НЕЛИНЕЙНО ОПТИМИРАНЕ

1. Формулировка на задачата

Нека са дадени подмножество X на евклидовото пространство $\mathbb{R}^n$ и функция f, дефинирана в X, със стойности реални числа. Общата задача на математическото оптимиране се формулира по следния начин: Да се намери такава точка $\mathbf{x}^*$ от множеството X, че за всяко $\mathbf{x} \in X$ да е изпълнено $f(\mathbf{x}^*) \leq f(\mathbf{x})$.

Едно от кратките записвания на тази формулировка е

$$f(\mathbf{x}) \to \min, \quad \mathbf{x} \in X. \tag{1}$$

Задачата за максимум $f(\mathbf{x}) \to \max$, $\mathbf{x} \in X$, е еквивалентна на задачата за минимум $-f(\mathbf{x}) \to \min$, $\mathbf{x} \in X$. Задачите за минимум и максимум се наричат задачи за екстремум или *екстремални задачи*.

Тук ще разглеждаме задачи, в които множеството X се задава с краен брой равенства и неравенства

$$X = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0,\ i = 1, \dots, m,\ h_j(\mathbf{x}) = 0,\ j = 1, \dots, s\}, \tag{2}$$

където g_i, $i = 1, \dots, m$, и h_j, $j = 1, \dots, s$, са функции, дефинирани в $\mathbb{R}^n$, със стойности в $\mathbb{R}$. Ще предполагаме, че функциите f, g_i и h_j имат дадени аналитични свойства (непрекъснатост, диференцуемост), но не са непременно от определен вид (линейни, квадратични и т. н.). За да различим този кръг задачи от останалите по-специфични класове, говорим за *задачи на нелинейното оптимиране*.

По-нататък в изложението ще използваме и други форми на записване на задача (1), например $\min\{f(\mathbf{x}) : \mathbf{x} \in X\}$. Ако X е от вида (2), понякога ще използваме следното записване:

$$f(\mathbf{x}) \to \min \tag{3}$$

при условия

$$\begin{aligned} &g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \\ &h_j(\mathbf{x}) = 0, \quad j = 1, \dots, s, \\ &\mathbf{x} \in \mathbb{R}^n. \end{aligned}$$

2. Класически теореми от анализа

Най-напред ще припомним някои теореми, които би трябвало да са известни от курса по математически анализ (диференциално смятане).

Теорема 1 (Вайерщрас)**.** Нека X е непразно компактно подмножество на $\mathbb{R}^n$ и функцията f е непрекъсната във всяка точка $\mathbf{x} \in X$. Тогава $f(\mathbf{x})$ достига инфимума и супремума си в X.

Следствие 1. Нека X е непразно затворено подмножество на $\mathbb{R}^n$, f е непрекъсната в X функция и освен това за всяка редица $\{\mathbf{x}_k\}$, $\mathbf{x}_k \in X$, $k = 1, 2, \dots$, за която $\lim\limits_{k\to+\infty} \|\mathbf{x}_k\| = +\infty$, е изпълнено $\lim\limits_{k\to+\infty} f(\mathbf{x}_k) = +\infty$ (наричано често *условие за коерцитивност*). Тогава задачата (1) има оптимално решение.

Теорема 2 (Ферма)**.** Нека $\mathbf{x}^*$ е точка на локален минимум на f в $\mathbb{R}^n$ и нека f е диференцуема в точката $\mathbf{x}^*$. Тогава $f'(\mathbf{x}^*) = \mathbf{0}$, където с $f'(\mathbf{x}^*) = (f'_{x_1}(\mathbf{x}^*), \dots, f'_{x_n}(\mathbf{x}^*))^{\mathrm{T}}$ означаваме производната (градиента) на f в точката $\mathbf{x}^*$.

Теоремата на Ферма е *необходимо условие* за локален екстремум. *Достатъчно условие* се дава от следната

Теорема 3. Нека f е два пъти непрекъснато диференцуема в околност на точката $\mathbf{x}^*$ и освен това

1. $f'(\mathbf{x}^*) = \mathbf{0}$;
2. матрицата от вторите производни $f''(\mathbf{x}^*) = \left[\dfrac{\partial^2 f(\mathbf{x}^*)}{\partial x_i \partial x_j}\right]_{(n\times n)}$ е положително дефинитна.

Тогава $\mathbf{x}^*$ е точка на локален минимум на f в $\mathbb{R}^n$.

3. Необходими условия за локален екстремум при наличие на ограничения

Ще разгледаме поотделно три случая: ограничения равенства, ограничения неравенства и накрая ограничения равенства и неравенства.

