Петър Кендеров
Георги Христов
Асен Дончев

Математическо оптимиране

София • 1989

В учебника „Математическо оптимиране“ са изложени теоретичните основи за решаване на задачи за минимум и максимум. Разгледани са свойствата на системите линейни неравенства и равенства, изпъкналите множества и изпъкналите функции и от тях са получени условия за екстремум. Отделено е специално внимание на линейното, хиперболичното и квадратичното оптимиране. Учебникът е предназначен за студенти от факултета по математика и информатика при Софийския университет „Климент Охридски“. За четенето му са необходими познания в рамките на основните курсове по линейна алгебра и математически анализ. Може да бъде използван от студенти в икономически и технически висши учебни заведения. Представлява интерес и за специалисти, чиято дейност е свързана с практически оптимизационни задачи.

Одобрено за печат от Министерството на културата, науката и просветата с решение № 131–00–85 от 3.IV.1989 г.

Съдържание

Използвани означения

$\mathbb{R}^n$	— n-мерно евклидово пространство
$\|x\|$	— евклидова норма на x
$\langle x, y\rangle$	— скаларно произведение на вектори
$\mathbb{R}^n_+$	— множеството от всички вектори в $\mathbb{R}^n$ с неотрицателни компоненти
A^{T}	— матрицата, транспонирана на матрицата A
$\overline{xy}$	— отсечка с краища векторите x и y
$\emptyset$	— празно множество
$\overline{X}$	— затворена обвивка на X
$\operatorname{int} X$	— вътрешност на X
∂X	— множеството от контурите точки на X
$\operatorname{co} X$	— изпъкнала обвивка на X
$\operatorname{aff} X$	— афинна обвивка на X
$\dim X$	— размерност на X
$\operatorname{ri} X$	— относителна вътрешност на X
$\mathrm{r}\partial X$	— множеството от всички истински контурни точки на X
$\widehat{X}$	— множеството от крайни точки на X
$K(X)$	— рецесивен конус на X
$S(X)$	— рецесивно подпространство на X
X^*	— спрегнат конус на X
$d(y, X)$	— разстоянието от точката y до множеството X
$f : X \to Y$	— функция, дефинирана в множеството X със стойности в множеството Y
$L(f, \alpha)$	— множеството на Лебег за функцията f
$\operatorname{epi} f$	— надграфика на функцията f
$\dfrac{\partial f}{\partial s}(x)$	— производна на функцията f в точката x по посока s
$f'(x)$	— производна (градиент) на функцията f в точката x
$\partial_f(x)$	— субдиференциал на функцията f в точката x
$l(x)$	— линейна функция
$H(x)$	— дробно-линейна функция
$Q(x)$	— квадратична функция

Предговор

В това учебно пособие са изложени теоретичните основи на математическото оптимиране. То е написано въз основа на лекции, четени от авторите пред студенти от факултета по математика и информатика при Софийския университет „Кл. Охридски“.

Математическото оптимиране е дял от математиката, в който се изучават задачи за намиране на минимум или максимум на функция. Оптимизационните задачи винаги са били актуални, а през последните десетилетия интересът към тях нарастна във връзка с изискванията на научно-технологическия прогрес и особено поради навлизането на компютрите в обществената практика.

В книгата се разглеждат задачи за намиране на минимум па функция на краен брой променливи в множество, дефинирано с краен брой равенства и (или) неравенства. Централно място е отделено на математическите идеи и резултати, които са основа за извеждане на различни необходими условия за екстремум.

Първата глава въвежда читателя в основните проблеми и резултати на математическото оптимиране. С помощта на сравнително несложен апарат — системи от линейни уравнения и неравенства — са получени необходими условия за екстремум и е обоснована двойнствеността в линейното оптимиране. Въведено е понятието седлова точка и са доказани достатъчни условия за екстремум. Изложен е подходът за решаване на оптимизационни задачи, известен като „метод за множителите на Лагранж“.

Изпъкналите множества и свързаните с тях понятия, идеи и методи са съществена част от апарата на математическото оптимиране. Основните дефиниции и резултати от теорията на изпъкналите множества и техните структурни свойства са събрани във втора и трета глава. В четвърта глава са изложени някои свойства на изпъкналите функции, а в пета глава е доказана (в различни варианти) теоремата на Кун— Такър — централен резултат в математическото оптимиране. Шеста глава е посветена на три специални и важни за практиката оптимизационни задачи: линейната, хиперболичната и квадратичната задача.

Поради ограничения обем в книгата не намериха място редица важни дялове на математическото оптимиране като дискретно оптимиране, стохастично оптимиране и многокритерийно оптимиране и динамично оптимиране. Все още предстои издаване на материал, предназначен за по-широка читателска аудитория и посветен на числените методи за оптимизация.

Книгата е предназначена за обучение на студенти съобразно учебните програми по математика и информатика. Може да бъде използвана и от студенти от икономически и технически специалности, а също представлява интерес и за специалисти, чиято дейност е свързана с приложения на математиката. За четенето ѝ са необходими познания в рамките на основните курсове по линейна алгебра и математически анализ.

Изложените в книгата математически факти са наричани леми (помощни твърдения), твърдения и теореми, последните от които съдържат основни за съответната тема резултати. Включени са решени примери и упражнения, илюстриращи материала, а също и задачи за самостоятелна подготовка.

Учебникът е набран на персонален компютър ПРАВЕЦ-16 и отпечатан с помощта на текстоформатиращата програма ФОРМАТЕКСТ [1.6] графичният шрифтов филтър LETTRIX на матричен принтер STAR SR-15, за което изказваме искрена благодарност на Антоанета Джераси, Евгени Белогай, Емил Стоянов и Михаела Иванова.

От авторите

София, април 1989 г.

Глава 1

Математическо оптимиране

Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е реалнозначна функция, дефинирана в $\mathbb{R}^n$, и нека X е подмножество на $\mathbb{R}^n$. Да разгледаме следната задача: Да се намери точка $x_0 \in X$, за която неравенството $f(x) \geq f(x_0)$ е в сила при всеки избор на $x \in X$.

Такава точка $x_0 \in X$ ще наричаме решение на задачата за минимум: $\min\{f(x) : x \in X\}$.

Прости примери показват, че дори функцията f да е ограничена отдолу върху непразното множество X, не винаги съществува най-малко число сред числата $\{f(x) : x \in X\}$. В такъв случай ще смятаме, че минимизационната задача е решена, ако успеем да намерим число $\alpha = \inf\{f(x) : x \in X\}$ и евентуално редица $\{x_k\}_{k=1}^{\infty}$, за която

$$\lim_{k\to\infty} f\{x_k\} = \alpha.$$

Когато X е празното множество, по определение ще считаме, че $\inf\{f(x) : x \in X\} = +\infty$ (аналогично $\sup\{f(x) : x \in X\} = -\infty$).

Често се разглежда и задачата за максимум: Да се намери точка $x_0 \in X$, за която $f(x) \leq f(x_0)$ при всеки избор на точката $x \in X$.

Понякога класът на всички минимизационни и максимизационни задачи се нарича още клас на екстремалните задачи или клас на оптимизационните задачи, т. е. задачи за намиране на екстремум (оптимум).

Като вземем предвид равенството $-\sup\{f(x) : x \in X\} = \inf\{-f(x) : x \in X\}$, лесно се вижда, че всяка максимизационна задача се свежда до минимизационна и обратното. За простота по-нататък ще разглеждаме главно минимизационни задачи. Прието е минимизационните задачи да се записват по следния начин:

$$\min_{x\in X} f(x), \quad \min\{f(x) : x \in X\}, \quad f(x) \underset{x\in X}{\longrightarrow} \min .$$

В § 1 на на тази глава ще опишем един достатъчно общ и често използван метод, чрез които се установява съществуване на решение на дадена оптимизационна задача. За съжаление този метод не ни дава ефективни средства за намиране на решението, чието съществуване вече е установено. Подобно е положението и при решаване на (системи от) уравнения. По един начин се установява съществуването на решение, а чрез други методи се характеризират тези решения.

В § 2 на тази глава са дадени техническите средства (система от линейни уравнения и неравенства), чрез които в следващите параграфи се доказват необходими условия, за да бъде една точка x_0 решение на съответната минимизационна задача. Множеството X се задава като съвкупност от решения на система от уравнения и неравенства. Доказаните твърдения обосновават използването на известния метод на множителите на Лагранж. Чрез функцията на Лагранж и чрез понятието седлова точка в § 6 е формулирано достатъчно условие, за да бъде една точка $x_0 \in X$ решение на дадена минимизационна задача.

1. Съществуване на решение на оптимизационни задачи

Един от основните въпроси в теорията на оптимизационните задачи е дали съществува решение. Задачата

$$\min\{f(x) : x \in X\}$$

ще има решение тогава (и само тогава), когато съществува точка $x_0 \in X$ такава, че

$$f(x_0) = \inf\{f(x) : x \in X\}.$$

Условия, гарантиращи съществуване на екстремум, се дават от известната теорема на Вайерщрас: Всяка непрекъсната функция, дефинирана в компактно множество X, достига както инфимума, така и супремума си в X.

По-нататък ще разгледаме едно обобщение на теоремата на Вайерщрас, което значително разширява сферата на нейната приложимост.

Определение 1. Функцията $f : X \to \mathbb{R}^1$ се нарича *полунепрекъсната отдолу* (*отгоре*) в точката $x_0 \in X$, ако за всяко $\varepsilon > 0$ съществува $\delta > 0$, зависещо от ε и такова, че всяко $x \in X$, за което $\|x - x_0\| < \delta$ е изпълнено неравенството

$$f(x) > f(x_0) - \varepsilon \qquad (\text{съответно } f(x) < f(x_0) + \varepsilon).$$

Определение 2. Функцията $f : X \to \mathbb{R}^1$ се нарича *полунепрекъсната отдолу* (*отгоре*) в точка $x_0 \in X$, ако за всяка редица от точки $\{x_k\}_{k=1}^{\infty} \to x_0$, $x_k \in X$, е изпълнено неравенството $f(x_0) \le \varliminf\limits_{k\to\infty} f(x_k)$ (съответно $f(x_0) \ge \varlimsup\limits_{k\to\infty} f(x_k)$).

Естествено е при две определения за едно и също нещо да се покаже тяхната еквивалентност.

И така нека функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу в точката $x_0 \in X$ според определение 1. Избираме произволна редица $\{x_k\}_{k=1}^{\infty} \to x_0$, $\{x_k\}_{k=1}^{\infty} \subset X$. Следователно за всяко $\varepsilon > 0$ съществува цяло положително $N(\varepsilon)$ такова, че при $k > N(\varepsilon)$ е изпълнено $\|x_k - x_0\| < \delta(\varepsilon)$. Оттук следват неравенствата

$$
\begin{aligned}
& f(x_k) > f(x_0) - \varepsilon \text{ при } k > N(\varepsilon), \\
& \varliminf_{k\to\infty} f(x_k) \geq f(x_0) - \varepsilon \text{ за всяко } \varepsilon > 0, \\
& \varliminf_{k\to\infty} f(x_k) \geq f(x_0),
\end{aligned}
$$

т. е. функцията f е полунепрекъсната отдолу в точката $x_0 \in X$ съгласно определение 2.

Нека сега функцията е полунепрекъсната отдолу в точката $x_0 \in X$ според определение 2. Допускаме, че не са изпълнени условията от определение 1. Тогава ще съществува такова $\varepsilon > 0$, че за всяко цяло положително число k има поне една точка $x_k \in X$, за която

$$\|x_k - x_0\| < \frac{1}{k}, \quad f(x_k) < f(x_0) - \varepsilon.$$

Но $\{x_k\}_{k=1}^{\infty} \to x_0$, а $\varliminf_{k\to\infty} f(x_k) \leq f(x_0) - \varepsilon < f(x_0)$.

Стигаме до противоречие с определение 2. С това еквивалентността на двете определения е доказана.

На фиг. 1 са дадени графики на функции: а) полунепрекъсната отгоре, б) полунепрекъсната отдолу, в) функцията не е полунепрекъсната.

Ще казваме, че функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу (отгоре) в множеството $A \subset X$, ако е полунепрекъсната отдолу (отгоре) във всяка негова точка.

