H(x^*)) < 0. \tag{1.6}$$

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

Алгоритм решения задачи (1.3):

1. Записать необходимые условия экстремума первого порядка в форме (1.4) и найти стационарные точки x^* в результате решения системы n в общем случае нелинейных алгебраических уравнений с n неизвестными.
2. В найденных стационарных точках x^* проверить выполнение достаточных, а если они не выполняются, то необходимых условий второго порядка.
3. Вычислить значения $f(x^*)$ в точках экстремума.

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

Для функции $f(x)$ одной переменной $(n = 1)$ можно сформулировать критерий безусловного экстремума.

Теорема 1.11 (критерий безусловного экстремума для функции одной переменной). . *Если функция $f(x)$ и ее производные непрерывны, то точка x^* является точкой экстремума тогда и только тогда, когда*

$$f'(x^*) = f''(x^*) = \cdots = f^{(m-1)}(x^*) = 0,\ f^{(m)}(x^*) \neq 0$$

и m – четно. При этом 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, то в точке x^* нет экстремума.

Пример 1.3. 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 $\mathbb{R}^3$.

Решение. В соответствии с алгоритмом решения задачи будем иметь:

1. Запишем необходимые условия экстремума первого порядка: $\frac{\partial f(x)}{\partial x_1} = 3x_1^2 - 3 = 0$, $\frac{\partial f(x)}{\partial x_2} = 2x_2 + x_3 + 6 = 0$, $\frac{\partial f(x)}{\partial x_3} = 2x_3 + x_2 = 0$. В результате решения системы получаем две стационарные точки $x^1 = (1, -4, 2)^T$, $x^2 = (-1, -4, 2)^T$.

2. Проверим выполнение достаточных и необходимых условий второго порядка в каждой стационарной точке.

   Исследуем точку $x^1 = (1, -4, 2)^T$. Матрица Гессе имеет вид:

$$H(x^1) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

Так как

$$\Delta_1 = 6 > 0, \quad \Delta_2 = \begin{vmatrix} 6 & 0 \\ 0 & 6 \end{vmatrix} = 12 > 0, \quad \Delta_3 = 18 > 0,$$

точка x^1 является точкой локального минимума.

Исследуем точку $x^2 = (-1, -4, 2)^T$. 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. Проверим необходимые условия экстремума второго порядка. Главные миноры матрицы Гессе первого порядка есть: -6, 2, 2. Главные миноры второго порядка: 3, -12, -12. Главный минор третьего порядка совпадает с $\Delta_3 = -18 < 0$. Отсюда следует, что необходимые условия экстремума второго порядка не выполняются. Следовательно, в точке x^2 нет экстремума.

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

Условный экстремум. Выпуклое программирование

1.4.1 Постановка задачи и основные определения

Мы будем рассматривать случай, когда множество допустимых точек X задается равенствами и неравенствами, то есть будем решать задачу

$$f(x^*) = \min_{x \in X} f(x) \quad \left(f(x^*) = \max_{x \in X} f(x) \right), \tag{1.7}$$

где

$$X = \{x \mid g_j(x) = 0,\ j = \overline{1,k},\ k < n;\ g_j(x) \le 0,\ j = \overline{k+1,m}\},$$

k и m – произвольные натуральные числа, причем k может принимать значение и 0.

Будем считать функции $f(x)$; $g_j(x)$, $j = \overline{1,m}$ дважды непрерывно дифференцируемыми на множестве $\mathbb{R}^n$. Задача (1.7) называется задачей на условный экстремум со *смешанными ограничениями*. При $m = k$ задача (1.7) преобразуется в задачу с *ограничениями типа равенств*, а при $k = 0$ в задачу с *ограничениями типа неравенств*.

Определение 1.12. Функция

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) \tag{1.8}$$

называется *функцией Лагранжа*, числа λ_0, $\lambda_1, \ldots, \lambda_m$ – *множителями Лагранжа*, $\lambda = (\lambda_1, \ldots, \lambda_m)^T$. *Классической функцией Лагранжа* $L(x, \lambda)$ называется функция $L(x, 1, \lambda)$, то есть

$$L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x). \tag{1.9}$$

Определение 1.13. *Градиентом функции Лагранжа по* x называется вектор-столбец, составленный из ее частных производных первого порядка по x_i, $i = \overline{1, n}$:

$$L'_x(x, \lambda_0, \lambda) = \left(\frac{\partial L(x, \lambda_0, \lambda)}{\partial x_1}, \ldots, \frac{\partial L(x, \lambda_0, \lambda)}{\partial x_n} \right)^T. \tag{1.10}$$

Определение 1.14. *Вторым дифференциалом функции Лагранжа по x* называется функция

$$d_x^2 L(x, \lambda_0, \lambda) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L(x, \lambda_0, \lambda)}{\partial x_i \partial x_j} dx_i dx_j. \qquad (1.11)$$

Таким образом, второй дифференциал функции Лагранжа является квадратичной формой с матрицей Гессе по x

$$L''_{xx}(x, \lambda, \lambda_0) = \begin{pmatrix} \frac{\partial^2 L(x_1,\lambda_0,\lambda)}{\partial x_1 \partial x_1} & \frac{\partial^2 L(x_1,\lambda_0,\lambda)}{\partial x_1 \partial x_n} \\ \dots\dots\dots\dots & \dots\dots\dots\dots \\ \frac{\partial^2 L(x_1,\lambda_0,\lambda)}{\partial x_n \partial x_1} & \frac{\partial^2 L(x_1,\lambda_0,\lambda)}{\partial x_n \partial x_n} \end{pmatrix}.$$

Определение 1.15. Ограничение $g_j(x) \le 0$ называется *активным* в точке x^*, если $g_j(x^*) = 0$. Если $g_j(x^*) < 0$, то ограничение называется *пассивным*.

Определение 1.16. Градиенты ограничений $g_1(x), \dots, g_m(x)$ называются *линейно независимыми* в точке x^*, если равенство

$$\lambda_1 q'_1(x^*) + \lambda_2 q'_2(x^*) + \dots + \lambda_m q'_m(x^*) = 0$$

выполняется только при $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$. Если существуют числа $\lambda_1, \lambda_2, \dots, \lambda_m$, одновременно не равные нулю, для которых равенство выполняется, то градиенты *линейно зависимы*.

Выпуклым программированием называется раздел математики, в котором исследуются задачи минимизации

$$f(x^*) = \min_{x \in X} f(x) \tag{1.12}$$

выпуклых функций $f(x)$ на выпуклых множествах $X \subseteq \mathbb{R}^n$.

Если множество X задается в виде

$$X = \{x \in P \mid g_j(x) \leq 0,\ j = \overline{1,k}\}, \tag{1.13}$$

где множество $P \subseteq \mathbb{R}^n$ выпуклое, а функции $g_j(x)$, $j = \overline{1,k}$ – выпуклые функции на P, то задачу (1.12)–(1.13) будем называть *основной задачей выпуклого программирования*, или, сокращенно, ОЗВП.

Как следует из теоремы 1.3, множество X, заданное в виде (1.13) с указанными условиями, будет выпуклым.

Будем называть задачу ОЗВП *гладкой*, если функции $f(x)$, $g_1(x)$, $\ldots$, $g_m(x)$ дважды непрерывно дифференцируемы на P.

Центральное место в теории выпуклого программирования занимает теорема Куна-Таккера. Для того чтобы сформулировать эту теорему, приведем некоторые определения и факты.

Для ОЗВП (1.12)–(1.13) составим классическую функцию Лагранжа

$$L(x,\lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

и рассмотрим ее при $x \in P$, $\lambda \geq 0$, $\lambda \in \mathbb{R}^m$.

Определение 1.17. Говорят, что точка (x^*, λ^*), $\lambda^* \geq 0$, $x^* \in P$ есть *седловая точка* функции Лагранжа $L(x,\lambda)$, если для всех $x \in P$ и $\lambda \geq 0$ выполняются неравенства

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*).$$

Определение 1.18. Говорят, что ОЗВП (1.12)–(1.13) удовлетворяет *условию Слейтера*, если существует такая точка $x^* \in X$, для которой выполняются неравенства

$$g_j(x^*) < 0, \quad j = \overline{1,k}.$$

Справедливо следующее достаточное условие глобального минимума, которое имеет место для ОЗВП в самой общей постановке.