3.1. Необходимо условие за локален екстремум за задачата с ограничения равенства

Теорема 4 (Лагранж)**.** Нека $\mathbf{x}^*$ е точка на локален минимум за функцията f при условия $h_j(\mathbf{x}) = 0$, $j = 1, \dots, s$. Да предположим, че функцията f е диференцуема в $\mathbf{x}^*$, а функциите h_j, $j = 1, \dots, s$, имат непрекъснати частни производни в околност на $\mathbf{x}^*$. Тогава съществуват числа $\lambda_0, \lambda_1, \dots, \lambda_s$, не всички равни на нула, такива че

$$\lambda_0 f'(\mathbf{x}^*) + \sum_{j=1}^{s} \lambda_j h_j'(\mathbf{x}^*) = \mathbf{0}.$$

Функцията $L(\mathbf{x}, \boldsymbol{\lambda}) = \lambda_0 f(\mathbf{x}) + \sum\limits_{j=1}^{s} \lambda_j h_j(\mathbf{x})$ се нарича *функция на Лагранж*, а числата λ_j, $j = 0, 1, \dots, s$ — *множители на Лагранж*. Теорема 4 е известна като *принцип на Лагранж*. Тази терминология се запазва и при другите задачи, които ще бъдат разгледани по-нататък.

Ако векторите $h_j'(\mathbf{x}^*)$, $j = 1, \dots, s$, са линейно независими, то $\lambda_0 \neq 0$. Ще отбележим, че векторът $(\lambda_0, \lambda_1, \dots, \lambda_s)^{\mathrm{T}}$ може да бъде избран с точност до ненулев реален множител. Следователно, ако $\lambda_0 \neq 0$, можем да приемем $\lambda_0 = 1$.

Пример 1. Да се реши задачата

$$f(x, y) = x^2 + y^2 \to \min, \quad (x, y) \in X = \{(x, y) : x^6 + y^6 = 1\}.$$

Решение. Множеството X е компактно (докажете), а f е непрекъсната функция. Следователно задачата има поне едно оптимално решение. Всяко оптимално решение ще удовлетворява принципа на Лагранж (теорема 4). Функцията на Лагранж е

$$L(x, y, \lambda_0, \lambda_1) = \lambda_0(x^2 + y^2) + \lambda_1(x^6 + y^6 - 1).$$

Следователно, ако (x, y) е оптимално решение, то съществуват множители на Лагранж $(\lambda_0, \lambda_1) \neq (0, 0)$, такива че

$$\left| \begin{array}{l} L_x' \equiv 2\lambda_0 x + 6\lambda_1 x^5 = 0 \\ L_y' \equiv 2\lambda_0 y + 6\lambda_1 y^5 = 0. \end{array} \right.$$

Ако $\lambda_0 = 0$, тъй като $\lambda_1 \neq 0$, получаваме $x = y = 0$, което е невъзможно. Следователно $\lambda_0 \neq 0$ и можем да приемем, че $\lambda_0 = 1$. И така получаваме системата

$$\left| \begin{array}{l} x + 3\lambda_1 x^5 = 0 \\ y + 3\lambda_1 y^5 = 0 \end{array} \right. \iff \left| \begin{array}{l} x(1 + 3\lambda_1 x^4) = 0 \\ y(1 + 3\lambda_1 y^4) = 0, \end{array} \right.$$

която има следните решения:

1) $x = 0$, $y = \pm 1$, $\lambda_1 = -\frac{1}{3}$;

2) $x = \pm 1$, $y = 0$, $\lambda_1 = -\frac{1}{3}$;

3) $x = \pm 1/\sqrt[6]{2}$, $y = \pm 1/\sqrt[6]{2}$, $\lambda_1 = -\frac{1}{3}\sqrt[3]{4}$.

Във всяка друга точка на X равенството в теоремата на Лагранж не е изпълнено. Следователно поне една от горните точки е точка на минимум. Остава да намерим стойностите на f в тези точки: $f(0, \pm 1) = f(\pm 1, 0) = 1$, $f(\pm 1/\sqrt[6]{2}, \pm 1/\sqrt[6]{2}) = \sqrt[3]{4} > 1$. Получаваме, че точките 1) и 2) са оптимални решения на задачата (останалите точки са точки на максимум).

Пример 2. Да се реши задачата

$$f(x, y) = y \to \min, \quad (x, y) \in X, \quad X = \{(x, y) : x^4 - y^5 = 0\}.$$

Решение. От ограничението следва, че $y = \sqrt[5]{x^4} \geq 0$. Тогава $f(x, y)$ е ограничена отдолу в X, при това $f(x, y) \geq 0$ за всяко $(x, y) \in X$. Очевидно оптималното решение е $(x^*, y^*) = (0, 0)$. Функцията на Лагранж е

$$L(x, y, \lambda_0, \lambda_1) = \lambda_0 y + \lambda_1(x^4 - y^5), \quad (\lambda_0, \lambda_1) \neq (0, 0).$$

Условието в теоремата на Лагранж има вида

$$4\lambda_1 x^3 = 0, \qquad \lambda_0 - 5\lambda_1 y^4 = 0.$$

Ако допуснем, че $\lambda_1 = 0$, веднага получаваме и $\lambda_0 = 0$, което е противоречие. Следователно $\lambda_1 \neq 0$. Тогава $x = 0$. От ограничението следва и $y = 0$. Множителите на Лагранж са $(0, \lambda_1)$, λ_1 — произволно реално число, различно от нула.