Упражнение 1. Докажете, че ако функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу (отгоре) в X и $\alpha \geq 0$, то функцията $g = \alpha f$ е полунепрекъсната отдолу (отгоре) в X.

Упражнение 2. Докажете, че ако функциите $f : X \to \mathbb{R}^1$ и $g : X \to \mathbb{R}^1$ са полунепрекъснати отдолу (отгоре) в X, то функцията $h = f + g$ е полунепрекъсната отдолу (отгоре) в X.

Упражнение 3. Докажете, че ако функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу в X, то функцията $-f$ е полунепрекъсната отгоре в X.

Теорема 1 (Вайерщрас)**.** Ако $X \neq \emptyset$ е компактно множество и функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу (отгоре) в X, тя достига инфимума (супремума) си в X.

Доказателство. Нека $\inf\limits_{x\in X} f(x) = \alpha$, $\alpha \geq -\infty$, и $\{x_k\}_{k=1}^{\infty} \subset X$ е редица от точки, за които $\lim\limits_{k\to\infty} f(x_k) = \alpha$ (такава редица е прието да се нарича *минимизираща редица* за функцията f в множеството X).

Редицата $\{x_k\}_{k=1}^{\infty}$ е ограничена (X е компактно). Следователно от нея можем да изберем сходяща подредица $\{x_{k_i}\}_{i=1}^{\infty} \to x_0$. Очевидно $x_0 \in X$ (X е затворено). От полунепрекъснатостта отдолу на функцията f в X следва $f(x_0) \leq \varliminf\limits_{i\to\infty} f(x_{k_i}) = \alpha$. Следователно $\alpha > -\infty$, т. е. f е ограничена отдолу в X. Но $f(x_0) \geq \alpha$, откъдето получаваме $f(x_0) = \alpha$.

За полунепрекъснати отгоре функции доказателството е аналогично. □

Теоремата остава в сила, ако заменим условието за компактност на X с условието „X е затворено и съществува ограничена минимизираща (максимизираща) редица за функцията f в множеството X“.

Следствие 1. Ако $X \neq \emptyset$ е затворено множество в $\mathbb{R}^n$, функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу в X и за всяка редица $\{x_k\}_{k=1}^{\infty} \subset X$, за която $\lim\limits_{k\to\infty} \|x_k\| = +\infty$, имаме $\lim\limits_{k\to\infty} f(x_k) = +\infty$, то f достига инфимума си в X.

Доказателство. От направените предположения следва, че всяка минимизираща редица за функцията f в X е ограничена, а оттук и верността на твърдението. □

Пример 1. Функцията $f(x) = e^x$ не достига инфимума си в $\mathbb{R}^1$. В случая условието $\lim\limits_{k\to\infty} f(x_k) = +\infty$, за всяка редица $\{x_k\}_{k=1}^{\infty} \subset \mathbb{R}^1$, $\{\|x_k\|\}_{k=1}^{\infty} \to +\infty$ (наричано често условие на коерцитивност), не е изпълнено.

Нека $f : X \to \mathbb{R}^1$ и $\alpha \in \mathbb{R}^1$. Множеството от вида $L(f,\alpha) = \{x \in X : f(x) \leq \alpha\}$ се нарича *множество на Лебег.*

Теорема 2. Функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу в X тогава и само тогава, когато $L(f,\alpha)$ е затворено множество в X (относно топологията в X) за всяко $\alpha \in \mathbb{R}^1$.

Доказателство. Нека f е полунепрекъсната отдолу в X и α е произволно избрано реално число. Разглеждаме $L(f,\alpha)$. Ако $L(f,\alpha) = \emptyset$, твърдението е доказано. Нека $L(f,\alpha) \neq \emptyset$. Избираме произволна редица $\{x_k\}_{k=1}^{\infty} \subset L(f,\alpha)$ и $\{x_k\}_{k=1}^{\infty} \to x_0$, $x_0 \in X$. От условието $x_k \in X$, $k = 1, 2, \dots$, от полунепрекъснатостта на f и от $f(x_k) \leq \alpha$ имаме $f(x_0) \leq \varliminf\limits_{k\to\infty} f(x_k) \leq \alpha$, т. е. $x_0 \in L(f,\alpha)$. Следователно $L(f,\alpha)$ е затворено.

Нека $L(f,\alpha)$ е затворено множество в X за всяко число α. Да вземем произволна точка $x_0 \in X$ и произволно число $\varepsilon > 0$. Множеството $L(f, f(x_0) - \varepsilon)$ е затворено и не съдържа точката x_0. Следователно съществува такова число $\delta > 0$, че ако $x \in X$ и $\|x - x_0\| < \delta$, то $x \notin L(f, f(x_0) - \varepsilon)$, т. е. за всяко $x \in X$

и $\|x - x_0\| < \delta$ имаме $f(x) > f(x_0) - \varepsilon$. Следователно f е полунепрекъсната отдолу в X. □

Забележка. Ако X е затворено множество, то $L(f, \alpha)$ е затворено както в X, така и в $\mathbb{R}^n$.

Следствие 2. Ако $X \neq \emptyset$ е затворено подмножество на $\mathbb{R}^n$, функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу в X и за някое $x^* \in X$ множеството $L(f, f(x^*))$ е ограничено, то f достига инфимума си в X.

Доказателство. Всяка минимизираща редица от известно място нататък се съдържа в $L(f, f(x^*))$ и следователно е ограничена. □

За произволна функция $f : X \to \mathbb{R}^1$ множеството

$$\operatorname{epi} f = \{(x, \alpha) : x \in X,\ \alpha \geq f(x)\} \subset \mathbb{R}^{n+1}$$

се нарича *надграфика* на функцията f. На фиг. 2 е илюстрирана надграфиката на функцията $f(x) = x^2$, дефинирана в $X = \{x \in \mathbb{R}^1 : x \geq 0\}$.

Теорема 3. Функцията $f : X \to \mathbb{R}^1$ е полунепрекъсната отдолу в X тогава и само тогава, когато $\operatorname{epi} f$ е затворено множество в $X \times \mathbb{R}^1$.

Доказателство. Нека f е полунепрекъсната отдолу в X. Ако $\{(x_k, \alpha_k)\}_{k=1}^{\infty} \subset \operatorname{epi} f$ и $\{(x_k, \alpha_k)\}_{k=1}^{\infty} \to (x_0, \alpha_0)$, то $x_k \in X$ и $\alpha_k \geq f(x_k)$, $k = 1, 2, \ldots$ Тъй като $\operatorname{epi} f$ трябва да е затворено в $X \times \mathbb{R}^1$, считаме, че $x_0 \in X$. Като вземем предвид и полунепрекъснатостта отдолу на f, ще получим

$$\alpha_0 \geq \varliminf_{k \to \infty} f(x_k) \geq f(x_0),$$

т. е. $(x_0, \alpha_0) \in \operatorname{epi} f$. Следователно $\operatorname{epi} f$ е затворено множество в $X \times \mathbb{R}^1$.

Нека сега $\operatorname{epi} f$ е затворено множество в $X \times \mathbb{R}^1$. Вземаме произволна редица $\{x_k\}_{k=1}^{\infty} \subset X$, $\{x_k\}_{k=1}^{\infty} \to X_0$ в X. Нека подредицата $\{x_{k_i}\}_{i=1}^{\infty}$ е такава, че $\varliminf_{i \to \infty} f(x_{k_i}) = \varliminf_{k \to \infty} f(x_k) = \alpha_0$. Очевидно

$$\{(x_{k_i}, f(x_{k_i}))\}_{i=1}^{\infty} \to (x_0, \alpha_0)$$

и

$$\{(x_{k_i}, f(x_{k_i}))\}_{i=1}^{\infty} \subset \operatorname{epi} f.$$

От затвореността на $\operatorname{epi} f$ получаваме $(x_0, \alpha_0) \in \operatorname{epi} f$, т. е. $f(x_0) \leq \alpha_0 = \varliminf_{k \to \infty} f(x_k)$. Следователно f е полунепрекъсната отдолу в X. □

Пример 2. Функцията

$$f_1(x) = \begin{cases} 1 & \text{за } x \leq 0, \\ 0 & \text{за } x > 0 \end{cases}$$

не е полунепрекъсната отдолу и нейната надграфика не е затворено множество (фиг. 3, а).

Ако обаче вземем

$$f_2(x) = \begin{cases} 1 & \text{за } x < 0, \\ 0 & \text{за } x \geq 0, \end{cases}$$

ще получим, че epi f_2 е затворено множество (фиг. 3, б).

Забележка. Ако X е затворено, то epi f е затворено както в $X \times \mathbb{R}^1$, така и в $\mathbb{R}^{n+1}$.

Задача 1. Нека $x \in \mathbb{R}^n$ и $X \subset \mathbb{R}^n$ е затворено множество. Докажете, че задачата

$$\|x - y\| \to \min, \quad y \in X,$$

има решение.

Задача 2. Нека X е множество в $\mathbb{R}^1$, което не е компактно. Намерете непрекъсната функция $f : X \to \mathbb{R}^1$, която не достига минимума си в X.

2. Еднородни системи линейни уравнения и системи линейни неравенства. Теорема на Фаркаш

В този параграф ще докажем някои резултати за системи линейни уравнения и линейни неравенства. Те са удобен инструмент, който ще се използва в по-нататъшните разглеждания.

Ако $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ са елементи от $\mathbb{R}^n$, с $\langle x, y \rangle$ ще означаваме скаларното им произведение:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

От линейната алгебра е известно, че ако a_j, $j = 1, 2, \dots, n$, са елементи на $\mathbb{R}^n$, то системата линейна уравнения $\langle a_j, x \rangle = b_j$, $j = 1, 2, \dots, n$, има решение при всеки избор на реалните $b_1, b_2, \dots, b_n$ тогава и само тогава, когато векторите a_j, $j = 1, 2, \dots, n$, са линейно независими (детерминантата на матрицата от координатите на a_j, $j = 1, 2, \dots, n$, е различна от нула). Решението на горната система може да се намери чрез формулите на Крамер. Известно е също така, че всяко множество a_i, $i = 1, 2, \dots, p < n$, от линейно независими вектори в $\mathbb{R}^n$ може да се допълни до система от n линейно независими вектори в $\mathbb{R}^n$. В частност, ако при някой избор на числата $b_1, b_2, \dots, b_p$ системата $\langle a_i, x \rangle = b_i$, $i = 1, 2, \dots, p$, няма решение, то векторите a_i, $i = 1, 2, \dots, p$, са линейно зависими.

Твърдение 1. Нека a_i, $i = 0, 1, \dots, p$, са такива вектори в $\mathbb{R}^n$, че $\langle a_0, x\rangle = 0$ всеки път, когато $\langle a_i, x\rangle = 0$ за $i = 1, 2, \dots, p$. Тогава съществуват такива числа λ_i, $i = 1, 2, \dots, p$, че

$$a_0 = \sum_{i=1}^{p} \lambda_i a_i.$$

Твърдението може да се изкаже още така: „Ако от $\langle a_i, x\rangle = 0$ за $i = 1, 2, \dots, p$ следва $\langle a_0, x\rangle = 0$, то $a_0 = \sum\limits_{i=1}^{p} \lambda_i a_i$ за някои числа λ_i, $i = 1, 2, \dots, p$.“ Или още така: „Ако всяко решение x на системата линейни уравнения $\langle a_i, x\rangle = 0$ е решение и на уравнението $\langle a_0, x\rangle = 0$, то a_0 е линейна комбинация на a_i, $i = 1, 2, \dots, p$.“

Доказателство. Без ограничение на общността можем да смятаме, че векторите a_i, $i = 1, 2, \dots, p$, са линейно независими. В противен случай (т. е. ако са линейно зависими) някои от уравненията в системата $\langle a_i, , x\rangle = 0$, $i = 1, 2, \dots, p$, ще са следствие от останалите и ако ги отстраним от системата, множеството от решенията няма да се промени.