Теорема 1.12. *Если (x^*, λ^*) есть седловая точка функции Лагранжа $L(x, \lambda)$, то x^* – глобальный минимум ОЗВП (1.12)–(1.13) и выполняются условия дополняющей нежесткости*

$$\lambda_j^* g_j(x^*) = 0, \quad j = \overline{1, m}. \tag{1.14}$$

Таким образом, согласно этой теореме для того, чтобы найти глобальный минимум ОЗВП (1.12)–(1.13), достаточно найти седловую точку функции Лагранжа. Однако не для каждой ОЗВП функция Лагранжа имеет седловую точку. В этом случае для поиска глобального минимума нужно привлекать другие условия.

Теорема 1.13 (Куна—Таккера). *Для существования глобального минимума x^* ОЗВП, удовлетворяющей условию регулярности Слейтера, необходимо и достаточно существование такого неотрицательного вектора $\lambda^* \geq 0$, что пара (x^*, λ^*) является седловой точкой функции Лагранжа $L(x, \lambda)$. При этом выполняются условия дополняющей нежесткости* (1.14).

В случае гладкой ОЗВП и специального вида выпуклого множества P можно сформулировать теорему Куна–Таккера не в терминах седловой точки, а в так называемой дифференциальной форме, которая более удобна при решении конкретных гладких задач. Соответствующие формулировки будут даны в дальнейшем.

1.4.2 Условный экстремум при ограничениях типа равенств

Пусть целевая функция $f(x)$ дважды, а функции ограничений $g_j(x)$, $j = \overline{1, k}$, определяющие множество допустимых точек X, один раз непрерывно дифференцируемы на некотором множестве, содержащем X.

Требуется исследовать функцию $f(x)$ на экстремум, то есть определить точки x^* ее локальных минимумов и максимумов на множестве X:

$$f(x^*) = \min_{x \in X} f(x) \quad \left(f(x^*) = \max_{x \in X} f(x) \right), \tag{1.15}$$

где

$$X = \{X \mid g_j(x) = 0, \quad j = \overline{1, k};\; k < n.\}$$

Точки x^* находятся с помощью необходимых и достаточных условий условного минимума и максимума первого и второго порядка при ограничениях типа равенств.

Теорема 1.14 (необходимые условия экстремума первого порядка). *Пусть x^* есть точка локального экстремума в задаче* (1.15). *Тогда найдутся числа $\lambda_0^*, \lambda_1^*, \ldots, \lambda_k^*$, не равные одновременно нулю, и такие, что выполняются следующие условия:*

– *условие стационарности функции Лагранжа по x:*

$$L'_x(x^*, \lambda_0^*, \lambda^*) = 0 \Longleftrightarrow \frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = \overline{1, n}; \tag{1.16}$$

– *условие допустимости решения:*

$$g_j(x^*) = 0, \quad j = \overline{1, k}. \tag{1.17}$$

Если при этом градиенты $g'_1(x^*), \ldots, g'_k(x^*)$ в точке x^* линейно независимы (выполняется условие регулярности), то $\lambda_0^* \neq 0$.

Система (1.16)–(1.17) содержит $n + k$ уравнений с $n + k + 1$ неизвестными $\lambda_0^*, \lambda_1^*, \ldots, \lambda_k^*, x_1^*, \ldots, x_n^*$. Точки $x^* = (x_1^*, \ldots, x_n^*)^T$, удовлетворяющие системе (1.16)–(1.17) при некоторых λ_0^*, λ^*, называются *условно стационарными*.

Проверка условия регулярности затруднена, так как точка x^* заранее не известна. Поэтому, как правило, рассматриваются два случая: $\lambda_0^* = 0$ и $\lambda_0^* \neq 0$. Если $\lambda_0^* \neq 0$, в (1.16) полагают $\lambda_0^* = 1$. Это эквивалентно делению соотношения (1.16) на λ_0^* и замене $\dfrac{\lambda_j^*}{\lambda_0^*}$ на λ_j^*. При этом функция Лагранжа становится классической. В этом случае в системе (1.16)–(1.17) число неизвестных становится равным числу уравнений, и эта система имеет вид:

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} = 0, \quad i = \overline{1, n}; \tag{1.18}$$

$$g_j(x^*) = 0, \quad j = \overline{1, k}. \tag{1.19}$$

Точка экстремума, удовлетворяющая системе (1.16)–(1.17) при $\lambda_0^* \neq 0$, называется *регулярной*, а при $\lambda_0^* = 0$ – нерегулярной. Случай $\lambda_0^* = 0$ отражает вырожденность ограничений. При этом в функции Лагранжа исчезает член, содержащий целевую функцию, а в необходимых условиях экстремума не используется информация, представляемая градиентом целевой функции.

Теорема 1.15 (необходимые условия условного экстремума второго порядка). *Пусть x^* – регулярная точка минимума (максимума) в задаче* (1.15) *и имеется решение x^*, λ^* системы* (1.18)–(1.19). *Тогда второй дифференциал классической функции Лагранжа, вычисленный в точке (x^*, λ^*), неотрицателен (неположителен)*

$$d_x^2 L(x^*, \lambda^*) \geq 0 \quad (d_x^2 L(x^*, \lambda^*) \leq 0) \tag{1.20}$$

для всех $dx = (dx_1, \ldots, dx_n)^T \in \mathbb{R}^n$ таких, что

$$dg_j(x^*) = \sum_{i=1}^{n} \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j = \overline{1,k}. \tag{1.21}$$

Теорема 1.16 (достаточные условия условного экстремума). *Пусть имеется точка (x^*, λ^*), удовлетворяющая системе* (1.18)–(1.19). *Если в этой точке $d_x^2 L(x^*, \lambda^*) > 0$ $(d_x^2 L(x^*, \lambda^*) < 0)$ для всех ненулевых значений $dx = (dx_1, \ldots, dx_n)^T \in \mathbb{R}^n$ таких, что выполняются соотношения* (1.21), *то точка x^* является точкой локального минимума (максимума) в задаче* (1.15).

Отметим, что достаточные и необходимые условия экстремума второго порядка проверяются в условно-стационарных точках, которые удовлетворяют системе (1.16)–(1.17) при $\lambda_0^* \neq 0$ или системе (1.18)–(1.19), так как для практики безусловно представляет интерес случай, когда в функции Лагранжа присутствует целевая функция, экстремум которой ищется.

Алгоритм решения задачи (1.15):

1. Составить функцию Лагранжа (1.8). Записать необходимые условия условного экстремума (1.16)–(1.17). Решить систему (1.16)–(1.17) для случаев а) $\lambda_0^* = 0$; б) $\lambda_0^* = 1$. В результате найти условно-стационарные точки x^*, выделив из них регулярные точки.

2. Для выделенных в предыдущем пункте точек проверить достаточные условия экстремума:

   а) записать выражение (1.11) для второго дифференциала по x классической функции Лагранжа в точке (x^*, λ^*);

   б) записать систему (1.21) в точке x^*;

   в) из предыдущей системы выразить любые k дифференциалов dx_i через остальные $(n-k)$ дифференциалов (будем называть их *независимыми*) и подставить в $d_x^2 L(x^*, \lambda^*)$. Получим $d_x^2 L(x^*, \lambda^*) = \tilde{dx}^T \Lambda \tilde{dx}$, где Λ – симметричная $(n-k) \times (n-k)$ матрица, а $\tilde{dx}$ – $(n-k)$-мерный вектор, содержащий дифференциалы независимых переменных x_j;

г) если $\Lambda > 0$, то в точке x^* – условный локальный минимум. Если $\Lambda < 0$, то в точке x^* – условный локальный максимум.

Если достаточные условия экстремума не выполняются, следует проверить выполнение необходимых условий второго порядка, следуя аналогичной процедуре. Если они выполняются, то требуется дополнительное исследование, а если они не выполняются, то в точке x^* нет условного экстремума.

3. Вычислить значение целевой функции в точках условного экстремума.

Проверку знакоопределенности построенной в п. 2 в) матрицы Λ можно осуществить по критерию Сильвестра.