3.2. Необходимо условие за локален екстремум за задачата с ограничения неравенства

Теорема 5 (Джон)**.** Нека $\mathbf{x}^*$ е точка на локален минимум за функцията f при условия $g_i(\mathbf{x}) \leq 0$, $i = 1, \ldots, m$. Ако $\mathbf{x}$ е допустима точка, с $I(\mathbf{x}) := \{i : g_i(\mathbf{x}) = 0\}$ означаваме множеството на активните в $\mathbf{x}$ ограничения. Да предположим, че функциите f и g_i, $i \in I(\mathbf{x}^*)$, са диференцуеми в $\mathbf{x}^*$, а функциите g_i, $i \notin I(\mathbf{x}^*)$, са непрекъснати в $\mathbf{x}^*$. Тогава съществуват неотрицателни числа μ_0, μ_i, $i \in I(\mathbf{x}^*)$, не всички равни на нула, такива че

$$\mu_0 f'(\mathbf{x}^*) + \sum_{i \in I(\mathbf{x}^*)} \mu_i g_i'(\mathbf{x}^*) = \mathbf{0}.$$

Ако g_i, $i \notin I(\mathbf{x}^*)$, също са диференцуеми в $\mathbf{x}^*$, теоремата на Джон изглежда така: съществуват неотрицателни числа μ_i, $i = 0, 1, \ldots, m$, не всички равни на нула, такива че

$$\mu_0 f'(\mathbf{x}^*) + \sum_{i=1}^{m} \mu_i g_i'(\mathbf{x}^*) = \mathbf{0}, \qquad \mu_i g_i(\mathbf{x}^*) = 0, \quad i = 1, \ldots, m.$$

Пример 3. Да се реши задачата

$$f(x, y) = x^2 + y^3 \to \min, \qquad (x, y) \in X,$$

където

а) $X = \{(x, y) : x^2 + y \le 1\}$;

б) $X = \{(x, y) : x^2 + y \le 1, \ -y \le 1\}$.

Решение. а) Множеството X не е ограничено, следователно не можем да приложим теоремата на Вайерщрас. Функцията f не е коерцитивна, защото можем да вземем редица от точки $(0, y_k)$ на X, $k = 1, 2, \ldots$, за която $y_k < 1$, $y_k \to -\infty$. Тогава $\lim\limits_{k\to+\infty} f(0, y_k) = \lim\limits_{k\to+\infty} y_k^3 \to -\infty$. Така не можем да приложим и следствие 1.

Все пак да се опитаме да намерим „подозрителни“ точки, които удовлетворяват условията на теорема 5. Функцията на Лагранж за тази задача е

$$L(x, y, \mu_0, \mu_1) = \mu_0(x^2 + y^3) + \mu_1(x^2 + y - 1),$$

като $\mu_0 \ge 0$, $\mu_1 \ge 0$, $(\mu_0, \mu_1) \ne (0, 0)$. Така получаваме

$$\begin{aligned} &\mu_0 2x + \mu_1 2x = 0, && \mu_0 3y^2 + \mu_1 = 0, \\ &\mu_1(x^2 + y - 1) = 0, && \mu_0 \ge 0, \ \mu_1 \ge 0. \end{aligned}$$

Ако $\mu_0 = 0$, то от второто равенство следва $\mu_1 = 0$, което противоречи на твърдението $(\mu_0, \mu_1) \ne (0, 0)$. Нека $\mu_0 = 1$. Тогава $\mu_1 = -3y^2 \le 0$. Следователно $\mu_1 = 0$ и получаваме $(x, y) = (0, 0)$. Тази точка обаче не е точка на локален минимум, тъй като $f(0, -\varepsilon) < f(0, 0) = 0$ за всяко $\varepsilon > 0$.

Всъщност f е неограничена отдолу в X, тъй като $\lim\limits_{\lambda\to+\infty} f(0, -\lambda) = -\infty$.

б) За разлика от предишния случай множеството X е компактно. Следователно, тъй като f е непрекъсната функция, задачата има оптимално решение. Прилагаме теорема 5. Функцията на Лагранж е

$$L(x, y, \mu_0, \mu_1, \mu_2) = \mu_0(x^2 + y^3) + \mu_1(x^2 + y - 1) + \mu_2(-y - 1),$$

като $\mu_0 \geq 0$, $\mu_1 \geq 0$, $\mu_2 \geq 0$, $(\mu_0, \mu_1, \mu_2) \neq (0,0,0)$, откъдето получаваме условията:

$$1) \begin{cases} 2\mu_0 x + 2\mu_1 x = 0 \\ 3\mu_0 y^2 + \mu_1 - \mu_2 = 0 \end{cases}, \qquad 2) \begin{cases} \mu_1(x^2 + y - 1) = 0 \\ \mu_2(-y-1) = 0 \end{cases}, \qquad 3)\ \mu_1 \geq 0, \mu_2 \geq 0.$$

Разглеждаме първо случая $\mu_0 = 0$. От 1) следва $\mu_1 x = 0$, $\mu_1 = \mu_2$. Възможно е само $\mu_1 \neq 0$. Но тогава $x = 0$ и системата 2) става противоречива. Следователно $\mu_0 \neq 0$.