От условието на твърдението е очевидно, че системата

$$\left|\begin{array}{l} \langle a_i, x\rangle = 0, \quad i = 1, 2, \dots, p, \\ \langle a_0, x\rangle = 1 \end{array}\right.$$

няма решение. Съгласно споменатите резултати от линейната алгебра векторите $a_0, a_1, \dots, a_p$ са линейно зависими, т. е. съществуват такива числа μ_i, $i = 0, 1, \dots, p$, не всички равни на нула, че $\sum\limits_{i=0}^{p} \mu_i a_i = 0$. При това $\mu_0 \neq 0$, защото векторите a_i, $i = 1, 2, \dots, p$, са линейно независими. Полагаме $\lambda_i = -\dfrac{\mu_i}{\mu_0}$, $i = 1, 2, \dots, p$, и получаваме $a_0 = \sum\limits_{i=0}^{p} \lambda_i a_i$. Твърдението е доказано. □

Твърдение 2 (теорема на Фаркаш)**.** Нека a_i, $i = 0, 1, \dots, p$, са такива вектори в $\mathbb{R}^n$, че $\langle a_0, x\rangle \leq 0$, когато $\langle a_i, x\rangle \leq 0$, $i = 1, 2, \dots, p$ (т. е. от неравенствата $\langle a_i, x\rangle \leq 0$, $i = 1, 2, \dots, p$, следва неравенството $\langle a_0, x\rangle \leq 0$). Тогава съществуват такива числа $\lambda_i \geq 0$, $i = 1, 2, \dots, p$, че $a_0 = \sum\limits_{i=1}^{p} \lambda_i a_i$.

Доказателство. Ако $a_0 = 0$, няма какво да доказваме, защото можем да изберем числата $\lambda_i = 0$, $i = 1, 2, \dots, p$. Затова ще предполагаме, че $a_0 \neq 0$. Ще докажем твърдението чрез индукция по p. При $p = 1$ то изглежда така: Ако от $\langle a_1, x\rangle \leq 0$ следва $\langle a_0, x\rangle \leq 0$, то $a_0 = \lambda_1 a_1$ за някое число $\lambda_1 \geq 0$.

Ще забележим най-напред, че от $\langle a_1, x\rangle = 0$ следват неравенствата $\langle a_0, x\rangle \leq 0$ и $\langle a_0, -x\rangle \leq 0$, т. е. от $\langle a_1, x\rangle = 0$ следва $\langle a_0, x\rangle = 0$. От предишното твърдение получаваме, че $a_0 = \lambda_1 a_1$. Остава да се убедим, че $\lambda_1 \geq 0$. Това може да стане по следния начин: числото $\langle a_0, a_1 \rangle$ е положително, защото, ако допуснем противното ($\langle a_0, a_1 \rangle \leq 0$), ще получим (според условието на твърдението), че $\|a_0\|^2 = \langle a_0, a_0 \rangle \leq 0$. Тогава от равенството $\langle a_0, a_1 \rangle = \lambda_1 \langle a_0, a_0 \rangle$ се вижда, че $\lambda_1 > 0$.

Нека твърдението е в сила за системи от не повече от $p-1$ неравенства и нека е дадено, че $\langle a_0, x \rangle \leq 0$ за всяко $x \in \mathbb{R}^n$, за което $\langle a_i, x \rangle \leq 0$, $i = 1, 2, \ldots, p$ (т. е. дадена е система от p неравенства). Отново забелязваме, че от $\langle a_i, x \rangle = 0$, $i = 1, 2, \ldots, p$, следва равенството $\langle a_0, x \rangle = 0$.

Съгласно твърдение 1 съществуват числа λ_i, $i = 1, 2, \ldots, p$, за които $a_0 = \sum_{i=1}^{p} \lambda_i a_i$.

Ще покажем, че сред всевъзможните представяния на a_0 като линейна комбинация на $\{a_i\}_{i=1}^{p}$ има и такова представяне, в което всички коефициенти са неотрицателни. Нека да допуснем, че в представянето $a_0 = \sum_{i=1}^{p} \lambda_i a_i$ има неотрицателни коефициенти λ_i. Без ограничение на общността можем да си мислим, че първите s, $s < p$, коефициента са отрицателни: $\lambda_1 < 0$, $\lambda_2 < 0$, $\ldots$, $\lambda_s < 0$, $\lambda_{s+1} \geq 0$, $\ldots$, $\lambda_p \geq 0$. Разглеждаме вектора $c = a_0 - \sum_{i=2}^{s} \lambda_i a_i = \lambda_1 a_1 + \sum_{i=s+1}^{p} \lambda_i a_i$. Ще докажем, че $\langle c, x \rangle \leq 0$ за всяко $x \in \mathbb{R}^n$, за което $\langle a_i, x \rangle \leq 0$, $i = 2, 3, \ldots, p$ (т. е. неравенството $\langle c, x \rangle \leq 0$ е следствие от системата $\langle a_i, x \rangle \leq 0$, $i = 2, 3, \ldots, p$, която се състои от $p - 1$ неравенства). Това ще ни даде възможност да използваме индуктивното предположение.

И така, нека за някое $x \in \mathbb{R}^n$ са в сила неравенствата $\langle a_i, x \rangle \leq 0$, $i = 2, 3, \ldots, p$. Възможни са два случая: 1) $\langle a_1, x \rangle \leq 0$; 2) $\langle a_1, -x \rangle \leq 0$.

В първия случай имаме $\langle a_i, x \rangle \leq 0$ за $i = 2, 3, \ldots, p$ и от условието на индуктивното предположение следва, че $\langle a_0, x \rangle \leq 0$. От представянето $c = a_0 - \sum_{i=2}^{s} \lambda_i a_i$, където $\lambda_i < 0$, се вижда, че

$$\langle c, x \rangle = \langle a_0, x \rangle - \sum_{i=2}^{s} \lambda_i \langle a_i, x \rangle \leq 0.$$

Ако е в сила 2), от представянето

$$c = \lambda_1 a_1 + \sum_{i=s+1}^{p} \lambda_i a_i, \quad \lambda_1 < 0, \quad \lambda_i \geq 0, \quad i = s+1, \ldots, p,$$

също се вижда, че $\langle c, x \rangle \leq 0$. Според индуктивното предположение $c =$

$\sum\limits_{i=2}^{p} \mu_i a_i$, където $\mu_i \geq 0$, $i = 2, 3, \ldots, p$. Тогава

$$c_0 = c + \sum_{i=2}^{s} \lambda_i a_i = \sum_{i=2}^{s} (\lambda_i + \mu_i) a_i + \sum_{i=s+1}^{p} \mu_i a_i.$$

В това представяне на a_0 като линейна комбинация от векторите $\{a_i\}_{i=1}^{p}$ броят на отрицателните коефициенти е най-много $s - 1$, т. е. с единица по-малък от първоначалния брой s на отрицателните коефициенти. Като продължаваме по подобен начин, след краен брой стъпки ще получим представянето на a_0 чрез линейна комбинация на $\{a_i\}_{i=1}^{p}$, в която всички коефициенти са неотрицателни. Твърдението е доказано. □

Следствие 1. Нека p, q, $0 \leq p < q$ са цели числа. Ако системата неравенства

(1) $$\begin{vmatrix} \langle a_i, x \rangle \leq 0\ , & i = 1, 2, \ldots, p, \\ \langle a_i, x \rangle < 0\ , & i = p + 1, \ldots, q, \end{vmatrix}$$

има решение и неравенството

(2) $$\langle a_0, x \rangle \leq 0$$

е следствие от нея, то съществуват такива числа $\lambda_i \geq 0$, $i = 1, 2, \ldots, q$, че

(3) $$a_0 = \sum_{i=1}^{q} \lambda_i a_i.$$

Забележка. Ако $p = 0$, в системата (1) участват само строги неравенства.

Доказателство. Ще докажем, че неравенството (2) е следствие и от системата

(4) $$\langle a_i, x \rangle \leq 0, \quad i = 1, 2, \ldots, q,$$

и след това ще приложим теоремата на Фаркаш.

Нека $x \in \mathbb{R}^n$ е решение на (4) и нека $h \in \mathbb{R}^n$ е решение на (1). Разглеждаме вектора $x + th$, където $t > 0$. От равенството $\langle a_i, x + th \rangle = \langle a_i, x \rangle + t \langle a_i, h \rangle$ се вижда, че $x + th$ е решение на (1). Следователно съгласно условието на доказаното твърдение $x + th$ ще удовлетворява неравенството (2):

(5) $$0 \geq \langle a_0, x + th \rangle = \langle a_0, x \rangle + t \langle a_0, h \rangle.$$

Тъй като t е произволно положително число, от (5) следва, че $\langle a_0, x \rangle \leq 0$.

От теоремата на Фаркаш се убеждаваме, че съществуват неотрицателни числа $\lambda_1, \lambda_2, \dots, \lambda_q$ такива, че

$$a_0 = \sum_{i=1}^{q} \lambda_i a_i.$$

Твърдението е доказано. □

Следствие 2. Ако системата неравенства (1) е несъвместима, то съществуват такива числа $\lambda_i \geq 0$, $i = 1, 2, \dots, q$, че $\sum_{i=1}^{q} \lambda_i a_i = 0$ и $\sum_{i=p+1}^{q} \lambda_i > 0$.

Доказателство. Без ограничение на общността можем да си мислим, че $a_i \neq 0$, $i = p+1, \dots, q$. Нека s е такова число, че подсистемата от първите s неравенства в (1) е съвместима (има решение), а системата от първите $s+1$ неравенства в (1) е несъвместима, т. е. няма решение. Това означава, че неравенството $\langle a_{s+1}, x\rangle \geq 0$ е следствие от първите s неравенства на (1). От следствие 1 се вижда, че съществуват числа $\lambda_i \geq 0$, $i = 1, 2, \dots, s$, за които $-a_{s+1} = \sum_{i=1}^{s} \lambda_i a_i$. Полагаме $\lambda_{s+1} = 1$, $\lambda_i = 0$ за $s+1 < i \leq q$ и получаваме $\sum_{i=1}^{q} \lambda_i a_i = 0$. Тъй като нулевият вектор е решение на системата $\langle a_i, x\rangle \leq 0$, $i = 1, 2, \dots, p$, то $s \geq p$. Следователно $\sum_{i=p+1}^{q} \lambda_i > 0$. Твърдението е доказано. □

Задача 1. Нека множеството X се състои от решенията на системата

$$\begin{aligned} \langle a_i', x\rangle &= 0, \quad i = 1, 2, \dots, l, \\ \langle a_j'', x\rangle &\leq 0, \quad j = 1, 2, \dots, m. \end{aligned}$$

Докажете, че ако $\langle a_0, x\rangle \leq 0$ за всяко $x \in X$, то $a_0 = \sum_{i=1}^{l} \lambda_i' a_i' + \sum_{j=1}^{m} \lambda_j'' a_j''$, където $\lambda_j'' \geq 0$, $j = 1, 2, \dots, m$.

Задача 2. Нека L е линейно подпространство на $\mathbb{R}^n$, а $a_0, a_1, \dots, a_m$ са такива вектори в $\mathbb{R}^n$, че за всяко $x \in L$ от $\langle a_j, x\rangle \leq 0$, $j = 1, 2, \dots, m$, следва неравенството $\langle a_0, x\rangle \leq 0$. Докажете, че съществуват такива числа $\lambda_j \geq 0$, $j = 1, 2, \dots, m$, че $\langle a_0, x\rangle = \sum_{j=1}^{m} \lambda_j \langle a_j, x\rangle$ за всяко $x \in L$.

Задача 3. Докажете, че във всеки от трите случая а), б), в) точно една от двете системи има решение:

а) 1) $Ax = c$, 2) $A^T y = 0$, $\langle c, y\rangle = 1$;

б) 1) $Ax \geq 0$, $x \geq 0$, $\langle c, x\rangle > 0$, 2) $A^T y \geq c$, $y \leq 0$;

в) 1) $Ax < 0$, 2) $A^{\mathrm{T}}y = 0$, $y \geq 0$, $y \neq 0$.

Задача 4. Докажете, че множеството $\{y \in \mathbb{R}^n : y = Ax,\ x \geq 0\}$ е затворено.

3. Приложения на теоремата на Фаркаш

Нека f е реална функция, дефинирана в $\mathbb{R}^n$ и нека a_j, b_j, $j = 1, 2, \dots, m$, са съответно вектори от $\mathbb{R}^n$ и реални числа.