Пример 1.4. . Найти условный экстремум в задаче

$$f(x) = x_1^2 + x_2^2 \to \text{extr}$$
$$g_1(x) = x_1 + x_2 - 2 = 0.$$

Решение. Проверим условие регулярности. Так как $g_1'(x) = (1,1)^T \neq 0$, то условие регулярности выполняется (отличный от нулевого вектор всегда образует линейно независимую систему векторов). Поэтому будем пользоваться классической функцией Лагранжа.

1. Составим функцию Лагранжа

$$L(x, \lambda_1) = x_1^2 + x_2^2 + \lambda_1(x_1 + x_2 - 2).$$

Выпишем необходимые условия экстремума первого порядка:

$$\begin{cases} \dfrac{\partial L(x,\lambda_1)}{\partial x_1} = 2x_1 + \lambda_1 = 0, \\ \dfrac{\partial L(x,\lambda_1)}{\partial x_2} = 2x_2 + \lambda_1 = 0, \\ g_1(x) = x_1 + x_2 - 2 = 0. \end{cases}$$

Решение этой системы: $x_1^* = x_2^* = 1$, $\lambda_1^* = -2$

2. Проверим достаточные условия экстремума. Так как

$$\frac{\partial^2 L(x,\lambda_1)}{\partial x_1^2} = \frac{\partial^2 L(x,\lambda_1)}{\partial x_2^2} = 2,$$
$$\frac{\partial^2 L(x,\lambda_1)}{\partial x_1 \partial x_2} = \frac{\partial^2 L(x,\lambda_1)}{\partial x_2 \partial x_1} = 0,$$

то

$$d_x^2 L(x^*, \lambda_1^*) = 2dx_1^2 + 2dx_2^2.$$

Далее, так как

$$\frac{\partial g_1(x)}{\partial x_1} = \frac{\partial g_1(x)}{\partial x_2} = 1,$$

то

$$dg_1(x^*) = dx_1 + dx_2 = 0.$$

Выражаем отсюда дифференциал dx_1 через dx_2: $dx_1 = -dx_2$ и подставим в $d_x^2 L$. Получим $4dx_2^2 > 0$ при $dx_2 \neq 0$. Следовательно, в точке $x^* = (1, 1)^T$ – регулярный локальный условный минимум.

3. Подсчитаем значение целевой функции в точке условного экстремума: $f(x^*) = 2$.

Пример 1.5. Найти условный экстремум в задаче

$$f(x) = x_1^2 + x_2^2 \to \text{extr},$$
$$g_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0.$$

Решение. Будем следовать алгоритму, не проверяя условия регулярности.

1. Составим функцию Лагранжа:

$$L(x, \lambda_0, \lambda_1) = \lambda_0(x_1^2 + x_2^2) + \lambda_1 \left((x_1 - 1)^2 + x_2^2 - 4\right).$$

Выпишем необходимые условия экстремума первого порядка:

$$\begin{cases} \frac{\partial L(x,\lambda_0,\lambda_1)}{\partial x_1} = 2\lambda_0 x_1 + 2\lambda_1(x_1 - 1) = 0, \\ \frac{\partial L(x,\lambda_0,\lambda_1)}{\partial x_2} = 2\lambda_0 x_2 + 2\lambda_1 x_2 = 0, \\ g_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0. \end{cases} \tag{1.22}$$

Решим эту систему для двух случаев.

В первом случае $\lambda_0 = 0$. Из (1.22) получим

$$2\lambda_1(x_1 - 1) = 0, \quad 2\lambda_1 x_2 = 0, \quad (x_1 - 1)^2 + x_2^2 - 4 = 0. \tag{1.23}$$

Так как λ_0 и λ_1 одновременно в нуль не обращаются, то $\lambda_1 \neq 0$. Тогда из первых двух соотношений (1.23) получим: $x_1 = 1$, $x_2 = 0$.

Однако при этих значениях не выполняется ограничение (см. третье соотношение в (1.23)). Следовательно, при $\lambda_0 = 0$ система (1.22) несовместна.

Во втором случае $\lambda_0 \neq 0$. В этом случае можно взять $\lambda_0 = 1$, и система (1.22) перепишется в виде:

$$\begin{cases} 2x_1 + 2\lambda_1(x_1 - 1) = 0, \\ 2x_2 + 2\lambda_1 x_2 = 0, \\ (x_1 - 1)^2 + x_2^2 - 4 = 0. \end{cases} \tag{1.24}$$

Рассмотрим второе уравнение. Если $x_2 = 0$, то из третьего уравнения следует $\lambda_1 = 3$, $x_1 = -1$, а из первого $\lambda_1 = -\dfrac{3}{2}$, $x_1 = -\dfrac{1}{2}$ соответственно. Если $x_2 \neq 0$, то из второго уравнения (1.24) получим $\lambda_1 = -1$ и первое уравнение будет иметь вид: $2 = 0$, то есть система несовместна. Таким образом, найдены две условно-стационарные точки:

A: $x_1^* = 3$, $x_2^* = 0$, $\lambda_1^* = -\dfrac{3}{2}$;

B: $x_1^* = -1$,$x_2^* = 0$,$\lambda_1^* = -\dfrac{1}{2}$.

2. Проверим достаточные условия экстремума, используя (1.22). Имеем:

$$d_x^2 L(x^*, \lambda^*) = 2(1+\lambda_1^*)dx_1^2 + 2(1+\lambda_1^*)dx_2^2;$$
$$dg_1(x^*) = 2(x_1^* - 1)dx_1 + 2x_2^* dx_2 = 0.$$

Исследуем точку A: $dg_1(A) = 4dx_1 + 0 = 0$, отсюда $dx_1 = 0$ и $\tilde{d}_x^{\,T} \Lambda \tilde{d}_x = -dx_2^2 < 0$ при $dx_2 \neq 0$. Поэтому в точке A – регулярный условный максимум.

Исследуем точку B: $dg_1(B) = -4dx_1 + 0 = 0$, откуда $dx_1 = 0$ и $\tilde{d}_x^{\,T} \Lambda \tilde{d}_x = dx_2^2 > 0$ при $dx_2 \neq 0$. Поэтому в точке B – регулярный условный минимум.

3. Подсчитаем значения целевой функции в точках экстремума:

$f(A) = 9$, $f(B) = 1$.

1.4.3 Условный экстремум при ограничениях типа неравенств

Пусть целевая функция $f(x)$ и функции ограничений типа неравенств $g_j(x)$, $j = \overline{1,m}$, определяющие множество допустимых точек X, дважды непрерывно дифференцируемы на некотором множестве, содержащем X.

Требуется исследовать функцию $f(x)$ на экстремум, то есть определить точки $x^* \in X$ ее локальных минимумов и максимумов на множестве X

$$f(x^*) = \min_{x \in X} f(x) \quad \left(f(x^*) = \max_{x \in X} f(x) \right), \tag{1.25}$$

где

$$X = \{x \mid g_j \le 0, \quad j = \overline{1,m}\}.$$

Точки x^* находятся с помощью необходимых и достаточных условий условного минимума и максимума первого и второго порядка при ограничениях типа неравенств.

Теорема 1.17 (необходимые условия минимума (максимума) первого порядка). *Пусть x^* – точка локального минимума (максимума) в задаче* (1.25). *Тогда найдется такое число $\lambda_0^* \geq 0$ и вектор $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)^T$, не равные одновременно нулю и такие, что выполняются следующие условия:*

– условие стационарности функций Лагранжа по x:

$$\frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = \overline{1,n}; \tag{1.26}$$

– условие допустимости решения:

$$g_j(x^*) \leq 0, \quad j = \overline{1,m}; \tag{1.27}$$

– условие неотрицательности для условного минимума:

$$\lambda_j^* \geq 0, \quad j = \overline{1,m} \tag{1.28}$$

(условие неположительности для условного максимума:

$$\lambda_j^* \leq 0, \quad j = \overline{1,m});$$

– условие дополняющей нежесткости:

$$\lambda_j^* g_(x^*) = 0, \quad j = \overline{1,m}. \tag{1.29}$$

Если при этом градиенты активных в точке x^* ограничений линейно независимы (выполняется условие регулярности), то $\lambda_0^* \neq 0$.