Нека $\mu_0 = 1$. От 1) и 3) следва $1 + \mu_1 \neq 0$. Следователно $x = 0$ и

$$3y^2 + \mu_1 - \mu_2 = 0, \qquad \mu_1(y-1) = 0, \qquad \mu_2(y+1) = 0.$$

Ако $\mu_2 \neq 0$, то $y = -1$. Следователно $\mu_1 = 0$, $3 - \mu_2 = 0$ и множителите на Лагранж са $\mu_0 = 1$, $\mu_1 = 0$, $\mu_2 = 3$.

Ако $\mu_1 \neq 0$, то $y = 1$, $\mu_2 = 0$, $3 + \mu_1 = 0$, което противоречи на 3).

И така, минимум се достига в точката $(0, -1)$ и $f(0, -1) = -1$.

3.3. Необходимо условие за локален екстремум за задачата с ограничения равенства и неравенства

Следващата теорема обединява теоремите на Лагранж и Джон за задачи с ограничения равенства и неравенства.

Теорема 6. Нека $\mathbf{x}^*$ е точка на локален минимум за функцията f при условия $g_i(\mathbf{x}) \leq 0$, $i = 1, \ldots, m$, $h_j(\mathbf{x}) = 0$, $j = 1, \ldots, s$. Да предположим, че функциите f, g_i, $i = 1, \ldots, m$, са диференцуеми в точката $\mathbf{x}^*$ и че функциите h_j, $j = 1, \ldots, s$, имат непрекъснати частни производни в околност на точката $\mathbf{x}^*$. Тогава съществуват реални числа $\mu_0, \mu_1, \ldots, \mu_m$, $\lambda_1, \ldots, \lambda_s$, не едновременно равни на нула, такива че са изпълнени условията

1) $\mu_0 f'(\mathbf{x}^*) + \sum\limits_{i=1}^{m} \mu_i g_i'(\mathbf{x}^*) + \sum\limits_{j=1}^{s} \lambda_j h_j'(\mathbf{x}^*) = \mathbf{0}$,

2) $\mu_i g_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$,

3) $\mu_i \geq 0$, $i = 0, \ldots, m$.

Ако векторите $g_i'(\mathbf{x}^*)$, $i \in I(\mathbf{x}^*)$, $h_j'(\mathbf{x}^*)$, $j = 1, \ldots, s$, са линейно независими, може да се приеме $\mu_0 = 1$.

Пример 4. Да се реши задачата

$$\begin{gathered} f(x_1, x_2) = x_1^2 + x_2^2 \to \min, \\ x_1^2 + x_2^2 \leq 5, \\ x_1 + 2x_2 = 4, \\ x_1 \geq 0,\ x_2 \geq 0. \end{gathered}$$

Решение. Очевидно допустимото множество е компакт и задачата има оптимално решение. Въвеждаме функциите $f(x_1, x_2) = x_1^2 + x_2^2$, $g_1(x_1, x_2) = x_1^2+x_2^2-5$, $g_2(x_1, x_2) = -x_1$, $g_3(x_1, x_2) = -x_2$, $h_1(x_1, x_2) = x_1+2x_2-4$. Функцията на Лагранж е

$$
\begin{aligned}
L(x_1, x_2, \mu_0, \mu_1, \mu_2, \mu_3, \lambda_1) = \mu_0(x_1^2 + x_2^2) + \mu_1(x_1^2 + x_2^2 - 5) \\
+ \mu_2(-x_1) + \mu_3(-x_2) + \lambda(x_1 + 2x_2 - 4),
\end{aligned}
$$

където $\mu_i \geq 0$, $i = 0, 1, 2, 3$, и $(\mu_0, \mu_1, \mu_2, \mu_3, \lambda_1) \neq \mathbf{0}$.

Условията на Джон са

$$2\mu_0 x_1 + 2\mu_1 x_1 - \mu_2 + \lambda_1 = 0, \tag{4}$$

$$2\mu_0 x_2 + 2\mu_1 x_2 - \mu_3 + 2\lambda_1 = 0, \tag{5}$$

$$\mu_1(x_1^2 + x_2^2 - 5) = 0, \tag{6}$$

$$\mu_2(-x_1) = 0, \tag{7}$$

$$\mu_3(-x_2) = 0, \tag{8}$$

$$\mu_0, \mu_1, \mu_2, \mu_3 \geq 0. \tag{9}$$

Добавяме и ограниченията на задачата

$$x_1^2 + x_2^2 \leq 5, \tag{10}$$

$$x_1 + 2x_2 = 4, \tag{11}$$

$$x_1 \geq 0,\ \ x_2 \geq 0. \tag{12}$$

Получаваме система от уравнения и неравенства. Забелязваме, че уравненията (6)–(8) представляват произведения от два множителя равни на нула. Комбинирайки по различен начин кой от множителите в отделните уравнения е равен на нула, разглеждаме следните пет случая:

1. Нека $\mu_1 = \mu_2 = \mu_3 = 0$. Тогава от (4), (5) и (11) получаваме

$$
\left|
\begin{aligned}
2\mu_0 x_1 + \lambda_1 &= 0 \\
2\mu_0 x_2 + 2\lambda_1 &= 0 \\
x_1 + 2x_2 &= 4.
\end{aligned}
\right.
$$

Ако $\mu_0 = 0$, то и $\lambda_1 = 0$. Така всички множители на Лагранж стават равни на нула, което е невъзможно. Да положим $\mu_0 = 1$. Тогава нещата се свеждат до следната линейна система от три уравнения с три неизвестни

$$
\left|
\begin{aligned}
2x_1 + \lambda_1 &= 0 \\
2x_2 + 2\lambda_1 &= 0 \\
x_1 + 2x_2 &= 4.
\end{aligned}
\right.
$$

Като я решим, получаваме $x_1 = \frac{4}{5}$, $x_2 = \frac{8}{5}$, $\lambda_1 = -\frac{8}{5}$.

2. Нека $\mu_1 = \mu_2 = x_2 = 0$. Тогава от (11) $x_1 = 4$. Получената точка $(4, 0)$ не е допустима, защото не удовлетворява (10).

3. Нека $\mu_1 = \mu_3 = x_1 = 0$. Тогава от (11) $x_2 = 2$. Получената точка $(0, 2)$ е допустима. Уравненията (4) и (5) стават

$$\left| \begin{aligned} &-\mu_2 + \lambda_1 = 0 \\ &4\mu_0 + 2\lambda_1 = 0, \end{aligned} \right.$$

откъдето $\mu_2 = \lambda_1 = -2\mu_0$. Така μ_0 и μ_2 не могат да бъдат едновременно неотрицателни, освен ако са равни на нула, но тогава всички множители на Лагранж се оказват равни на нула, което е противоречие.

4. Случаят $\mu_1 = x_1 = x_2 = 0$ е невъзможен, тъй като тогава не е изпълнено (11).

5. Нека сега вторият множител в (6) $x_1^2 + x_2^2 - 5 = 0$. Това заедно с (11) $x_1 + 2x_2 = 4$ дава система от две уравнения с две неизвестни. Тя има две решения $x_1 = -\frac{2}{5}$, $x_2 = \frac{11}{5}$ и $x_1 = 2$, $x_2 = 1$. Първото е недопустимо поради (12). Второто е допустимо и заедно със (7) и (8) дава $\mu_2 = \mu_3 = 0$. Заместваме в (4) и (5) и получаваме

$$\left| \begin{aligned} &4(\mu_0 + \mu_1) + \;\;\lambda_1 = 0 \\ &2(\mu_0 + \mu_1) + 2\lambda_1 = 0. \end{aligned} \right.$$

Съкращаваме на 2 второто уравнение и го изваждаме от първото. Така стигаме до $\mu_0 + \mu_1 = 0$. Отново тези два множителя на Лагранж не могат да бъдат едновременно неотрицателни, освен ако не са равни на нула. Но тогава всички множители на Лагранж стават равни на нула, което е противоречие.

Окончателно оптималното решение на задачата е $\mathbf{x}^* = \left(\frac{4}{5}, \frac{8}{5}\right)$. Стойността на целевата функция е равна на $\frac{16}{5}$.

Пример 5. Да се реши задачата

$$f(\mathbf{x}) = \sum_{i=1}^{n} \frac{c_i}{x_i} \to \min, \qquad \mathbf{x} \in X = \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^{n} a_i x_i = b, \; x_i \geq 0 \right\},$$

където a_i, b и c_i, $i = 1, \ldots, n$, са положителни реални числа.

Решение. Задачата има оптимално решение (защо?). Функцията на Лагранж е

$$L(\mathbf{x}, \boldsymbol{\mu}, \lambda_1) = \mu_0 \sum_{i=1}^{n} \frac{c_i}{x_i} + \sum_{i=1}^{n} \mu_i(-x_i) + \lambda_1 \left(\sum_{i=1}^{n} a_i x_i - b \right),$$

където $\boldsymbol{\mu} \in \mathbb{R}^{n+1}$, $\boldsymbol{\mu} \geq \mathbf{0}$, $\lambda_1 \in \mathbb{R}$ и не всички множители на Лагранж са равни на нула. Всяка точка $\mathbf{x}^*$ на локален минимум ще удовлетворява условията на теорема 6. Тогава

$$-\mu_0 \frac{c_i}{x_i^2} - \mu_i + \lambda_1 a_i = 0, \qquad \mu_i x_i = 0, \qquad \mu_i \geq 0, \qquad i = 0, 1, \dots, n.$$