Да разгледаме следната оптимизационна задача: да се намери минимумът на f в множеството

$$X = \{x \in \mathbb{R}^n : \langle a_j, x\rangle \leq b_j,\ j = 1, 2, \dots, m\},$$

т. е.

$$\begin{aligned} &f(x) \to \min, \\ &\langle a_j, x\rangle \leq b_j, \quad j = 1, 2, \dots, m. \end{aligned} \tag{1}$$

Неравенствата $\langle a_j, x\rangle \leq b_j$, $j = 1, 2, \dots, m$, се наричат *ограничения* в задачата (1). Нека $x^0 \in X$. Ограничението $\langle a_j, x\rangle \leq b_j$ се нарича *активно* в точката x^0, ако $\langle a_j, x^0\rangle = b_j$.

Ще припомним понятието производна на функции на много променливи:

Определение 1. Нека X е отворено подмножество на $\mathbb{R}^n$ и $f : X \to \mathbb{R}^1$. Казваме,че функцията f е *диференцуема* в точката $x^0 \in X$, ако съществува такъв вектор $f'(x^0)$, наричан *производна* или *градиент* на f в x^0, че да е изпълнено

$$\lim_{\substack{x \to x^0 \\ x \neq x^0}} \frac{f(x) - f(x^0) - \langle f'(x^0), x - x^0\rangle}{\|x - x^0\|} = 0.$$

Ако функцията f е диференцуема в точката x^0, тя има единствен градиент в тази точка, който е равен на

$$f'(x^0) = \left[\frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0)\right],$$

където $\dfrac{\partial f}{\partial x_i}(x^0)$ е частната производна на f по x_i в точката x^0, $i = 1, 2, \dots, n$. Ако f има непрекъснати частни производни в околност на точката x^0, то f е диференцуема в x^0.

Теорема 1. Нека f е диференцуема в точката x^0. За да бъде x^0 решение на задачата (1), необходимо е да съществуват такива неотрицателни числа λ_j, $j = 1, 2, \ldots, m$, че

$$f'(x^0) + \sum_{j=1}^{m} \lambda_j a_j = 0, \tag{2}$$

$$\lambda_j \left[\langle a_j, x^0\rangle - b_j\right] = 0, \quad j = 1, 2, \ldots, m. \tag{3}$$

Доказателство. Нека x^0 е решение на (1). Това ще рече, че $\langle a_j, x^0\rangle - b_j \le 0$ за $j = 1, 2, \ldots, m$ и $f(x) \ge f(x^0)$ за всяко x от множеството X. Да означим $J(x^0) = \{j : \langle a_j, x^0\rangle - b_j = 0\}$, т. е. $J(x^0)$ е (възможно и празното) множество от индексите на активните в точката x^0 ограничения. Да разгледаме произволен вектор $y \neq 0$, за който $\langle a_j, y\rangle \le 0$ при $j \in J(x^0)$. Ще докажем, че $\langle -f'(x^0), y\rangle \le 0$ и ще приложим теоремата на Фаркаш за векторите $-f'(x^0)$ и a_j, $j \in J(x^0)$.

Да разгледаме вектора $z = x^0 + ty$, където $t > 0$. Ще се убедим най-напред, че z удовлетворява ограниченията $\langle a_j, z\rangle - b_j \le 0$, $j = 1, 2, \ldots, m$, когато числото $t > 0$ е достатъчно близко до 0.

Наистина, ако $j \in J(x^0)$, то за всяко $t \ge 0$

$$b_j = \langle a_j^0, x_j\rangle \ge \langle a_j, x^0\rangle + t\langle a_j, y\rangle = \langle a_j, z\rangle.$$

За индекси j извън множеството $J(x^0)$ са в сила неравенствата $\langle a_j, x^0\rangle - b_j < 0$. От съображения за непрекъснатост при достатъчно малки числа $t > 0$ ще е в сила и неравенството $\langle a_j, z_t\rangle = \langle a_j, x^0 + ty\rangle < b_j$. Тъй като в точката x^0 функцията f достига минимума си в множеството X, в сила е неравенството $0 \le f(x^0 + ty) - f(x^0) = f(z_t) - f(x^0)$. Тогава $0 \le \dfrac{f(x^0 + ty) - f(x^0)}{t}$. След граничен преход по $t \to 0$, $t > 0$, получаваме $-\langle f'(x^0), y\rangle \le 0$. Следователно от $\langle a_j, y\rangle \le 0$ за $j \in J(x^0)$ получаваме $\langle -f'(x^0), y\rangle \le 0$. Това означава, че $-f'(x^0) = \sum_{j\in J(x^0)} \lambda_j a_j$ за някои числа $\lambda_j \ge 0$, $j \in J(x^0)$. Полагаме $\lambda_j = 0$ за $j \notin J(x^0)$ и тогава $f'(x^0) + \sum_{j=1}^{m} \lambda_j a_j = 0$. Освен това е ясно, че $\lambda_j(\langle a_j, x^0\rangle - b_j) = 0$ за всяко $j = 1, 2, \ldots, m$. □

Следствие 1 (Теорема на Ферма). Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е диференцуема в точката x^0. Ако точката x^0 е решение на задачата $\min\{f(x) : x \in \mathbb{R}^n\}$, то $f'(x^0) = 0$.

Определение 2. Нека X е множество в $\mathbb{R}^n$ и $x^0 \in X$. Векторът $y \in \mathbb{R}^n$ се нарича *допустима посока* за X в точката x^0, ако съществува такова число $\varepsilon > 0$, че $x^0 + ty \in X$ за всяко t, $0 \le t \le \varepsilon$. С други думи y е допустимата посока в точката x^0, ако тръгвайки от x^0 по посока y, известно време оставаме в множеството X.

От доказателството на теорема 1 се вижда, че множеството от допустимите за $X = \{x \in \mathbb{R}^n : \langle a_j, x\rangle \le b_j, j = 1, 2, \ldots, m\}$ посоки в една точка x^0 се състои от решенията на системата неравенства

$$\langle a_j, y\rangle \le 0, \quad j \in J(x^0).$$

Всъщност в доказателството на теоремата бе установено,че ако x^0 е решение на задачата (1), то $\langle f'(x^0), y\rangle \ge 0$ за всяка допустима в x^0 посока. Това е математически израз на интуитивно очевидния факт, че тръгвайки от точката на минимума в допустима посока, стойностите на функцията могат само да нарастват.

Пример 1. Да се реши задачата

$$\begin{aligned} &x_1^2 + x_2^2 + x_3^2 \to \min, \\ &\left|\begin{array}{l} 2x_1 - x_2 + x_3 \le 5, \\ x_1 + x_2 + x_3 \ge 3. \end{array}\right. \end{aligned} \tag{4}$$

Най-напред записваме

$$f(x) \to \min, \quad \langle a_1, x\rangle \le 5, \quad \langle a_2, x\rangle \le -3,$$

където $x = (x_1, x_2, x_3)$, $f(x) = x_1^2 + x_2^2 + x_3^2 = \|x\|^2$, $a_1 = (2, -1, 1)$, $a_2 = (-1, -1, -1)$. Тук $f'(x) = (2x_1, 2x_2, 2x_3)$. Ако $x = (x_1, x_2, x_3)$ е решение на задачата, то съществуват неотрицателни числа λ_1, λ_2, за които

$$\lambda_1(2x_1 - x_2 + x_3 - 5) = 0, \tag{5}$$

$$\lambda_2(-x_1 - x_2 - x_3 + 3) = 0, \tag{6}$$

$$f'(x) + \lambda_1 a_1 + \lambda_2 a_2 = 0.$$

В координати последното равенство може да се запише така:

$$2x_1 + 2\lambda_1 - \lambda_2 = 0, \tag{7}$$

$$2x_2 - \lambda_1 - \lambda_2 = 0, \tag{8}$$

$$2x_3 + \lambda_1 - \lambda_2 = 0. \tag{9}$$

Ако допуснем ,че $\lambda_1 = 0$, $\lambda_2 = 0$, от (7), (8) и (9) ще получим,че $x_1 = x_2 = x_3 = 0$. Тази точка обаче не удовлетворява условието $\langle a_2, x\rangle \le -3$. Следователно поне едно от двете числа λ_1, λ_2 е различно от нула. Възможни са следните три случая:

(i) $\lambda_1 \neq 0$, $\lambda_2 = 0$,

(ii) $\lambda_1 \neq 0$, $\lambda_2 \neq 0$,

(iii) $\lambda_1 = 0$, $\lambda_2 \neq 0$.

Разглеждаме случай (i). От (5), (7), (8) и (9) получаваме

$$\left|\begin{array}{l} 2x_1 - x_2 + x_3 = 5, \\ 2x_1 + 2\lambda_1 = 0, \\ 2x_2 - \lambda_1 = 0, \\ 2x_3 + \lambda_1 = 0, \end{array}\right. \tag{10}$$

т. е. $x_1 = -\lambda_1$, $x_2 = \dfrac{\lambda_1}{2}$, $x_3 = -\dfrac{\lambda_1}{2}$. Заместваме в първото равенство на (10) и получаваме

$$-2\lambda_1 - \frac{\lambda_1}{2} - \frac{\lambda_1}{2} = 5, \text{ т. е. } \lambda_1 = -\frac{5}{3}.$$

Това е противоречие, защото $\lambda_1 \geq 0$.

Разглеждаме случай (ii). Сега от (5), (7), (8) и (9) съответно получаваме

$$\left|\begin{array}{l} 2x_1 - x_2 + x_3 = 5, \\ x_1 + x_2 + x_3 = 3, \\ 2x_1 + 2\lambda_1 - \lambda_2 = 0, \\ 2x_2 - \lambda_1 - \lambda_2 = 0, \\ 2x_3 + \lambda_1 - \lambda_2 = 0. \end{array}\right.$$

Тази система има пет уравнения и пет неизвестни. От последните три уравнения определяме x_1, x_2 и x_3:

$$x_1 = \frac{1}{2}(\lambda_2 - 2\lambda_1),$$
$$x_2 = \frac{1}{2}(\lambda_1 + \lambda_2),$$
$$x_3 = \frac{1}{2}(\lambda_2 - \lambda_1).$$

Като заместим в първите две уравнения, получаваме

$$\lambda_3 - 3\lambda_1 = 5 \quad \text{и} \quad \frac{3}{2}\lambda_2 - \lambda_1 = 3, \quad \text{т. е.} \quad \frac{15}{2} + \frac{7}{2}\lambda_1 = 3.$$

Значи $\lambda_1 < 0$, а това е противоречие. Този случай също е невъзможен.

Остава да разгледаме случая (*iii*). Сега $\lambda_2 > 0$ и затова $x_1 + x_2 + x_3 = 3$. Освен това (7), (8), (9) стават съответно

$$2x_1 = \lambda_2, \quad 2x_2 = \lambda_3, \quad 2x_3 = \lambda_2,$$

значи $x_1 = x_2 = x_3 = 1$. Следователно решение на задачата може да бъде само точката $(1, 1, 1)$.

Досега сме доказали само, че ако задачата (4) има решение, то е $x^0 = (1, 1, 1)$. Следователно за пълното решение на задачата е необходимо още да се убедим, че решение съществува. Този момент не може да се подценява. В процеса на решаване на задачата ние намерихме всъщност две решения на системата (5), (6), (7), (8) и (9): $\lambda_1 = \lambda_2 = x_1 = x_2 = x_3 = 0$ и $\lambda_1 = 0$, $\lambda_2 = \frac{1}{2}$, $x_1 = x_2 = x_3 = 1$. Първото от тях се оказа, че не е решение на задачата (4), защото точката $x = (0, 0, 0)$ не удовлетворява ограниченията на задачата, т. е. не е допустима точка. Точката $x^0 = (1, 1, 1)$ удовлетворява ограниченията на задачата (и следователно е допустима точка), но от това все още не следва, че е решение. Последната може изобщо да няма решение. Ако обаче от някакви други съображения ни е известно, че задачата (1) има решение, то това решение обезателно ще се среща (според доказаната теорема) и сред решенията на системата (5)–(9) и следователно това решение ще бъде точно точката x^0.

Твърдения за съществуване на решения на екстремални задачи бяха разгледани в § 1. В дадения случай съществуването на решение се установява лесно, защото множествата на Лебег $L(f, \alpha)$ са ограничени (следствие 1 от § 2).