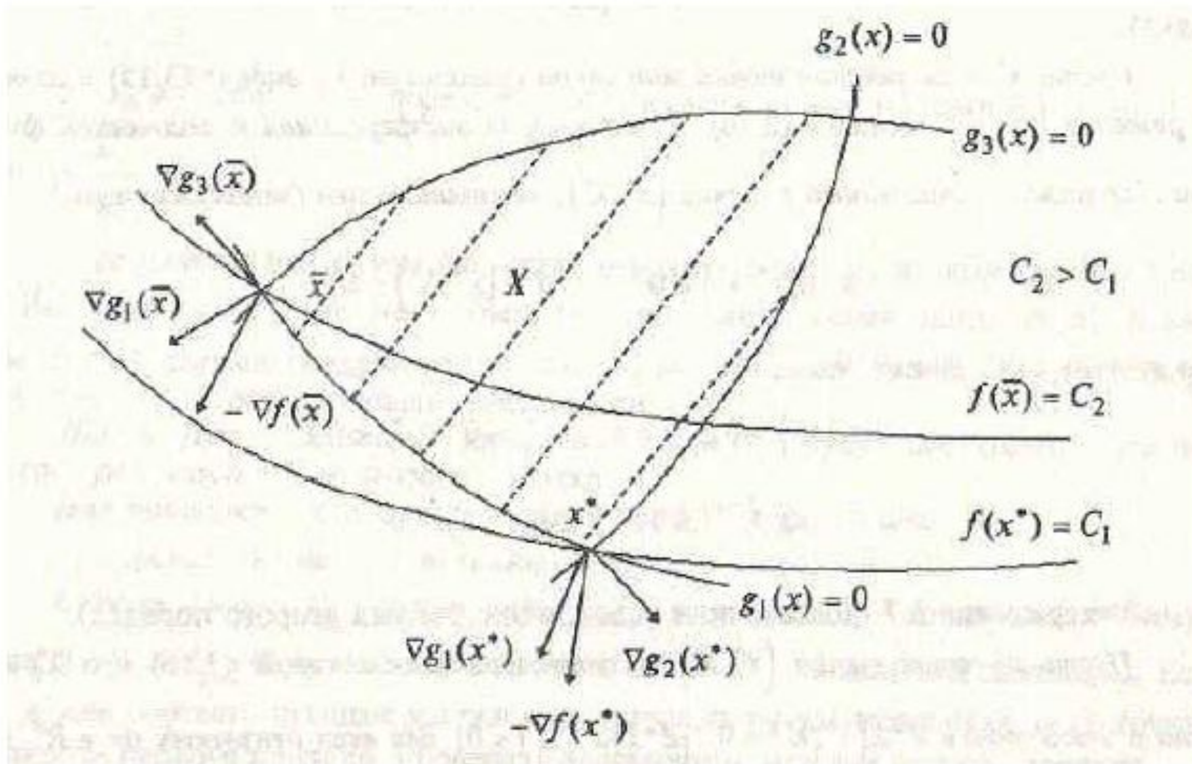

Рис. 1.2:

Будем, как и раньше, называть точки x^*, удовлетворяющие системе (1.26)–(1.29), *условно-стационарными.*

Если в решаемой задаче (1.25) ограничение записано в форме $g_j \geq 0$, то его необходимо записать в виде $-g_j \leq 0$.

Далее будем использовать $J_a(x^*)$ – *множество индексов ограничений, активных в точке* x^*.

Так как точка x^* заранее неизвестна, то проверка условия регулярности затруднена. Поэтому рекомендуется рассматривать два случая: а) $\lambda_0^* = 0$ и б) $\lambda_0^* = 1$, как и раньше. Точка экстремума, удовлетворяющая системе (1.26)–(1.29) при $\lambda_0^* \neq 0$, называется *регулярной*, а при $\lambda_0^* = 0$ – *нерегулярной*. Последний случай отражает вырожденность ограничений.

Условие (1.26) в регулярной точке экстремума x^* отражает тот факт, что антиградиент целевой функции является неотрицательной (положительной в случае максимума) линейной комбинацией градиентов функций, образующих активные ограничения в точке x^* (см. рис. 1.2). Действительно, условие (1.26) с учетом условия (1.29) можно переписать в виде:

$$-f'(x^*) = \sum_{j=1}^{m} \lambda_j^* g_j'(x^*) = \sum_{j \in J_a} \lambda_j^* g_j'(x^*).$$

В случае гладкой ОЗВП $P = \mathbb{R}^n$ и $\lambda_0 \neq 0$, условия (1.26)–(1.29) достаточны для того, чтобы допустимая точка x^* была решением задачи. Таким образом, получаем необходимые и достаточные условия решения гладкой задачи ОЗВП, которые называются теоремой Куна-Таккера в дифференциальной форме. Для того чтобы $\lambda_0 \neq 0$, достаточно выполнения условия Слейтера.

Из условия дополняющей нежесткости следует, что если ограничение в точке x^* пассивное, то есть $g_j(x^*) < 0$, то $\lambda_j^* = 0$, а если – активное, то есть $g_j(x^*) = 0$, то $\lambda_j^* \geq 0$ (для минимума) и $\lambda_j^* \leq 0$ (для максимума).

Теорема 1.18 (достаточные условия минимума (максимума) первого порядка). *Пусть имеется точка* (x^*, λ^*), *удовлетворяющая системе* (1.26)–(1.29) *при* $\lambda_0^* \neq 0$, *число активных ограничений в точке* x^* *совпадает с числом* n *переменных (при этом условие регулярности выполняется). Если* $\lambda_j^* > 0$ *для всех* $j \in J_a$, *то точка* x^* – *точка условного локального минимума. Если* $\lambda_j^* < 0$ *для всех* $j \in J_a$, *то* x^* – *точка условного локального максимума в задаче* (1.25).

Теорема 1.19 (необходимое условие минимума (максимума) второго порядка). *Пусть* x^* – *регулярная точка минимума (максимума) в задаче* (1.25) *и имеется решение* (x^*, λ^*) *системы* (1.26)–(1.29). *Тогда второй дифференциал классической функции Лагранжа по* x, *вычисленный в точке* (x^*, λ^*), *неотрицателен (неположителен):*

$$d_x^2 L(x^*, \lambda^*) \geq 0 \quad \left(d_x^2 L(x^*, \lambda^*) \leq 0\right)$$

для всех $dx \in \mathbb{R}^n$ *таких, что*

$$\begin{aligned} &dg_j(x^*) = 0, \quad j \in J_a, \quad \lambda_j^* > 0 (\lambda_j^* < 0); \\ &dg_j(x^*) \leq 0, \quad j \in J_a, \quad \lambda_j^* = 0. \end{aligned} \tag{1.30}$$

Теорема 1.20 (достаточные условия экстремума второго порядка). *Пусть имеется точка* (x^*, λ^*), *удовлетворяющая системе* (1.26)–(1.29) *при* $\lambda_0^* \neq 0$. *Если в этой точке* $d_x^2 L(x^*, \lambda^*) > 0$ $(d_x^2 L(x^*, \lambda^*) < 0)$ *для всех ненулевых* $dx \in \mathbb{R}^n$ *таких, что выполняются соотношения* (1.30), *то точка* x^* *является точкой локального минимума (максимума) в задаче* (1.25).

Алгоритм решения задачи (1.25):

1. Составить функцию Лагранжа

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^{m} \lambda_j g_j(x).$$

   Записать необходимые условия минимума (максимума) первого порядка (1.26)–(1.29). Решить полученную систему для двух случаев: а) $\lambda_0^* = 0$; б) $\lambda_0^* = 1$. В результате найти условно-стационарные точки x^*, выделив из них полученные при $\lambda_0^* = 1$ (они могут быть регулярными точками экстремума). В каждом из двух случаев решение системы (1.26)–(1.29) следует начинать с рассмотрения 2^m вариантов удовлетворения условия (1.29) дополняющей нежесткости.

2. Для выделенных на шаге 1 точек проверить достаточные условия экстремума первого или второго порядка.