Нека $\mu_0 = 0$. Тогава $\mu_i = \lambda_1 a_i$ и $\lambda_1 \neq 0$ (иначе всички множители на Лагранж са равни на нула). Заместваме a_i с $\dfrac{\mu_i}{\lambda_1}$ в ограничението равенството и получаваме $\dfrac{1}{\lambda_1} \sum\limits_{i=1}^{n} \mu_i x_i = b$. Тъй като $\mu_i x_i = 0$, следва $b = 0$ — противоречие. Да положим $\mu_0 = 1$. Тогава $x_i > 0$ и $\mu_i = 0$. Следователно $x_i = \sqrt{\dfrac{c_i}{\lambda_1 a_i}}$, откъдето $\lambda_1 = \dfrac{1}{b^2} \left(\sum\limits_{i=1}^{n} \sqrt{a_i c_i} \right)^2$ и накрая $x_i^* = \dfrac{b}{\sum\limits_{i=1}^{n} \sqrt{a_i c_i}} \sqrt{\dfrac{c_i}{a_i}}$ и $f(\mathbf{x}^*) = \dfrac{1}{b} \left(\sum\limits_{i=1}^{n} \sqrt{a_i c_i} \right)^2$.

4. Доказване на неравенства с използването на екстремални задачи

Тук ще покажем как могат да бъдат доказвани неравенства, като се използват оптималните решения на екстремални задачи. Започваме с едно класическо неравенство.

Пример 6. Да се реши задачата

$$\min f(\mathbf{x}) = \sum_{j=1}^{n} x_j, \qquad \prod_{j=1}^{n} x_j = c, \quad x_j \geq 0, \quad j = 1, \dots, n,$$

където c е положителна константа. Като следствие от решението на задачата да се получи неравенството между средното аритметично и средното геометрично на n неотрицателни числа $a_1, \dots, a_n$, а именно

$$\frac{1}{n} \sum_{j=1}^{n} a_j \geq \left(\prod_{j=1}^{n} a_j \right)^{1/n}.$$

Решение. Тъй като x_j са неотрицателни, то произволна редица от допустими точки, за която $\lim\limits_{k\to\infty} \|\mathbf{x}_k\| = +\infty$, се характеризира с това, че за поне една координата i на векторите $\mathbf{x}_k$ съответната редица x_{i_k} расте неограничено. Тогава целевата функция $f(\mathbf{x})$ е коерцитивна. Според следствие 1 задачата има оптимално решение.

Функцията на Лагранж за дадената задача е

$$L(\mathbf{x}, \boldsymbol{\mu}, \lambda) = \mu_0 \sum_{j=1}^{n} x_j + \sum_{j=1}^{n} \mu_j(-x_j) + \lambda \left(\prod_{j=1}^{n} x_j - c \right), \qquad \boldsymbol{\mu} \in \mathbb{R}_+^{n+1}$$

и множителите на Лагранж не са едновременно нули. Като приравним частните производни на функцията на Лагранж по отношение на променливите x_i на нула и напишем условията за допълнителност, получаваме

$$\left| \begin{array}{l} \mu_0 - \mu_i + \lambda \dfrac{c}{x_i} = 0, \quad i = 1, \ldots, n, \\ \mu_i x_i = 0, \quad i = 1, \ldots, n. \end{array} \right.$$

Тъй като c е положително, никое x_j не е равно на нула. Тогава от условията за допълнителност $\mu_i = 0$, $i = 1, \ldots, n$. Сега ако допуснем, че $\mu_0 = 0$, излиза и $\lambda = 0$, т. е. всички множители на Лагранж са равни на нула, което е противоречие. Полагаме $\mu_0 = 1$. Така $x_i = -\lambda c$, $i = 1, \ldots, n$, т. е. всички x_i са равни. От ограничението равенство намираме $x_i^* = \sqrt[n]{c}$, $i = 1, \ldots, n$. Множителят на Лагранж $\lambda = -c^{1/n-1}$. Минималната стойност на целевата функция е $n\sqrt[n]{c}$.

Сега да докажем неравенството. Ако някое от числата a_j е равно на нула, неравенството е тривиално. Тогава да разгледаме случая, когато никое от тях не е равно на нула. В този случай полагаме $c = \prod\limits_{j=1}^{n} a_j > 0$. Така точката $\mathbf{a} = (a_1, \ldots, a_n)$ е допустима за разгледаната екстремална задача и следователно

$$\frac{1}{n} \sum_{j=1}^{n} a_j = \frac{1}{n} f(\mathbf{a}) \geq \frac{1}{n} f(\mathbf{x}^*) = \sqrt[n]{c} = \left(\prod_{j=1}^{n} a_j \right)^{1/n}.$$

Пример 7 (неравенство на Hölder)**.** Нека $p > 1$, $q > 1$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Да се докаже, че ако a_i, b_i, $i = 1, \ldots, n$, са реални числа, то

$$\sum_{i=1}^{n} |a_i b_i| \leq \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |b_i|^q \right)^{1/q}.$$

Решение. Разглеждаме задачата

$$\min_{\mathbf{x}\in X} f(\mathbf{x}) = \sum_{i=1}^{n} x_i^p, \quad X = \left\{\sum_{i=1}^{n} b_i x_i = c,\ x_i \geq 0,\ i = 1,\dots,n\right\},$$