Упражнение 1. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е диференцуема в $\mathbb{R}^n$, $\lim\limits_{\|x\| \to +\infty} f(x) = +\infty$ и уравнението $f'(x) = 0$ има единствено решение x^0 в $\mathbb{R}^n$. Докажете, че x^0 е единствената точка на минимум на f в $\mathbb{R}^n$.

Задача 1.

$$x_1^2 + x_2^2 + x_3^2 \to \min, \quad x_1 + x_2 + x_3 \geq 3.$$

Задача 2.

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 &\to \min, \\ x_1 + x_2 + x_3 &\geq 3, \\ 2x_1 + 4x_2 - x_3 &\leq 7, \\ 3x_1 - 5x_2 + 2x_3 &\leq 1. \end{aligned}$$

Задача 3. Намерете най-малкото разстояние от точката $(1, 1)$ до полупространството $X = \{x \in \mathbb{R}^2,\ x_1 + x_2 \leq 0\}$.

Задача 4. Намерете най-малкото разстояние от дадена точка в $\mathbb{R}^3$ до дадено полупространство.

Задача 5. Намерете необходимото условие за минимум, аналогично на това в теорема 1, за задачата:

$$f(x) \to \min, \quad \langle a_j, x\rangle \le b_j, \ j = 1, \dots, m, \qquad \langle a_j, x\rangle = b_j, \ j = m+1, \dots, s.$$

Упътване. Представете всяко равенство като две неравенства.

4. Обща задача на математическото оптимиране. Теорема на Джон. Множители на Лагранж

Нека са дадени функциите

$$\begin{aligned} &f : \mathbb{R}^n \to \mathbb{R}^1, \\ &g_j : \mathbb{R}^n \to \mathbb{R}^1, \quad j = 1, 2, \dots, m, \\ &h_k : \mathbb{R}^n \to \mathbb{R}^1, \quad k = 1, 2, \dots, s. \end{aligned}$$

Задачата

$$\begin{gathered} f(x) \to \min, \\ x \in X = \{x \in \mathbb{R}^n : g_j(x) \le 0, \ j = 1, 2, \dots, m, \ h_k(x) = 0, \ k = 1, 2, \dots, s\} \end{gathered} \tag{1}$$

се нарича *задача на математическото оптимиране*. В общия случай функциите f, g_j и h_k са нелинейни, затова тази задача се нарича още задача на *нелинейното оптимиране*.

Ще посочим необходими условия, за да бъде една точка x^0 решение на задачата (1). Ще разгледаме най-напред частния случай, когато множеството се задава само с неравенства

$$f(x) \to \min, \quad g_j(x) \le 0, \ j = 1, 2, \dots, m, \tag{2}$$

т. е. $X = \{x : g_j \le 0, \ j = 1, 2, \dots, m\}$. За $x \in X$ полагаме $J(x) = \{j : g_j(x) = 0\}$.

Теорема 1 (Джон)**.** Нека x^0 е решение на задача (2). Нека предположим, че f и g_j, $j \in J(x^0)$, са диференцуеми в x^0 и g_j, $j \notin J(x^0)$, са непрекъснати в x^0. Тогава съществуват неотрицателни числа μ_0, μ_j, $j \in J(x^0)$, не всички равни на нула и такива, че

$$\mu_0 f'(x^0) + \sum_{j \in J(x^0)} \mu_j g_j'(x^0) = 0.$$

В доказателството ще използваме следната

Лема 1. Нека функцията $p : \mathbb{R}^n \to \mathbb{R}^1$ е диференцуема в точката x^0 и $\langle p'(x^0), d\rangle < 0$ за всяко $d \in \mathbb{R}^n$. Тогава съществува такова число $\delta_0 > 0$, че за всяко $\delta \in (0, \delta_0)$ е изпълнено неравенството $p(x^0 + \delta d) < p(x^0)$.

Доказателство. Ще изхождаме от определението за диференцуемост (вж. опр. 1, § 3): функцията p е диференцуема в точката x^0, ако съществува такъв вектор $p'(x^0)$, че

$$\lim_{\substack{x \to x^0 \\ x \neq x^0}} \frac{p(x) - p(x^0) - \langle p'(x^0), x - x^0\rangle}{\|x - x^0\|} = 0.$$

Полагаме $x = x^0 + td$, $t > 0$. Тогава

$$\lim_{\substack{t \to 0 \\ t > 0}} \left(\frac{p(x^0 + td) - p(x^0)}{t\|d\|} - \left\langle p'(x^0), \frac{d}{\|d\|} \right\rangle \right) = 0.$$

Следователно

$$\lim_{\substack{t \to 0 \\ t > 0}} \frac{p(x^0 + td) - p(x^0)}{t\|d\|} = \left\langle p'(x^0), \frac{d}{\|d\|} \right\rangle.$$

Оттук следва, че при достатъчно малки положителни числа t е в сила неравенството $p(x^0 + td) < p(x^0)$. Лемата е доказана. □

Доказателство на теоремата на Джон. Ще забележим най-напред, че системата линейни неравенства

$$\langle f'(x^0), y\rangle < 0, \quad \langle g_j'(x^0), y\rangle < 0, \quad j \in J(x^0), \tag{3}$$

(с неизвестно y) няма решение.

Наистина, ако предположим противното, ще получим от предишната лема, че съществуват вектор y и число $\delta_0 > 0$ такива, че

$$f(x^0 + \delta y) < f(x^0), \tag{4}$$

$$g_j(x^0 + \delta y) < g_j(x^0) = 0 \text{ за } j \in J(x^0) \text{ и } \delta \in (0, \delta_0). \tag{5}$$

От $g_j(x^0) < 0$ и непрекъснатостта на g_j, $j \notin J(x^0)$, в точката x^0 следва, че за достатъчно малки числа δ е в сила неравенството $g_j(x^0 + \delta y) < 0$, $j \notin J(x^0)$. Това показва, че за достатъчно малко число $\delta > 0$ точката $x = x^0 + \delta y$ е от множеството $X = \{x \in \mathbb{R}^n : g_j(x) \leq 0,\ j = 1, 2, \ldots, m\}$. От друга страна, неравенството $f(x^0 + \delta y) < f(x^0)$ показва, че x^0 не е решение на задача (2).

Допускането, че системата (3) има решение, ни доведе до противоречие. Следователно (3) няма решение.

От следствие 2 от § 2 се вижда, че съществуват неотрицателни числа μ_0 и μ_j, $j \in J(x^0)$, не всички равни на нула, за които е в сила равенството

$$\mu_0 f'(x^0) + \sum_{j \in J(x^0)} \mu_j g_j'(x^0) = 0. \tag{6}$$ □

Ще отбележим, че ако функциите $g_j(x)$, $j \notin J(x^0)$ са диференцуеми, условието (6) може да се запише така:

$$\left| \begin{aligned} &\mu_0 f'(x^0) + \sum_{j=1}^{m} \mu_j g_j'(x^0) = 0, \\ &\mu_j g_j(x^0) = 0, \ \ j = 1, 2, \dots, m. \end{aligned} \right. \tag{7}$$

Достатъчно е да положим $\mu_j = 0$ за $j \notin J(x^0)$.

Забележка. От доказателството на теоремата на Джон се вижда, че векторите y, които удовлетворяват системата неравенства

$$\langle g_j'(x^0), y \rangle < 0, \ \ j \in J(x^0),$$

определят допустима посока в точката x_0 за множеството

$$X = \{x \in \mathbb{R}^n : g_j(x) < 0, \ \ j = 1, 2, \dots, m\}.$$

Теоремата на Джон всъщност показва, че променяйки се от точката на минимума x^0 по допустима посока, стойностите на функцията $f(x)$ не намаляват ($\langle f'(x^0), y \rangle \geq 0$).

Следствие 1. Нека са изпълнени условията от теоремата на Джон, функциите g_j, $j = 1, 2, \dots, m$, са диференцуеми в точката x^0 и нека системата

$$\langle g_j'(x^0), y \rangle < 0, \quad j \in J(x^0),$$

има решение (такъв е например случаят, когато векторите $g_j'(x^0)$, $j \in J(x^0)$, са линейно независими). Тогава съществуват такива неотрицателни числа $\lambda_1, \lambda_2, \dots, \lambda_m$, че

$$\left| \begin{aligned} &f'(x^0) + \sum_{j=1}^{m} \lambda_j g_j'(x^0) = 0, \\ &\lambda_j g_j(x^0) = 0, \quad j = 1, 2, \dots, m. \end{aligned} \right. \tag{8}$$

Доказателство. От условието на твърдението и от доказателството на теоремата на Джон се вижда, че системата

$$\langle g_j'(x^0), y\rangle < 0, \quad j \in J(x^0),$$

има решение и че всички нейни решения удовлетворяват неравенството $\langle -f'(x^0), y\rangle \leq 0$. Съгласно следствие 1 от § 2 ще съществуват неотрицателни числа λ_j, $j \in J(x^0)$, за които $-f'(x^0) = \sum_{j \in J(x^0)} \lambda_j g_j'(x^0)$. Полагаме $\lambda_j = 0$ за $j \notin J(x^0)$ и получаваме (8). Твърдението е доказано. □

Да разгледаме сега другия частен случай, когато в общата задача на нелинейното оптимиране има само ограничения от тип равенства

$$f(x) \to \min, \quad x \in X = \{x \in \mathbb{R}^n : h_k(x) = 0,\ k = 1, 2, \ldots, s\}. \tag{9}$$

Теорема 2 (за множителите на Лагранж)**.** Нека x^0 е решение на задачата (9) и нека всички функции h_k, $k = 1, 2, \ldots, s$, имат непрекъснати частни производни в околност на x^0. Тогава съществуват числа $\lambda_1, \lambda_2, \ldots, \lambda_s$, не всички равни на нула, за които

$$\lambda_0 f'(x^0) + \lambda_1 h_1'(x^0) + \cdots + \lambda_s h_s'(x^0) = 0. \tag{10}$$

Доказателство. Ще отбележим най-напред, че случаят, когато векторите $\left\{h_k'(x^0)\right\}_{k=1}^{s}$ са линейно зависими, не е интересен, защото съществуването на константите $\lambda_0, \lambda_1, \ldots, \lambda_s$, наричани още *множители на Лагранж*, е тривиално. Достатъчно е да вземем числата $\lambda_0, \lambda_1, \ldots, \lambda_s$, за които $\sum_{k=1}^{s} \lambda_k h_k'(x^0) = 0$, и да добавим числото $\lambda_0 = 0$, за да получим равенството (10). Затова ще предполагаме, че $\left\{h_k'(x^0)\right\}_{k=1}^{s}$ са линейно независими вектори. По подобни причини не е интересен и случаят $f'(x^0) = 0$. Затова ще предполагаме, че $f'(x^0) \neq 0$. При доказателството на теоремата ще използваме следната

Лема 2. При условията на Теорема 2, ако за някое $y \in \mathbb{R}^n$ са в сила равенствата $\langle h_k'(x^0), y\rangle = 0$, $k = 1, 2, \ldots, s$, то $\langle f'(x^0), y\rangle = 0$.

Преди да докажем тази лема, ще покажем как от нея следва твърдението на теоремата. Наистина от теоремата за еднородни системи линейни уравнения получаваме

$$f'(x^0) = \sum_{k=1}^{s} \lambda_k h_k'(x^0)$$

за някои числа $\lambda_1, \lambda_2, \ldots, \lambda_s$. Следователно бихме могли да положим $\lambda_0 = -1$ и да получим твърдението на теоремата.

Доказателство на лема 2. Нека y е такъв вектор, че $\langle h_k'(x^0), y\rangle = 0$ за $k = 1, 2, \dots, s$. Тъй като векторите $\left\{h_k'(x^0)\right\}_{k=1}^{s}$ са линейно независими, за всяко $k = 1, 2, \dots, s$ може да се намери вектор $y_k \in \mathbb{R}^n$ такъв, че

$$\begin{aligned} \langle h_i'(x^0), y_k\rangle &= 0 \quad \text{при } i \neq k, \\ \langle h_i'(x^0), y_k\rangle &= 1 \quad \text{при } i = k. \end{aligned}$$

Следователно квадратната матрица

$$\left\{\langle h_i'(x^0), y_k\rangle\right\}_{k,i=1}^{s} = \begin{pmatrix} 1 \ 0 \ 0 \dots 0 \\ 0 \ 1 \ 0 \dots 0 \\ \dots\dots\dots\dots \\ 0 \ 0 \ 0 \dots 1 \end{pmatrix}$$

е неизродена.