   Для проверки достаточных условий первого порядка:

   а) определить число l активных в точке x^* ограничений;

   б) если $l = n$ и $\lambda_j^* > 0$ для всех $j \in J_a$, то в точке x^* – локальный минимум. Если $l = n$ $\lambda_j^* \le 0$ для всех $j \in J_a$, то в точке x^* – локальный максимум. Если $l < n$ или $l = n$, но соответствующие множители Лагранжа не удовлетворяют достаточным условиям первого порядка, проверить достаточные условия второго порядка.

Для проверки достаточных условий второго порядка следует:

а) записать выражение для второго дифференциала классической функции Лагранжа в точке (x^*, λ^*):

$$d_x^2 L(x^*, \lambda^*) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_i \partial x_j} dx_i dx_j;$$

б) записать условия, накладываемые на первые дифференциалы активных ограничений:

$$\begin{cases} dg_j(x^*) = \sum\limits_{i=1}^{n} \dfrac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j \in J_a; \quad \lambda_j^* > 0 \ (\lambda_j^* < 0); \\ dg_j(x^*) = \sum\limits_{i=1}^{n} \dfrac{\partial g_j(x^*)}{\partial x_i} dx_i \leq 0, \quad j \in J_a, \quad , \lambda_j^* = 0. \end{cases} \tag{1.31}$$

в) исследовать знак второго дифференциала функции Лагранжа по x для ненулевых dx, удовлетворяющих системе (1.31). Если $d_x^2 L(x^*, \lambda^*) > 0$, то в точке x^* – условный локальный минимум. Если $d_x^2 L(x^*, \lambda^*) < 0$, то в точке x^* – условный локальный максимум.

Если достаточные условия первого и второго порядка не выполняются, следует проверить выполнение необходимых условий второго порядка (1.19), следуя аналогичной процедуре. Если они выполняются, то требуется дополнительное исследование, а если не выполняются, то в точке x^* нет условного экстремума.

3. Вычислить значения целевой функции в точках условного экстремума.

Пример 1.6. Найти условный минимум в задаче

$$\begin{aligned} f(x) &= x_1^2 + (x_2 - 2)^2 \to \min \\ g_1(x) &= x_1^2 + x_2^2 - 1 \leq 0, \\ g_2(x) &= -x_1 \leq 0, \\ g_3(x) &= -x_2 \leq 0. \end{aligned}$$

Решение. Решаем задачу в соответствии с алгоритмом.

1. Составим функцию Лагранжа:

$$L(x, \lambda_0, \lambda) = \lambda_0 \left(x_1^2 + (x_2 - 2)^2\right) + \lambda_1(x_1^2 + x_2^2 - 1) + \lambda_2(-x_1) + \lambda_3(-x_2).$$

Выпишем необходимые условия минимума первого порядка:

а)

$$\begin{aligned} \tfrac{\partial L(x,\lambda_0,\lambda)}{\partial x_1} &= 2\lambda_0 x_1 + 2\lambda_1 x_1 - \lambda_2 = 0, \\ \tfrac{\partial L(x,\lambda_0,\lambda)}{\partial x_2} &= 2\lambda_0(x_2 - 2) + 2\lambda_1 x_2 - \lambda_3 = 0; \end{aligned}$$

б) $x_1^2 + x_2^2 - 1 \leq 0$, $-x_1 \leq 0$, $-x_2 \leq 0$;

в) $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, $\lambda_3 \geq 0$;

г) $\lambda_1(x_1^2 + x_2^2 - 1) = 0$, $\lambda_2(-x_1) = 0$, $\lambda_3(-x_2) = 0$.

Решим систему а)–г) для двух случаев.

В первом случае $\lambda_0 = 0$. Тогда условия а) запишутся в виде

$$\begin{cases} 2\lambda_1 x_1 - \lambda_2 = 0, \\ 2\lambda_1 x_2 - \lambda_3 = 0. \end{cases}$$

Рассмотрим восемь вариантов выполнения условий г) дополняющей нежесткости:

1) $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 0$. В этом случае не удовлетворяются требования (1.17), так как $\lambda^* = (\lambda_1^*, \lambda_2^*, \lambda_3^*) \neq 0$;
2) $\lambda_1 \neq 0$, $\lambda_2 = 0$, $\lambda_3 = 0$. Тогда $x_1 = x_2 = 0$ из условия а), но первое условие дополняющей нежесткости не удовлетворяется;
3) $\lambda_1 = 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$. Тогда из первого уравнения в условии а) имеем $\lambda_2 = 0$, то есть имеем противоречие;
4) $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 \neq 0$. Тогда из второго уравнения в условии а) имеем $\lambda_3 = 0$, то есть имеем противоречие;
5) $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$. Тогда $x_1 = 0$ из условия г), и из первого уравнения в условии а) имеем $\lambda_2 = 0$, то есть имеем противоречие;
6) $\lambda_1 \neq 0$, $\lambda_2 = 0$, $\lambda_3 \neq 0$. Тогда $x_2 = 0$ из условия г), и из второго уравнения в условии а) имеем $\lambda_3 = 0$, то есть также имеем противоречие;
7) $\lambda_1 = 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$. Тогда не выполняются оба уравнения в условии а);
8) $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$.

Тогда уравнения $x_1 = x_2 = 0$, $x_1^2 + x_2^2 - 1 = 0$, следующие из условия г), вместе не выполняются.

Таким образом, в случае $\lambda_0 = 0$ условно-стационарные точки не найдены.

Во втором случае $\lambda_0 \neq 0$. Как уже было отмечено раньше, нужно взять $\lambda_0 = 1$. Получим:

а)
$$\frac{\partial L(x,\lambda)}{\partial x_1} = 2x_1 + 2\lambda_1 x_1 - \lambda_2 = 0,$$
$$\frac{\partial L(x,\lambda)}{\partial x_2} = 2(x_2 - 2) + 2\lambda_1 x_2 - \lambda_3 = 0;$$

б) $x_1^2 + x_2^2 - 1 \leq 0$, $-x_1 \leq 0$, $-x_2 \leq 0$;

в) $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, $\lambda_3 \geq 0$;

г) $\lambda_1(x_1^2 + x_2^2 - 1) = 0$, $\lambda_2(-x_1) = 0$, $\lambda_3(-x_2) = 0$.

Рассмотрим восемь вариантов выполнения условий дополняющей нежесткости:

1) $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 0$. Тогда из условий а) получим $x_1 = 0$, $x_2 = 2$, и не выполняется первое ограничение в условии б)
2) $\lambda_1 \neq 0$, $\lambda_2 = 0$, $\lambda_3 = 0$. Тогда $x_1^2 + x_2^2 - 1 = 0$, $2x_1(1 + \lambda_1) = 0$, $2(x_2 - 2) + 2\lambda_2 x_1 = 0$. Если $\lambda_1 = -1$, то третье уравнение не удовлетворяется. Если $x_1 = 0$, то $x_2 = \pm 1$. Ограничениям в условии б) удовлетворяет $x_2 = 1$. При этом $\lambda_1 = 1$. Получили условно-стационарную точку A: $x_1* = 0$, $x_2^* = 1$, при этом $\lambda_1^* = 1$, $\lambda_2^* = 0$, $\lambda_3 = 0$;
3) $\lambda_1 = 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$. Тогда $x_1 = 0$, $2x_1 - \lambda_2 = 0$, $2(x_2 - 2) = 0$. Получаем $\lambda_2 = 2x_1 = 0$, что противоречит условию $\lambda_2 \neq 0$;
4) $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 \neq 0$. Тогда $x_2 = 0$, $2x_1 = 0$, $2(x_2 - 2) - \lambda_3 = 0$. Получаем $\lambda_3 = -4 < 0$, что противоречит условию в);
5) $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$. Тогда $x_1^2 + x_2^2 - 1 = 0$, $x_1 = 0$, $2x_1 + 2\lambda_1 x_1 = 0$, $2(x_2 - 2) + 2\lambda_1 x_2 = 0$. Из третьего соотношения следует, что $\lambda_2 = 0$, то есть имеется противоречие;
6) $\lambda_1 \neq 0$, $\lambda_2 = 0$, $\lambda_3 \neq 0$. Тогда $x_1^2 + x_2^2 - 1 = 0$, $x_2 = 0$, $2x_1 + 2\lambda_1 x_1 = 0$, $2(x_2 - 2) + 2\lambda_1 x_2 - \lambda_3 = 0$. Из последнего соотношения следует, что $\lambda_3 = -4 < 0$. Но это противоречит условию в);
7) $\lambda_1 = 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$. Тогда $x_1 = 0$, $x_2 = 0$, $2x_1 - \lambda_2 = 0$. Из второго соотношения следует, что $\lambda_2 = 0$. Это противоречит тому, что в данном случае $\lambda_2 \neq 0$;
8) $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$. Тогда из условия г) следует, что $x_1^2 + x_2^2 - 1 = 0$, $x_1 = 0$, $x_2 = 0$. Но эта система несовместна.