където $b_i > 0$, $i = 1,\dots,n$, $c > 0$ са константи. Множеството X е компакт, а f е непрекъсната. По теоремата на Вайерщрас съществува точка на минимум $\mathbf{x}^* \in X$. Според теорема 6 съществуват числа $\mu_i \geq 0$, $i = 0,1,\dots,n$, λ, не всичките нули, такива че

$$(13) \qquad \left|\begin{array}{l} \mu_0 f'(\mathbf{x}^*) + \sum\limits_{i=1}^{n} \mu_i g_i(\mathbf{x}^*) + \lambda h(\mathbf{x}^*) = \mathbf{0}, \\ u_i g_i(\mathbf{x}^*) = 0,\ i = 1,\dots,n, \end{array}\right.$$

където $g_i(\mathbf{x}) = -x_i$, $i = 1,\dots,n$, $h(\mathbf{x}) = \sum\limits_{i=1}^{n} b_i x_i = c$. От (13) получаваме

$$(14) \qquad \left|\begin{array}{l} \mu_0 p x_i^{*\,p-1} - \mu_i + \lambda b_i = 0,\ i = 1,\dots,n, \\ \mu_i x_i^* = 0,\ i = 1,\dots,n. \end{array}\right.$$

Да допуснем, че $\mu_0 = 0$. Тогава обезателно $\mu_i > 0$, $i = 1,\dots,n$, защото в противен случай ако $\mu_{i_0} = 0$ за някое $i_0 = 1,\dots,n$, то от първите уравнения в (14) за $i = i_0$ получаваме $\lambda = 0$, а след това от останалите уравнения и $\mu_i = 0$ за $i = 1,\dots,n$, което е противоречие. Сега от условията за допълнителност следва, че $x_i^* = 0$, $i = 1,\dots,n$. Но нулевият вектор не е допустим. Следователно $\mu_0 \neq 0$.

Полагаме $\mu_0 = 1$. От (14) получаваме

$$x_i^* = \left(-\frac{\lambda}{p} b_i\right)^{\frac{1}{p-1}} \quad \text{или} \quad x_i^* = 0.$$

Означаваме $I = \{i : x_i^* \neq 0\}$. Тогава

$$(15) \qquad \begin{aligned} c = \sum_{i=1}^{n} b_i x_i^* &= \sum_{i\in I} b_i \left(-\frac{\lambda}{p} b_i\right)^{\frac{1}{p-1}} = \left(-\frac{\lambda}{p}\right)^{\frac{1}{p-1}} \sum_{i\in I} b_i^{1+\frac{1}{p-1}} \\ &= \left(-\frac{\lambda}{p}\right)^{\frac{1}{p-1}} \sum_{i\in I} b_i^q. \end{aligned}$$

От друга страна

$$(16) \qquad \left(\sum_{i=1}^{n} x_i^{*p}\right)^{1/p} = \left(\sum_{i\in I} \left(-\frac{\lambda}{p} b_i\right)^{\frac{1}{p-1}}\right)^{1/p} = \left(-\frac{\lambda}{p}\right)^{\frac{1}{p-1}} \left(\sum_{i\in I} b_i^q\right)^{1/p}.$$

Сега да докажем неравенството. Без ограничение на общността можем да смятаме, че числата a_i и b_i, $i = 1, \dots, n$, са неотрицателни. Нека $c = \sum_{i=1}^{n} a_i b_i$. Тогава векторът $(a_1, \dots, a_n)$ е допустим и

$$\begin{aligned}
\left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} b_i^q\right)^{1/q} &\geq \left(\sum_{i=1}^{n} x_i^{*p}\right)^{1/p} \left(\sum_{i\in I} b_i^q\right)^{1/q} \\
&\overset{(16)}{=} \left(-\frac{\lambda}{p}\right)^{\frac{1}{p-1}} \left(\sum_{i\in I} b_i^p\right)^{1/p} \left(\sum_{i\in I} b_i^q\right)^{1/q} \\
&= \left(-\frac{\lambda}{p}\right)^{\frac{1}{p-1}} \sum_{i\in I} b_i^q \overset{(15)}{=} c = \sum_{i=1}^{\nu} a_i b_i.
\end{aligned}$$

Задачи

1. Да се покаже (с пример), че:

а) условието $f'(\mathbf{x}^*) = 0$ не е достатъчно, за да бъде $\mathbf{x}^*$ точка на локален екстремум на f в $\mathbb{R}^n$;

б) условието $f'(\mathbf{x}^*) = 0$ и $f''(\mathbf{x}^*)$ да е положително дефинитна не е необходимо, за да бъде $\mathbf{x}^*$ точка на локален екстремум на f в $\mathbb{R}^n$.

2. Да се състави задача, за която се удовлетворяват условията на теоремата на Лагранж с $\mu_0 = 0$.