Въвеждаме нови реални променливи $\varepsilon, t_1, t_2, \dots, t_s$ и разглеждаме системата уравнения

$$\begin{aligned} &h_1(x^0 + \varepsilon y + t_1 y_1 + t_2 y_2 + \cdots + t_s y_s) = 0, \\ &h_2(x^0 + \varepsilon y + t_1 y_1 + t_2 y_2 + \cdots + t_s y_s) = 0, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &h_s(x^0 + \varepsilon y + t_1 y_1 + t_2 y_2 + \cdots + t_s y_s) = 0. \end{aligned}$$

Левите страни на тези уравнения са функции на $s + 1$ променливи: ε, t_1, t_2, $\dots$, t_s. Якобианът на тази система в точката $\varepsilon = t_1 = t_2 = \cdots = t_s = 0$ е точно детерминантата на посочената матрица, за която видяхме, че е различна от нула. Можем да приложим теоремата за неявните функции. В този случай тя гласи, че съществуват s непрекъснати функции на една променлива $t_1(\varepsilon), t_2(\varepsilon), \dots, t_s(\varepsilon)$, които са дефинирани в $(-\varepsilon_0, \varepsilon_0)$, удовлетворяват условията $t_k(0) = 0$, $k = 1, 2, \dots, s$, и

$$h_k\left(x^0 + \varepsilon y + \sum_{i=1}^{s} t_i(\varepsilon) y_i\right) = 0,\ k = 1, 2, \dots, s,\ t \in (-\varepsilon_0, \varepsilon_0)\,. \tag{11}$$

Полагаме $z(\varepsilon) = \sum\limits_{i=1}^{s} t_i(\varepsilon) y_i$. Ще пресметнем производните на функциите $t_i(\varepsilon)$ за $\varepsilon = 0$. Диференцираме по ε тъждеството (11):

$$0 = \left\langle h_k'\left(x^0 + \varepsilon y + z(\varepsilon)\right), y + \sum_{i=1}^{s} t_i'(\varepsilon) y_i \right\rangle.$$

При $\varepsilon = 0$ получаваме

$$0 = \left\langle h_k'(x^0), y + \sum_{i=1}^{s} t_i'(0) y_i \right\rangle = \langle h_k'(x^0), y\rangle + t_k'(0). \tag{12}$$

Нека сега си спомним, че за y имаме $\langle h_k'(x^0), y\rangle = 0$ при $k = 1, 2, \ldots, s$. Тогава $t_k'(0) = 0$, $k = 1, 2, \ldots, s$. Разглеждаме функцията $f_1(\varepsilon) = f(x^0 + \varepsilon y + z(\varepsilon))$. Тъй като $x^0 + \varepsilon y + z(\varepsilon) \in X$ (т. е. $x^0 + \varepsilon y + z(\varepsilon)$ е допустима точка), то $f_1(\varepsilon)$ има минимум в $\varepsilon = 0$. Следователно $f_1'(0) = 0$, където

$$f_1'(\varepsilon) = \left\langle f'\left(x^0 + \varepsilon y + z(\varepsilon)\right), y + \sum_{i=1}^{s} t_i'(0) y_i \right\rangle.$$

Като вземем предвид, че $t_i'(0) = 0$, $i = 1, 2, \ldots, s$, и $f_1'(0) = 0$, получаваме $\langle f'(x^0), y\rangle = 0$. Лемата, а с това и теоремата е доказана. $\square$

Забележка. Смисълът на доказаната теорема е илюстриран на фиг. 4. В $\mathbb{R}^2$ е разгледана задачата

$$f(x) \to \min, \quad h(x) = 0.$$

С H_0 е означено множеството $\{x \in \mathbb{R}^2 : h(x) = 0\}$, а с F_α — множествата („линиите на ниво“) $\{x \in \mathbb{R}^2 : f(x) = \alpha\}$, $\alpha \in \mathbb{R}^1$. Ако $x^0 \in H_0$ е решение на задачата, линиите на ниво H_0 и $F_{f(x^0)}$ имат обща допирателна. Градиентите $f'(x^0)$ и $h'(x^0)$ са ортогонални на тази допирателна и затова са колинеарни: $f'(x^0) = \lambda_1 h'(x^0)$.

Пример 1.

$$x_1^2 + x_2^2 \to \min, \quad x_1^4 + x_2^4 = 1.$$

Съставяме функция на Лагранж

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

Необходимото условие (10) има вида

$$\left| \begin{aligned} &\lambda_0 x_1 + 2\lambda_1 x_1^3 = 0, \\ &\lambda_0 x_2 + 2\lambda_1 x_2^3 = 0. \end{aligned} \right.$$

Ако $\lambda_0 = 0$, то $\lambda_1 \neq 0$, откъдето получаваме $x_1 = x_2 = 0$. Точката $(0, 0)$ обаче не е допустима. Следователно $\lambda_0 = 1$. Получаваме следните решения на системата:

$$(0, \pm 1), \ (\pm 1, 0), \ \left(\pm\frac{1}{\sqrt[4]{2}}, \pm\frac{1}{\sqrt[4]{2}}\right).$$

По теоремата на Вайерщрас задачата за минимум (и за максимум) има решение. С непосредствена проверка виждаме, че точките $(0, \pm 1)$ и $(\pm 1, 0)$ са точки на минимум.

Пример 2. Нека A е симетрична матрица с размери $n \times n$. Ще решим задачата

$$\langle x, Ax \rangle \to \min, \quad \langle x, x \rangle = 1.$$

Съставяме функция на Лагранж

$$L(x, \lambda) = \lambda_0 \langle x, Ax \rangle + \lambda_1 \langle x, x \rangle.$$

Необходимите условия дават системата

$$\lambda_0 Ax + \lambda_1 x = 0.$$

Ако $\lambda_0 = 0$, то $\lambda_1 x = 0$, т. е. $x = 0$. Нулата не е допустима точка, следователно $\lambda_0 \neq 0$. Полагаме $\lambda_0 = 1$. Тогава решението x^0 удовлетворява

$$Ax = -\lambda_1 x,$$

т. е. x^0 е собствен вектор на матрицата A, а $-\lambda_1$ е съответното собствено число. Като умножим двете страни на това равенство с x, ще получим

$$\min_{\|x\|^2=1} \langle Ax, x \rangle = -\lambda_1 \langle x^0, x^0 \rangle = -\lambda_1.$$

Това означава, че решение на задачата е собственият вектор, отговарящ на най-малкото собствено число, което от своя страна е равно на минималната стойност на функцията.

Задача 1 (Задача на Ферма)**.** Намерете правоъгълен триъгълник с максимално лице, сумата от дължините на катетите на който е равна на дадено число.

Задача 2 (Задача на Кеплер)**.** Измежду цилиндрите, вписани в единичното кълбо, намерете цилиндър с максимален обем.

Задача 3.

$$\begin{aligned} &x_1 x_2 x_3 \to \min, \\ &x_1^2 + x_2^2 + x_3^2 \leq 1. \end{aligned}$$

Задача 4.

$$\begin{aligned} &x_1 x_2 x_3 \to \min, \\ &2x_1 x_2 + x_2 x_3 = 12, \\ &2x_1 - x_2 = 8. \end{aligned}$$

Задача 5.

$$\begin{aligned}
&3x_1 + 4x_2 \to \max, \\
&x_1^2 + x_2^2 \le 25, \\
&x_1 x_2 \ge 4, \\
&x_1 \ge 0, \ x_2 \ge 0.
\end{aligned}$$

5. Нелинейно оптимиране (общ случай)

Ще разгледаме задачата

$$\begin{aligned}
&f(x) \to \min, \\
&g_j(x) \le 0, \quad j = 1, 2, \dots, m, \\
&h_k(x) = 0, \quad k = 1, 2, \dots, s,
\end{aligned} \tag{1}$$

където функциите f, g_j и h_k са дефинирани в $\mathbb{R}^n$ със стойности в $\mathbb{R}^1$.

Дотук разгледахме частните случаи, когато ограниченията са само равенства или само неравенства. Сега ще разгледаме общия случай.

Теорема 1. Нека точката x^0 е решение на задачата (1)) и нека функциите f, $\{g_j\}_{j=1}^m$ и $\{h_k\}_{k=1}^s$ имат непрекъснати производни в околност на точка x^0. Тогава съществуват неотрицателни числа $\{\mu_j\}_{j=0}^m$ и числа $\{\lambda_k\}_{k=1}^s$ такива, че

$$\mu_0 f'(x^0) + \sum_{j=1}^m \mu_j g_j'(x^0) + \sum_{k=1}^s \lambda_k h_k'(x^0) = 0, \tag{2}$$

$$\mu_j g_j(x^0) = 0, \quad j = 1, 2, \dots, m, \tag{3}$$

(4) не всички числа $\{\lambda_k\}_{k=1}^s$, $\{\mu_j\}_{j=0}^m$ са равни на нула.

Доказателство. Без ограничения на общността можем да считаме, че векторите $\{h_k'(x^0)\}_{k=1}^s$ са линейно независими. Можем да смятаме още, че $f'(x^0) \ne 0$. Нека $J(x^0) = \{j : g_j(x^0) = 0\}$.

Лема 1. При условията на теорема 1 системата

$$\left|\begin{aligned}
&\langle h_k'(x^0), y\rangle = 0, \quad k = 1, 2, \dots, s, \\
&\langle g_j'(x^0), y\rangle < 0, \quad j \in J(x^0), \\
&\langle f'(x^0), y\rangle < 0
\end{aligned}\right. \tag{5}$$

няма решение.

Доказателство. Нека допуснем, че y е решение. Както видяхме в доказателството на теоремата за множителите на Лагранж, съществуват вектори $y_1, y_2, \dots, y_s$ и непрекъснати функции $t_1(\varepsilon), t_2(\varepsilon), \dots, t_s(\varepsilon)$, дефинирани в $(-\varepsilon, \varepsilon)$ и такива, че

$$h_k(x^0 + \varepsilon y + z(\varepsilon)) = 0, \quad k = 1, 2, \dots s,$$

където $z(\varepsilon) = \sum_{i=1}^{s} t_i(\varepsilon) y_i$, $t_k(0) = 0$, $t_k'(0) = 0$, $k = 1, 2, \dots, s$. От тези условия следва, че $\dfrac{\|z(\varepsilon)\|}{\varepsilon} \to 0$, когато $\varepsilon \to 0$. Тъй като y по допускане удовлетворява неравенството $\langle f'(x^0), y\rangle < 0$, то за достатъчно малко $\varepsilon > 0$ ще е в сила и неравенството $\left\langle f'(x^0), y + \dfrac{z(\varepsilon)}{\varepsilon}\right\rangle < 0$. Поради диференцуемостта на f в x^0 е в сила съотношението

$$\lim_{\varepsilon\to 0} \frac{f\left(x^0 + \varepsilon\left(y + \frac{z(\varepsilon)}{\varepsilon}\right)\right) - f(x^0)}{\left\|\varepsilon(y + \frac{z(\varepsilon)}{\varepsilon})\right\|} = \left\langle f'(x^0), \frac{y}{\|y\|}\right\rangle < 0.$$

Следователно за достатъчно малки числа $\varepsilon > 0$ е в сила неравенството $f(x^0 + \varepsilon y + z(\varepsilon)) < f(x^0)$. По подобен начин можем да се убедим, че $g_j(x^0 + \varepsilon y + z(\varepsilon)) < g_j(x^0) = 0$ за $j \in J(x^0)$. При $j \notin J(x^0)$ от съображения за непрекъснатост се вижда, че при малки $\varepsilon > 0$ е в сила неравенството $g_j(x^0 + \varepsilon y + z(\varepsilon)) < 0$. Следователно при малки $\varepsilon > 0$ точката $x^0 + \varepsilon y + z(\varepsilon)$ е допустима (удовлетворява всички ограничения на задачата (1)) и освен това $f(x^0 + \varepsilon y + z(\varepsilon)) < f(x^0)$. Това противоречи на допускането, че x^0 е решение на задачата (1). Лемата е доказана. □