2. Проверим достаточные условия минимума первого порядка для точки A. В этой точке имеется два активных ограничения, то есть $l = 2 = n$. Так как $\lambda_1^* = 1 > 0$, $\lambda_2^* = 0$, то достаточные условия минимума первого порядка не выполняются ввиду того, что требуется строгая положительность соответствующих множителей Лагранжа.

   Проверим достаточные условия минимума второго порядка для точки A. Имеем

$$d_x^2 L(A, \lambda^*) = (2 + 2\lambda_1^*)dx_1^2 + (2 + 2\lambda_1^*)dx_2^2.$$

   Так как в точке A два активных ограничения и для одного из них $\lambda_1^* > 0$, а для другого $\lambda_2^* = 0$, то применим условия (1.31):

$$dg_1(A) = 2x_1^* dx_1 + 2x_2^* dx_2 = 2dx_2 = 0,\ \ \lambda_1^* > 0;$$
$$dg_2(A) = -dx_1 \le 0,\ \ \lambda_2^* = 0.$$

   В результате $d_x^2 L(A, \lambda^*) = 4dx_1^2 > 0$ при $dx_1 \ge 0$ и $dx_1 \ne 0$. Поэтому в точке A – локальный условный минимум. С другой стороны, целевая функция и множество допустимых решений выпуклые (теорема 1.3). Поэтому в точке A достигается глобальный минимум (теорема 1.4).

3. Вычислим значение функции в точке глобального минимума: $f(A) = 1$.

1.4.4 Условный экстремум при смешанных ограничениях

Пусть целевая функция $f(x)$ и функции ограничений типа равенств и неравенств $g_j(x)$, $j = \overline{1, m}$, определяющие множество допустимых точек X, дважды непрерывно дифференцируемых на некотором множестве, содержащем X. Требуется исследовать функцию $f(x)$ на экстремум, то есть определить точки $x^* \in X$ ее локальных минимумов и максимумов на множестве X:

$$f(x^*) = \min_{x \in X} f(x) \qquad \left(f(x^*) = \max_{x \in X} f(x) \right), \tag{1.32}$$

где

$$X = \{x \mid g_j(x) \leq 0, \ j = \overline{1, k}; \ g_j(x) = 0, \ \overline{j = k+1, m}\}. \tag{1.33}$$

Точки x^* находятся с помощью необходимых и достаточных условий Условного минимумам и максимума первого и второго порядка.

Теорема 1.21 (Необходимое условие экстремума первого порядка). *Пусть x^* – точка локального минимума (максимума) в задаче* (1.32)–(1.33). *Тогда найдется такое число $\lambda_0^* \geq 0$ и вектор $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)^T$, не равные одновременно нулю и такие, что выполняются следующие условия:*

– *условие стационарности функции Лагранжа по x:*

$$\frac{\partial L(x*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = \overline{1, n}; \tag{1.34}$$

– *условие допустимости решения:*

$$g_j(x) \leq 0, \; j = \overline{1, k}; \quad g_j(x) = 0, \; \overline{j = k+1, m}; \tag{1.35}$$

– *условие неотрицательности для условного минимума:*

$$\lambda_j^* \geq 0, \quad j = \overline{1, k} \tag{1.36}$$

(условие неотрицательности для условного максимума:

$$\lambda_j^* \leq 0, \quad j = \overline{1, k};$$

– *условие дополняющей нежесткости:*

$$\lambda_j^* g_j(x^*) = 0, \quad j = \overline{1, k}. \tag{1.37}$$

В случае гладкой ОЗВП

$$f(x) \to \min, \ x \in X. \tag{1.38}$$

$$X = \{x \in P \mid g_j(x) \le 0, \ j = \overline{1,k}\}, \tag{1.39}$$

где

$$P = \{x \in X \mid g_j = a_j^T x - b_j \ j = \overline{k+1,m}\},$$

при $\lambda_0^* \neq 0$ необходимые условия, даваемые теоремой 1.21, являются и достаточными, а именно, справедлива следующая теорема.

Теорема 1.22 (Куна-Таккера, дифференциальная форма). *Для того, чтобы регулярная точка x^*, была решением задачи* (1.38)–(1.39), *где $g_j(x)$, $j = \overline{1,k}$, есть непрерывно-дифференцируемые выпуклые функции в окрестности точки x^*, а $g_j(x)$, $j = \overline{k+1,m}$, — линейные функции, необходимо и достаточно, чтобы существовал такой вектор $\lambda^* = (\lambda_1^*, \ldots, \lambda_m)$, что для классической функции Лагранжа $L(x, \lambda)$ выполнялись условия* (1.34)–(1.37) *теоремы* 1.21.

Задача (1.38)–(1.39) будет регулярной, если для нее выполняется условие Слейтера.

Отметим, что выпуклость множества X следует из теоремы 1.3.

Теорема 1.23 (Достаточные условия экстремума первого порядка). *Дана точка* (x^*, λ^*), *удовлетворяющая системе*(1.34)–(1.37) *при* $\lambda_0^* \neq 0$, *суммарное число активных ограничений-неравенств в точке* $x*$ *и ограничений-равенств совпадает с числом* n *переменных. Если* $\lambda_j^* > 0$ *для всех* $j \in J_a$, *то точка* x^* – *точка локального минимума в задаче* (1.32)–(1.33). *Если* $\lambda_j^* < 0$ *для всех* $j \in J_a$, *то* $x*$ – *точка условного максимума.*

Теорема 1.24 (необходимые условия экстремума второго порядка). *Пусть* x^* – *регулярная точка минимума (максимума) в задаче* (1.32)–(1.33) *и имеется решение* (x^*, λ^*) *системы* (1.34)–(1.37). *Тогда второй дифференциал по* x *классической функции Лагранжа* $L(x, \lambda)$, *вычисленный в точке* (x^*, λ^*), *неотрицателен (неположителен):*

$$d_x^2 L(x^*, \lambda^*) \geq 0 \quad (d_x^2 L(x^*, \lambda^*) \leq 0)$$

для всех $dx \in \mathbb{R}_n$ *таких, что*

$$\begin{cases} dg_j(x^*) = \sum\limits_{i=1}^{n} \dfrac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \;\; j = \overline{k+1, m} \text{ и } j \in J_a, \;\; \lambda_j^* > 0 \; (\lambda_j^* < 0); \\ dg_j(x^*) = \sum\limits_{i=1}^{n} \dfrac{\partial g_j(x^*)}{\partial x_i} dx_i \leq 0, \;\; j \in J_a, \; \lambda_j^* = 0. \end{cases} \tag{1.40}$$

Теорема 1.25 (достаточные условия экстремума второго порядка). *Пусть имеется точка* $(x*, \lambda^*)$, *удовлетворяющая системе* (1.34)–(1.37) *при* $\lambda_0^* \neq 0$. *Если в этой точке*

$$d_x^2 L(x^*, \lambda^*) > 0 \quad (d_x^2 L(x^*, \lambda^*) < 0)$$

для всех ненулевых $dx \in \mathbb{R}_n$ *таких, что выполняются соотношения* (1.40), *то точка* x^* *является точкой локального минимума (максимума) в задаче* (1.32)–(1.33).