3. Да се намерят локалните и глобалните екстремуми на следните функции при съответните условия:

3.1. $xy + \dfrac{50}{x} + \dfrac{20}{y} \to \mathrm{extr}$;

3.2. $x^2 - y^2 - 4x + 6y \to \mathrm{extr}$;

3.3. $x^2 + y^2 \to \mathrm{extr}$ при ограничение $3x + 4y = 1$;

3.4. $xy^2z^3 \to \mathrm{extr}$ при ограничение $x + y + z = 1$;

3.5. $xyz \to \mathrm{extr}$ при ограничение $x^2 + y^2 + z^2 \leq 1$;

3.6. $x^2 + y^2 \to \min$ при ограничение $x^4 + y^4 = 1$;

3.7. $e^{xy} \to \mathrm{extr}$ при ограничение $x + y = 1$.

4. Да се докаже, че минималната стойност на функцията $f(\mathbf{x}) = \mathbf{xAx}$ в единичната сфера $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$, където $\mathbf{A}$ е симетрична матрица, е най-малката собствена стойност на $\mathbf{A}$.

5. Да се намерят екстремумите в следните задачи:

5.1. $\sum\limits_{i=1}^{n} x_i^2 \to \text{extr}, \ \sum\limits_{i=1}^{n} x_i^4 \leq 1;$

5.2. $\sum\limits_{i=1}^{n} x_i^4 \to \text{extr}, \ \sum\limits_{i=1}^{n} x_i^2 \leq 1.$

6. Да се намери правоъгълен триъгълник с най-голямо лице, ако сумата от дължините на катетите му е дадено число d.

7. Да се впише в кръг триъгълник с максимална сума от квадратите на страните.

8. Да се намерят стойностите на x_1, x_2 (и x_3), които са оптимални решения на следните задачи:

8.1. $\displaystyle\int_{-1}^{1} (t^2 + x_1 t + x_2)^2 \, dt \to \inf;$

8.2. $\int_{-1}^{1} (t^3 + x_1 t^2 + x_2 t + x_3)^2 \, dt \to \inf.$

Подинтегралните функции с коефициенти, при които се достига инфимумът, се наричат *полиноми на Льожандър* от втора и трета степен съответно.

9. Разглеждаме общата задача на нелинейното оптимиране (1) с ограничения: $X = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0\}$. Нека $L(\mathbf{x}, \boldsymbol{\mu})$ е съответната функция на Лагранж и нека $\psi(\boldsymbol{\mu}) = \inf\{L(\mathbf{x}, \boldsymbol{\mu}) : \mathbf{x} \in \mathbb{R}^n\}$. Да се докаже, че за всяка точка $\mathbf{x} \in X$ и за всяко $\boldsymbol{\mu} \geq 0$ е изпълнено $f(\mathbf{x}) \geq \psi(\boldsymbol{\mu})$.

10. Ако $X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{y}, \ \mathbf{y} \in \mathbb{R}^m, \ \mathbf{y} \geq 0\}$, където $\mathbf{A}$ е матрица $m \times n$ и $f(\mathbf{x}) = \|\mathbf{x}\|$, да се докаже, че задачата $\min\{f(\mathbf{x}) : \mathbf{x} \in X\}$ има оптимално решение.

Отговори и решения

3.1. Локален минимум в $(5, 2)$; $\min = -\infty$; $\max = \infty$.

3.2. $\max = \infty$; $\min = -\infty$, няма локален екстремум.

3.3. Глобален минимум в $(\frac{3}{25}, \frac{4}{25})$; $\min = \frac{1}{25}$; $\max = \infty$.

3.4. Локален максимум в $(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$ и $(t, 0, 1-t)$ за $t > 1$, $t < 0$; локален минимум в $(t, 0, 1-t)$ за $0 < t < 1$; $\min = -\infty$; $\max = \infty$.

3.5. Ако $\sigma = -\frac{1}{\sqrt{3}}$,

- глобален минимум в $(-\sigma, -\sigma, -\sigma)$, $(\sigma, \sigma, -\sigma)$, $(\sigma, -\sigma, \sigma)$, $(-\sigma, \sigma, \sigma)$;
- глобален максимум в (σ, σ, σ), $(\sigma, -\sigma, -\sigma)$, $(-\sigma, \sigma, -\sigma)$ и $(-\sigma, -\sigma, \sigma)$;
- $\max = -\min = \frac{\sigma}{3}$.

3.6. Глобален минимум в $(\pm 1, 0)$, $(0, \pm 1)$; $\min = 1$.

3.7. Глобален максимум в $(\frac{1}{2}, \frac{1}{2})$; $\max = e^{\frac{1}{4}}$, $\inf = 0$ (не се достига).

5.1. Глобален минимум в $(0, \dots, 0)$, $\min = 0$;
глобален максимум в $(\pm n^{-\frac{1}{4}}, \dots, \pm n^{-\frac{1}{4}})$, $\max = \sqrt{n}$.

5.2. Глобален минимум в $(0, \dots, 0)$, $\min = 0$;
глобален максимум в $(\pm 1, 0, \dots, 0)$, $\dots$, $(0, \dots, \pm 1)$, $\max = 1$.

8.1. $t^2 - \frac{1}{3}$.

8.2. $t^3 - \frac{3}{5}t$.

10. **Упътване.** Покажете, че X е затворено множество.