Да се върнем към доказателството на теорема 1. Системата (1) е еквивалентна на системата

$$\left|\begin{array}{ll} \langle h_k'(x^0), y\rangle \le 0, & \\ \langle -h_k'(x^0), y\rangle \le 0, & k = 1, 2, \dots, s, \\ \langle g_j'(x^0), y\rangle < 0, & j \in J(x^0), \\ \langle f'(x^0), y\rangle < 0, & \end{array}\right.$$

която също няма решение. Съгласно следствие 2 от § 2 съществуват такива неотрицателни числа $\{\lambda_k', \lambda_k''\}_{k=1}^{s}$ и $\{\mu_0, \mu_j\}_{j\in J(x^0)}$, не всички равни на нула, че

$$\mu_0 f'(x^0) + \sum_{j\in J(x^0)} \mu_j g_j'(x^0) + \sum_{k=1}^{s} (\lambda_k' - \lambda_k'') h'(x^0) = 0, \quad \mu_0 + \sum_{j\in J(x^0)} \mu_j > 0.$$

Полагаме $\lambda_k = \lambda_k' - \lambda_k''$, $k = 1, 2, \dots, s$, и $\mu_j = 0$ за $j \notin J(x^0)$. Тогава са в сила равенствата (2), (3) и условието (4). □

Пример 1. Да се реши задачата

$$\begin{aligned} x_1^2 + x_2^2 + x_3^2 &\to \min, \\ 2x_1 - x_2 + x_3 &\le 5, \\ x_1 + x_2 + x_3 &= 3. \end{aligned}$$

Решение. Съставяме функция на Лагранж

$$L = \lambda_0(x_1^2 + x_2^2 + x_3^2) + \lambda_1(2x_1 - x_2 + x_3 - 5) + \lambda_2(x_1 + x_2 + x_3 - 3).$$

Необходимите условия имат вида

$$\left| \begin{array}{l} 2\lambda_0 x_1 + 2\lambda_1 + \lambda_2 = 0, \\ 2\lambda_0 x_2 - \lambda_1 + \lambda_2 = 0, \\ 2\lambda_0 x_3 + \lambda_1 + \lambda_2 = 0, \\ \lambda_1(2x_1 - x_2 + x_3 - 5) = 0, \\ \lambda_1 \ge 0. \end{array} \right.$$

Ако положим $\lambda_0 = 0$, ще получим $\lambda_1 = \lambda_2 = 0$, т. е. противоречие. Нека $\lambda_0 = \frac{1}{2}$. От горната система следват

$$x_1 = x_2 = x_3 = 1, \ \lambda_1 = 0, \ \lambda_2 = -1.$$

Задачата има решение (докажете). Следователно $(1, 1, 1)$ е единствената точка на минимум.

Задача 1.

$$\begin{aligned} &2x_1^2 + 2x_1 + 4x_2 - 3x_3 \to \min, \\ &8x_1 - 3x_2 + 3x_3 \le 40, \ x_2 \ge 0. \end{aligned}$$

Задача 2.

$$\begin{aligned} &e^{x_1 - x_2} - x_1 - x_2 \to \min, \\ &x_1 + x_2 = 1, \ x_1 \ge 0, \ x_2 \ge 0. \end{aligned}$$

6. Достатъчни условия за екстремум. Седлова точка на функцията на Лагранж

Да разгледаме оптимизационната задача

$$\begin{aligned} &f(x) \to \min, \\ &g_j(x) \le 0, \ j = 1, 2, \dots, m, \\ &h_k(x) = 0, \ k = 1, 2, \dots, s, \\ &x \in X, \end{aligned} \tag{1}$$

където X е подмножество на $\mathbb{R}^n$, а f, g_j и h_k са функции, дефинирани в X със стойности с $\mathbb{R}^1$.

Нека $\mathbb{R}^m_+$ е множеството на всички вектори в $\mathbb{R}^m$ с неотрицателни компоненти и нека $M = \mathbb{R}^m_+ \times \mathbb{R}^s$. Елементите на M са точки от вида $u = (\mu, \lambda)$, $\mu \in \mathbb{R}^m_+$, $\lambda \in \mathbb{R}^s$. За $x \in X$ и $u \in M$ дефинираме

$$L(x,u) = f(x) + \sum_{j=1}^{m} \mu_j g_j(x) + \sum_{k=1}^{s} \lambda_k h_k(x).$$

Определение 1. Функцията $L(x,u)$, дефинирана в $X \times M$, се нарича *функция на Лагранж* за задачата (1).

Определение 2. Точката $(x^0, u^0) \in X \times M$ се нарича *седлова* за функцията на Лагранж в областта $X \times M$, ако следните неравенства са изпълнени за всяко $x \in X$ и $u \in M$:

$$L(x^0, u) \le L(x^0, u^0) \le L(x, u^0). \tag{2}$$

Дясното неравенство показва, че функцията $L(x, u^0)$ достига минимума си по $x \in X$ в точката $x = x^0$, а от лявото неравенство се вижда, че u^0 „доставя" максимум за $L(x^0, u)$ в M. На фиг. 5 е изобразена графиката на функцията $L(x,u)$, за която (x^0, u^0) е седлова точка.

Теорема 1. Достатъчно условие, за да бъде $x^0 \in X$ решение на (1), е да съществува точка $u^0 \in M$, за която (x^0, u^0) е седлова точка на $L(x, y)$ в областта $X \times M$.

Доказателство. От първото неравенство в условието за седлова точка (2) имаме

$$f(x^0) + \sum_{j=1}^{m} \mu_j g_j(x^0) + \sum_{k=1}^{s} \lambda_k h_k(x^0) \le f(x^0) + \sum_{j=1}^{m} \mu_j^0 g_j(x^0) + \sum_{k=1}^{s} \lambda_k^0 h_k(x^0). \tag{3}$$

Вижда се, че $g_j(x^0) \le 0$ и $h_k(x^0) = 0$ за $j = 1, 2, \dots, m$ и $k = 1, 2, \dots, s$. Наистина, ако допуснем, че $g_{j_0}(x^0) > 0$, то като подберем достатъчно голямо число $\mu_{j_0} > 0$, лявата страна на неравенството ще стане по-голяма от дясната $L(x^0, u^0)$, което е противоречие. Ако $h_{k_0}(x^0) \ne 0$, при подходящ избор на λ_{k_0} отново ще е нарушено неравенството (3). С това доказахме, че точката x^0 е допустима, т. е. че удовлетворява ограниченията $g_j(x^0) \le 0$, $j = 1, 2, \dots, m$, $h_k(x^0) = 0$, $k = 1, 2, \dots, s$, и $x^0 \in X$. В частност, $\sum\limits_{j=1}^{m} \mu_j^0 g_j(x^0) \le 0$. При $\mu_j = 0$, $j = 1, 2, \dots, m$, от първото неравенство в (2) получаваме още $0 \le \sum\limits_{j=1}^{m} \mu_j^0 g_j(x^0)$.

Следователно $\sum_{j=1}^{m} \mu_j^0 g_j(x^0) = 0$. Тогава второто неравенство в (2) добива вида

$$(4) \qquad f(x^0) \le f(x) + \sum_{j=1}^{m} \mu_j^0 g_j(x) + \sum_{k=1}^{s} \lambda_k^0 h_k(x).$$

По условие ни е дадено, че то е изпълнено за всяко $x \in X$. Ако $x \in X$ е точка, за която са изпълнени и условията

$$g_j(x) \le 0, \quad j = 1, 2, \dots, m,$$
$$h_k(x) = 0, \quad k = 1, 2, \dots, s,$$

то от неравенството (4) ще следва, че $f(x^0) \le f(x)$, т. е. че x^0 е решение на задачата (1). □

Ако означим с $g(x)$ и $h(x)$ съответно векторите $(g_1(x), g_2(x), \dots, g_m(x)) \in \mathbb{R}^m$ и $(h_1(x), h_2(x), \dots, h_s(x)) \in \mathbb{R}^s$, то функцията на Лагранж се записва и така:

$$L(x, u) = L(x, \mu, \lambda) = f(x) + \langle \mu, g(x) \rangle + \langle \lambda, h(x) \rangle.$$

Съществува клас задачи, за които достатъчното условие, съдържащо се в теорема 1, е и необходимо условие. С този клас ще се запознаем подробно в пета глава. Тук ще илюстрираме такава ситуация, като разгледаме задачата за минимум на линейна функция при линейни ограничения:

$$(5) \qquad \langle a_0, x \rangle \to \min, \quad \langle a_j, x \rangle \le b_j, \ j = 1, 2, \dots, m, \ x \in \mathbb{R}^n,$$

където $a_j \in \mathbb{R}^n$, $j = 0, 1, \dots, m$. В теорема 1 на § 3 доказахме, че ако x^0 е решение на (5), то съществуват такива неотрицателни числа λ_j^0, $j = 1, 2, \dots, m$, че са изпълнени равенствата

$$(6) \qquad a^0 + \sum_{j=1}^{m} \lambda_j^0 a_j = 0,$$

$$(7) \qquad \lambda_j^0 \left(\langle a_j, x^0 \rangle - b_j \right) = 0, \ j = 1, 2, \dots, m.$$

Ще покажем, че от тези условия следва, че точката (x^0, λ^0), където $\lambda^0 = (\lambda_1^0, \dots, \lambda_m^0)$ е седлова точка на функцията на Лагранж в областта $\mathbb{R}^n \times \mathbb{R}^m_+$. В случая тази функция има вида

$$L(x, \lambda) = \langle a_0, x \rangle + \sum_{j=1}^{m} \lambda_j \left(\langle a_j, x \rangle - b_j \right).$$

Действително, от (6) получаваме, че за всяко $x \in \mathbb{R}^n$

$$L(x^0, \lambda^0) = L(x, \lambda^0) = -\sum_{j=1}^{m} \lambda_j^0 b_j,$$

т. е. $L(x, \lambda^0)$ е константа. От друга страна, тъй като $\langle a_j, x^0 \rangle \le b_j$, $j = 1, 2, \dots, m$, то за всяко $\lambda \in R_+^m$ е изпълнено

$$\lambda_j \left(\langle a_j, x^0 \rangle - b_j \right) \le 0.$$

Следователно, вземайки предвид (7),

$$L(x^0, \lambda) \le L(x^0, \lambda^0)$$

за всяко $\lambda \in \mathbb{R}_+^m$, т. е. (x^0, λ^0) е седлова точка. И така доказахме, че на линейната задача (5) точката x^0 е решение тогава и само тогава, когато съществува такова $\lambda^0 \in \mathbb{R}_+^m$, че (x^0, λ^0) е седлова точка за функцията на Лагранж в областта $\mathbb{R}^n \times \mathbb{R}_+^m$.

7. Нееднородни системи линейни неравенства. Двойнственост в задачата на линейното оптимизиране

Да означим с A матрицата с размер $m \times n$, чиито редове са векторите a_j, $j = 1, \dots, m$, и нека b е вектор с компоненти b_j, $j = 1, \dots, m$. Разглеждаме задачата

$$\text{(1)} \qquad \langle a_0, x \rangle \to \min, \quad Ax \le b.$$

Функцията на Лагранж има вида

$$L(x, \lambda) = \langle a_0, x \rangle + \langle \lambda, Ax - b \rangle$$

и удовлетворява равенството

$$L(x, \lambda) = -L_1(\lambda, x),$$

където функцията

$$L_1(y, x) = \langle b, y \rangle + \left\langle x, -A^{\mathrm{T}} y - a_0 \right\rangle.$$

Ясно е, че ако (x^0, λ^0) е седлова точка на $L(x, \lambda)$ в областта $\mathbb{R}^n \times \mathbb{R}_+^m$, то (x^0, λ^0) е седлова точна на $L_1(y, x)$ в областта $\mathbb{R}_+^m \times \mathbb{R}^n$, т. е.

$$L_1(\lambda^0, x) \le L_1(\lambda^0, x^0) \le L_1(y, x^0)$$

за всяко $y \in \mathbb{R}^m_+$ и $x \in \mathbb{R}^n$. Но тогава от теорема 1 на предишния параграф следва, че λ^0 е решение на следната задача:

$$\langle b, y\rangle \to \min, \quad A^{\mathrm{T}} y + a_0 = 0, \quad y \geq 0.$$

Тази задача се нарича *двойнствена* на задачата (1). Ние показахме, че ако задачата (1) има решение, то двойнствената задача също има решение, при това нейното решение е лагранжовият множител на задачата (1).