Алгоритм решения задачи (1.32)–(1.33):

1. Составить функцию Лагранжа:

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^{m} \lambda_j g_j(x).$$

Записать необходимые условия условного минимума (максимума) первого порядка (1.34)–(1.37). Решить систему (1.34)–(1.37) для двух случаев: а)$\lambda_0^* = 0$; б)$\lambda_0^* = 1$. В результате найти условно-стационарные точки, выделив из них полученные при $\lambda_0^* = 1$ (они могут быть регулярными точками экстремума). В каждом из двух случаев следует начинать с рассмотрения 2^{m-k} вариантов удовлетворения условия (1.37) дополняющей нежесткости.

2. Для получения в пункте 1 условно-стационарных точек проверить достаточные условия экстремума первого или второго порядка. Для проверки достаточных условий первого порядка следует:

   1) найти число l ограничений-равенств и активных ограничений-неравенств;

   2) если $l = n$ и $\lambda_j^* > 0$ для всех $j \in J_a$, то есть для всех активных ограничений-неравенств, то в точке x^* – локальный минимум. Если $l = n$ и $\lambda_j^* < 0$ для всех $j \in J_a$, то в точке x^* – локальный максимум. Если $l < n$ или $l = n$, но соответствующие множители Лагранжа не удовлетворяют достаточным условиям первого порядка, проверить достаточные условия второго порядка.

Для проверки достаточных условий второго порядка следует:

   1) записать выражение для второго дифференциала классической функции Лагранжа $L(x, \lambda)$ в точке (x^*, λ^*):

$$d_x^2 L(x^*, \lambda_*) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_i \partial x_j} dx_i dx_j;$$

   2) записать условия, накладываемые на первые дифференциалы ограничений-равенств и активных в точке x^* ограничений неравенств (1.40);

3) исследовать знак второго дифференциала функции Лагранжа для ненулевых dx, удовлетворяющих соотношениям (1.40). Если $d_x^2 L(x^*, \lambda^*) > 0$, в точке x^* – условный локальный минимум. Если $d_x^2 L(x^*, \lambda^*) < 0$, то в точке x^* – условный локальный максимум. Если достаточные условия условного экстремума не выполняются, следует проверить выполнение необходимых условий экстремума второго порядка, следуя аналогичной процедуре. Если они выполняются, то требуется дополнительное исследование, а если нет, то в точке x^* нет условного экстремума.

3. Вычислить значения функции в точках условного экстремума.

Пример 1.7. Найти условный экстремум в задаче

$$
\begin{aligned}
&f(x) = x_1^2 - x_2 \to \text{extr},\\
&g_1(x) = x_1 + x_2 - 6 = 0,\\
&g_2(x) = 1 - x_1 \le 0,\\
&g_3(x) = x_1^2 + x_2^2 - 26 \le 0.
\end{aligned}
$$

Решение. Рассмотрим соответствуюшие пункты предложенного алгоритма решения задач со смешанными ограничениями.
1. Составим функцию Лагранжа

$$L(x, \lambda_0, \lambda) = \lambda_0(x_1^2 - x_2) + \lambda_1(x_1 + x_2 - 6) + \lambda_2(1 - x_1) + \lambda_3(x_1^2 + x_2^2 - 26).$$

Выпишем необходимые условия экстремума первого порядка:

а) $\frac{\partial L(x,\lambda_0,\lambda)}{\partial x_1} = 2\lambda_0 x_1 + \lambda_1 - \lambda_2 + 2\lambda_3 x_1 = 0$, $\frac{\partial L(x,\lambda_0,\lambda)}{\partial x_2} = -\lambda_0 + \lambda_1 + 2\lambda_3 x_2 = 0$;

б) $x_1 + x_2 - 6 = 0$, $1 - x_1 \leq 0$, $x_1^2 + x_2^2 - 26 \leq 0$;

в) $\lambda_2 \geq 0$, $\lambda_3 \geq 0$ (для минимума), $\lambda_2 \leq 0$, $\lambda_3 \leq 0$ (для максимума);

г) $\lambda_2(1 - x_1) = 0$, $\lambda_3(x_1^2 + x_2^2 - 26) = 0$.

Решим систему для двух случаев.
В первом случае $\lambda_0 = 0$. Тогда условия а) имеют вид:

$$\lambda_1 - \lambda_2 + 2\lambda_3 x_1 = 0, \quad \lambda_1 + 2\lambda_3 x_2 = 0.$$

Рассмотрим четыре варианта удовлетворения условий г) дополняющей нежесткости:

1) $\lambda_2 = 0$, $\lambda_3 = 0$. Тогда $\lambda_1 = 0$ и не удовлетворяется необходимое условие экстремума, даваемое теоремой 4.8;

2) $\lambda_2 \neq 0$, $\lambda_3 = 0$. Тогда $\lambda_1 = 0$, а отсюда и $\lambda_2 = 0$, что противоречит условию $\lambda_2 \neq 0$;

3) $\lambda_2 = 0$, $\lambda_3 \neq 0$. Тогда

$$\begin{aligned} &x_1^2 + x_2^2 - 26 = 0, \\ &x_1 + x_2 - 6 = 0, \\ &\lambda_1 + 2\lambda_3 x_1 = 0, \\ &\lambda_1 + 2\lambda_3 x_2 = 0. \end{aligned}$$

Из двух последних уравнений следует: $2\lambda_3(x_2 - x_1) = 0$. Так как $\lambda_3 \neq 0$, то $x_1 = x_2$. Из двух первых уравнений следует: $x_1 = 1$, $x_2 = 5$ или $x_1 = 5$, $x_2 = 1$, то есть $x_1 \neq x_2$. Поэтому система несовместна;

4) $\lambda_2 \neq 0$, $\lambda_3 \neq 0$. Тогда

$$\begin{aligned} &x_1^2 + x_2^2 - 26 = 0, \\ &1 - x_1 = 0, \\ &x_1 + x_2 - 6 = 0. \end{aligned}$$

Система удовлетворяется в точке $x_1 = 1$, $x_2 = 5$. Условия а) примут вид

$$\begin{aligned} &\lambda_1 - \lambda_2 + 2\lambda_3 = 0, \\ &\lambda_1 + 10\lambda_3 = 0. \end{aligned}$$

Отсюда $\lambda_1 = -10\lambda_3$ и $\lambda_2 = -8\lambda_3$. Так как $\lambda_3 \neq 0$, то λ_2 и λ_3 имеют разные знаки, что противоречит условию и минимума, и максимума.

В втором случае $\lambda_0 \neq 0$. Поделим соответствующие уравнения системы а) – г) на λ_0 и заменим $\frac{\lambda_1}{\lambda_0}$ на λ_1, $\frac{\lambda_2}{\lambda_0}$ на λ_2, $\frac{\lambda_3}{\lambda_0}$ на λ_3. Условие а) примет вид

$$2x_1 + \lambda_1 - \lambda_2 + 2\lambda_3 x_1 = 0, \quad -1 + \lambda_1 + 2\lambda_3 x_2 = 0.$$

Условия б) – г) сохраняют вид.

Рассмотрим четыре варианта удовлетворения условий г) дополняющей нежесткости:

1) $\lambda_2 = 0$, $\lambda_3 = 0$. Тогда $\lambda_1 = 1$, а $x_1 = -\frac{1}{2}$, что не удовлетворяет ограничениям б);

2) $\lambda_2 \neq 0$, $\lambda_3 = 0$. Тогда $x_1 = 1$, $x_2 = 5$, $\lambda_1 = 1$, $\lambda_2 = 3$. Получена условно-стационарная точка A: $x_1^* = 1$, $x_2^* = 5$, $\lambda_1^* = 1$, $\lambda_2^* = 3$, в которой удовлетворяются необходимые условия минимума;

3) $\lambda_2 = 0$, $\lambda_3 \neq 0$. Тогда

$$\begin{aligned}
&x_1^2 + x_2^2 - 26 = 0,\\
&x_1 + x_2 - 6 = 0,\\
&2x_1 + \lambda_1 + 2\lambda_3 x_1 = 0,\\
&-1 + \lambda_1 + 2\lambda_3 x_2 = 0.
\end{aligned}$$

Отсюда получаем точки с координатами $x_1 = 1$, $x_2 = 5$ и $x_1 = 5$, $x_2 = 1$. В первой точке имеем

$$\begin{aligned}
&2 + \lambda_1 + 2\lambda_3 = 0,\\
&-1 + \lambda_1 + 10\lambda_3 = 0,
\end{aligned}$$

откуда $\lambda_1 = -\frac{11}{4}$, $\lambda_3 = \frac{3}{8}$. Во второй точке имеем

$$\begin{aligned}
&10 + \lambda_1 + 10\lambda_3 = 0,\\
&-1 + \lambda_1 + 2\lambda_3 = 0,
\end{aligned}$$

откуда $\lambda_1 = \frac{15}{4}$, $\lambda_3 = -\frac{11}{8}$. Получены условно-стационарные точка A': $x_1^* = 1$, $x_2^* = 5$, $\lambda_1^* = -\frac{11}{4}$, $\lambda_2^* = 0$, $\lambda_3^* = \frac{3}{8}$, в которой удовлетворяются необходимые условия минимума, и точка B: $x_1^* = 5$, $x_2^* = 1$, $\lambda_1^* = \frac{15}{4}$, $\lambda_2^* = 0$, $\lambda_3^* = -\frac{11}{8}$, в которой удовлетворяются необходимые условия максимума;

4) $\lambda_2 \neq 0$, $\lambda_3 \neq 0$. Тогда получаем систему

$$\begin{aligned} &x_1^2 + x_2^2 - 26 = 0, \\ &1 - x_1 = 0, \\ &x_1 + x_2 - 6 = 0, \end{aligned}$$

которая удовлетворяется в точке $x_1 = 1, x_2 = 5$. Условия а) примут вид

$$\begin{aligned} &2 + \lambda_1 - \lambda_2 + 2\lambda_3 = 0, \\ &-1 + \lambda_1 + 10\lambda_3 = 0. \end{aligned}$$

Отсюда $\lambda_1 = 1 - 10\lambda_3$ и $3 - \lambda_2 - 8\lambda_3 = 0$. Так как $\lambda_2 \neq 0$ и $\lambda_3 \neq 0$, а также они должны быть одного знака, то последнее равенство выполняется только при $\lambda_2 > 0$ и $\lambda_3 > 0$, в частности, при $\lambda_2 = 2,2$ и $\lambda_3 = 0,1$. При этом $\lambda_1 = 0$. Получили ту же условно-стационарную точку A'': A': $x_1^* = 1$, $x_2^* = 5$, $\lambda_1^* = 0$, $\lambda_2^* = 2,2$, $\lambda_3^* = 0,1$.

2. Проверим выполнение достаточных условий экстремума первого порядка. Ограничение-равенство в точках A и B естественно выполняется. В точке A активно второе ограничение и, следовательно, $l = 2 = n$. Так как $\lambda_2^* = 3 > 0$, то в точке A — условный локальный минимум (см. теорему). В точке A' активно третье ограничение и, поэтому, $l = 2 = n$. Так как $\lambda_3^* = \frac{3}{8} > 0$, то в точке A' — условный локальный минимум. В точке B активно третье ограничение и, поэтому, $l = 2 = n$. Так как $\lambda_3^* = -\frac{11}{8} < 0$, то в точке B — условный локальный максимум (см. теорему).

Проверим выполнение достаточных условий экстремума второго порядка из методических соображений. Имеем:

$$d_x^2 L(x^*, \lambda^*) = (2 + 2\lambda_3^*)dx_1^2 + 2\lambda_3^* dx_2^2.$$

В точке A активно второе ограничение. Поэтому:

$$\begin{aligned} dg_1(A) &= dx_1 + dx_2 = 0, \\ dg_2(A) &= -dx_1 = 0. \end{aligned}$$

Отсюда $dx_1 = dx_2 = 0$ и, таким образом, $d_x^2 L(A) \equiv 0$. Поэтому требуется дополнительное исследование.

В точке A' активно третье ограничение. Поэтому:

$$\begin{aligned} dg_1(A') &= dx_1 + dx_2 = 0, \\ dg_3(A') &= 2x_1^* dx_1 + 2x_2^* dx_2 = 2dx_1 + 10dx_2 = 0. \end{aligned}$$

Отсюда $dx_1 = dx_2 = 0$ и, таким образом, $d_x^2 L(A') \equiv 0$. Поэтому тоже требуется дополнительное исследование.

В точке B активно третье ограничение. Поэтому:

$$dg_1(B) = dx_1 + dx_2 = 0,$$
$$dg_3(B) = 2x_1^* dx_1 + 2x_2^* dx_2 = 10dx_1 + 2dx_2 = 0.$$

Отсюда $dx_1 = dx_2 = 0$ и, таким образом, $d_x^2 L(B) \equiv 0$. Поэтому требуется дополнительное исследование.

В точке A'' активны второе и третье ограничения. Поэтому:

$$dg_1(A'') = dx_1 + dx_2 = 0,$$
$$dg_2(A'') = -dx_1 = 0,$$
$$dg_3(A') = 2x_1^* dx_1 + 2x_2^* dx_2 = 2dx_1 + 10dx_2 = 0.$$

Отсюда $dx_1 = dx_2 = 0$ и, таким образом, $d_x^2 L(A'') \equiv 0$. Поэтому тоже требуется дополнительное исследование.

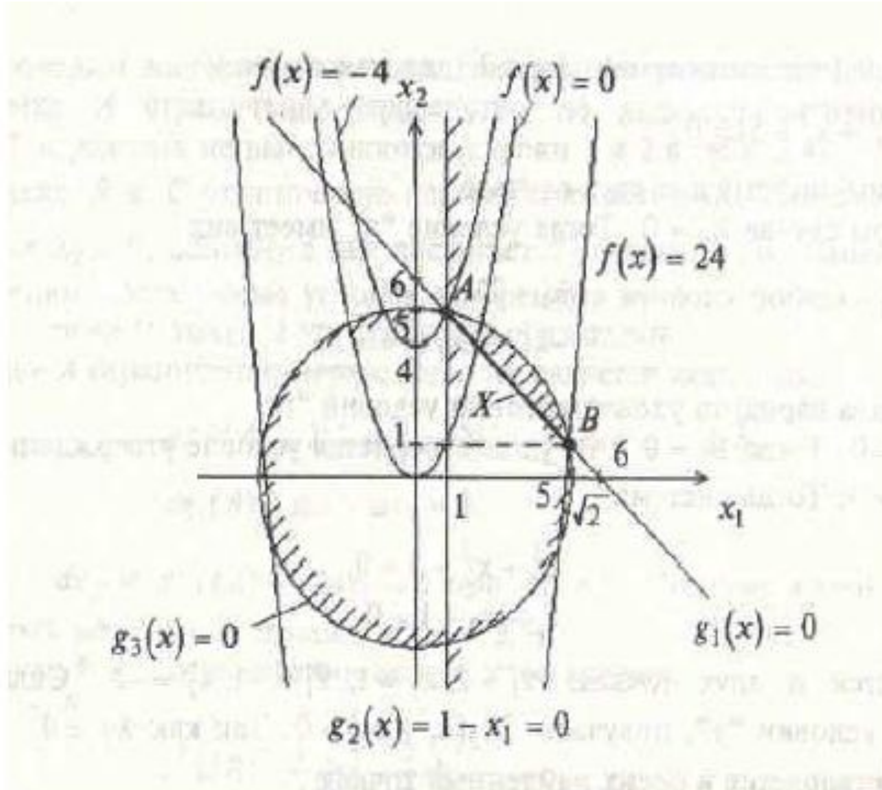

Рис. 1.3:

Из рис. 1.3 следует, что в точках A и B — соответственно глобальный минимум и максимум.

Теперь исследуем свойства целевой функции и ограничений. Ограничение-равенство — линейное. Так как целевая функция и функции второго и третьего ограничений удовлетворяют условиям

$$H_f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \geq 0,\ H_{g_2}(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \geq 0,\ H_{g_3}(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} > 0,$$

то они выпуклы (см.). Поэтому в точке A — глобальный минимум (см. теорему). Так как функция $-f(x) = -x_1^2 + x_2$ не является выпуклой, то вывода о глобальном максимуме с помощью необходимых условий первого порядка сделать нельзя.

3. Целевая функция в точках условного экстремума принимает следующие значения: $f(A) = -4$, $f(B) = 24$. $\square$

Продолжение следует…