Ще изясним понятието двойственост, като използуваме едно обобщение на доказаните в § 2 резултати за еднородни системи линейни неравенства.

Твърдение 1. Нека точките $a_0, a_1, \ldots, a_p$ в $\mathbb{R}^n$ и числата $b_0, b_1, \ldots, b_p$ са такива, че множеството от решения на системата неравенства $\langle a_j, x\rangle \leq b_j$, $j = 1, 2, \ldots, p$, не е празно и се съдържа в множеството от решения на неравенството $\langle a_0, x\rangle \leq b_0$ (т. е. от $\langle a_j, x\rangle \leq b_j$, $j = 1, 2, \ldots, p$, следва $\langle a_0, x\rangle \leq b_0$). Тогава съществуват такива числа $\mu_j \geq 0$, $j = 1, 2, \ldots, p$, че $a_0 = \sum_{j=1}^{p} \mu_j a_j$ и $\sum_{j=1}^{p} \mu_j b_j \leq b_0$.

Доказателство. Ще се убедим, че това твърдение следва от теоремата на Фаркаш, т. е. от свойствата на системата еднородни линейни неравенства. Нека x^0 е решение на системата неравенства

$$\langle a_j, x\rangle \leq b_j, \quad j = 1, 2, \ldots, p. \tag{2}$$

Разсъжденията се състоят от отделни стъпки. Две от тях са обособени в леми.

Лема 1. При условията на твърдение 1, ако $\langle a_j, h\rangle \leq 0$ за $j = 1, 2, \ldots, p$, то $\langle a_0, h\rangle \leq 0$.

Доказателство. Наистина, ако $\langle a_j, h\rangle \leq 0$ за $j = 1, 2, \ldots, p$, то за всяко число $t > 0$ ще са в сила неравенствата $\langle a_j, x^0 + th\rangle = \langle a_j, x^0\rangle + t\langle a_j, h\rangle \leq b_j$, $j = 1, 2, \ldots, p$. Следователно точката $x^0 + th$ е решение на системата (2). Съгласно условието на доказваното твърдение ще е в сила и неравенството $b_0 \geq \langle a_0, x^0 + th\rangle = \langle a_0, x^0\rangle + t\langle a_0, h\rangle$. Това е възможно само ако $\langle a_0, h\rangle \leq 0$. Лемата е доказана. □

Лема 2. При условията на твърдение 1 всяко решение $(x, \xi) \in \mathbb{R}^{n+1}$ на системата неравенства

$$\langle a_j, x\rangle - b_j \xi \leq 0, \ j = 1, 2, \ldots, p, \quad \xi \geq 0 \tag{3}$$

е решение и на неравенството $\langle a_0, x\rangle - b_0 \xi \leq 0$.

Доказателство. Наистина нека (x, ξ) е решение на (3). Възможни са два случая:

а) $\xi = 0$ — в този случай твърдението на лема 2 се свежда до доказателството в лема 1;

б) $\xi > 0$ — сега точката x/ξ е решение на системата (2), а следователно и на неравенството $\left\langle a_0, \frac{x}{\xi} \right\rangle - b_0 \le 0$, т. е. $\langle a_0, x\rangle - b_0\xi \le 0$; лема 2 е доказана. □

Ще запишем сега лема 2 в друг вид, по-удобен за използване на по-рано доказаните твърдения. За целта записваме точките на $\mathbb{R}^{n+1}$ във вида (x, ξ), където $x \in \mathbb{R}^n$, а $\xi \in \mathbb{R}^1$. Да разгледаме в $\mathbb{R}^{n+1}$ точките $(a_j, -b_j)$, $j = 1, 2, \dots, p$, и $(0, -1)$, където 0 е нулевата точка в $\mathbb{R}^n$. Тогава лема 2 може да се изкаже още така: решенията на системата еднородни линейни неравенства

$$\left| \begin{array}{l} \langle (a_j, -b_j), (x, \xi)\rangle \le 0, \;\; j = 1, 2, \dots, p, \\ \langle (0, -1), (x, \xi)\rangle \le 0 \end{array} \right.$$

са решения и на неравенството $\langle (a_0, -b_0), (x, \xi)\rangle \le 0$. Като приложим теоремата за еднородни системи линейни неравенства, получаваме, че съществуват такива числа $\mu_j \ge 0$, $j = 0, 1, \dots, p$, че

$$\mu_0(0, -1) + \sum_{j=1}^{p} \mu_j(a_j, -b_j) = (a_0, -b_0).$$

Това равенство е еквивалентно на следните две:

$$\sum_{j=1}^{p} \mu_j a_j = a_0 \quad \text{и} \quad \mu_0 + \sum_{j=1}^{p} \mu_j b_j = b_0.$$

Като вземем предвид, че $\mu_0 \ge 0$, от второто равенство получаваме $\sum_{j=1}^{p} \mu_j b_j \le b_0$. Твърдението е доказано. □

Следствие 1. Нека J_1 и J_2 са крайните множества без обща точка и нека $\{a_j\}_{j\in J_1\cup J_2} \subset \mathbb{R}^n$ и $\{b_j\}_{j\in J_1\cup J_2} \subset \mathbb{R}^1$ са такива, че множеството от решения на системата

$$\left| \begin{array}{l} \langle a_j, x\rangle \le b_j, \;\; j \in J_1, \\ \langle a_j, x\rangle = b_j, \;\; j \in J_2, \end{array} \right. \tag{4}$$

не е празно и се съдържа в множеството от решенията на неравенството $\langle a_0, x\rangle \le b_0$ (т. е. от $\langle a_j, x\rangle \le b_j$ за $j \in J_1$ и $\langle a_j, x\rangle = b_j$ за $j \in J_2$ следва, че $\langle a_0, x\rangle \le b_0$). Тогава съществуват такива числа $\{\mu_j\}_{j\in J_1\cup J_2}$, че

а) $\mu_j \ge 0$ за $j \in J_1$;

б) $a_0 = \sum\limits_{j\in J_1\cup J_2} \mu_j a_j$;

в) $\sum\limits_{j\in J_1\cup J_2} \mu_j b_j \le b_0$.

Доказателство. Системата (4) може да се представи още във вида

$$\begin{aligned} \langle a_j, x\rangle &\le b_j,\ \ j \in J_1, \\ \langle a_j, x\rangle &\le b_j,\ \ j \in J_2, \\ \langle -a_j, x\rangle &\le -b_j,\ \ j \in J_2. \end{aligned} \tag{4'}$$

Решенията на (4′) са решения и на неравенството $\langle a_0, x\rangle \le b_0$. Съгласно твърдение 1 съществуват неотрицателни числа $\{\mu_j\}_{j\in J_1}$, $\{\mu'_j\}_{j\in J_2}$ и $\{\mu''_j\}_{j\in J_2}$ такива, че

$$a_0 = \sum_{j\in J_1} \mu_j a_j + \sum_{j\in J_2} (\mu'_j - \mu''_j) a_j \quad \text{и} \quad \sum_{j\in J_1} \mu_j b_j + \sum_{j\in J_2} (\mu'_j - \mu''_j) b_j \le b_0.$$

Полагаме $\mu_j = \mu'_j - \mu''_j$ за $j \in J_2$. Следствието е доказано. □

Във връзка с този резултат е уместно да разгледаме въпроса за намиране на най-малкото число (да го означим с b_0^*), за което неравенството $\langle a_0, x\rangle \le b_0$ е следствие от системата (4), т. е.

$$b_0^* = \inf\{b_0 : (4)\} \Rightarrow \langle a_0, x\rangle \le b_0. \tag{5}$$

Ясно е, че $b_0^* = \sup\limits_{x\in X} \langle a_0, x\rangle$, където X е множеството от решенията на системата (4). Ако инфимумът в (5) се достига, то съществува $x^0 \in X$ такова, че $\langle a_0, x^0\rangle = b_0^* = \sup\limits_X \langle a_0, x\rangle$.

Да разгледаме оптимизационната задача

$$\langle a_0, x\rangle \to \max, \quad x \in X = \{x \in \mathbb{R}^n : \langle a_j, x\rangle \le b_j,\ j \in J_1,\ \langle a_j, x\rangle = b_j,\ j \in J_2\}. \tag{6}$$

Тя се нарича *задача на линейното оптимиране*, защото както подлежащата на минимизиране функция $\langle a_0, x\rangle$, наричана *целева функция*, така и левите части на *условията* или *ограниченията* в (6) са линейни функции на x.

Ще покажем, че b_0^* е оптимална стойност и на една друга задача на линейното оптимиране.

Нека задачата (6) има решение, откъдето следва, че $X \neq \emptyset$ и инфимумът в (5) се достига. От следствие 1 знаем, че съществуват такива числа $\{\mu_j^0\}_{j\in J=J_1\cup J_2}$, че:

а) $\mu_j^0 \geq 0,\ j \in J_1$,

б) $\sum\limits_{j\in J} \mu_j^0 a_j = a_0$,

в) $\sum\limits_{j\in J} \mu_j^0 b_j \leq b_0^*$.

От друга страна, ако числата $\{\mu_j\}_{j\in J}$ удовлетворяват условията а) и б) и точката $x \in X$, то е в сила неравенството

$$\langle a_0, x\rangle = \sum_{j\in J} \mu_j \langle a_j, x\rangle \leq \sum_{j\in J} \mu_j b_j.$$

Следователно

$$b_0^* = \sup_{x\in X} \langle a_0, x\rangle \leq \inf_{y\in Y} \langle b, y\rangle \leq b_0^*,$$

където y и b са вектори съответно с координати μ_j и b_j, $j \in J$, а Y е множеството от вектори y, които са решение на системата

$$\left|\begin{array}{l} \sum\limits_{j\in J} y_j a_j = a_0, \\ y_j \geq 0,\ j \in J. \end{array}\right.$$

Задачата

$$\langle b, y\rangle \to \min, \quad y \in Y$$

също е задача на линейното оптимиране. От казаното дотук се вижда, че ако задачата (6) има решение x^0, то (7) също има решение, равно на $y^0 = \{\mu_j^0\}_{j\in J}$ и

$$\inf_{y\in Y} \langle b, y\rangle = b_0^* = \langle b, y^0\rangle = \langle a_0, x_0\rangle = \sup_{x\in X} \langle a_0, x\rangle. \tag{7}$$

Задачата (7) е двойствена на задачата (6). Свойствата на задачите (6) и (7) ще бъдат описани по-подробно в гл. 4.

Задача 1. Да се докаже, че непразното множество от решения на системата $\langle a_j, x\rangle \leq b_j,\ j \in J$ е ограничено тогава и само тогава, когато за всяка точка $x^0 \in \mathbb{R}^n$, $x^0 \neq 0$ съществува такова $j \in J$, че $\langle a_j, x^0\rangle > 0$.

Задача 2. Да се докаже, че точно една от следните две системи (съответно с неизвестни $x = \{x_i\}_{i=1}^n$ и $y = \{y_j\}_{j\in J}$) има решение

1) $\langle a_j, x\rangle \leq b_j,\ j \in J$,

$$
2) \left| \begin{array}{l} \sum\limits_{j \in J} y_j a_j = 0, \\ y_j \geq 0, \ \ j \in J, \\ \sum\limits_{j \in J} y_j b_j > 0. \end{array} \right.
$$

Упътване. Докажете най-напред, че 1) и 2) не могат едновременно да имат решение. Ако 2) няма решение, от условията $y_j \geq 0$, $j \in J$, и от $\sum\limits_{j \in J} y_j a_j = 0$ следва $\sum\limits_{j \in J} y_j b_j \leq 0$. Приложете теоремата на Фаркаш.

Задача 3. Да се докаже, че системата

$$
\langle a_j, x \rangle < b_j, \quad j \in J,
$$

има решение тогава и само тогава, когато от $\sum\limits_{j \in J} \mu_j a_j = 0$, $\mu_j \geq 0$, $j \in J$, следва $\sum\limits_{j \in J} \mu_j b_j \geq 0$ и освен това от $\sum\limits_{j \in J} \mu_j > 0$ следва $\sum\limits_{j \in J} \mu_j b_j > 0$.

Глава 2

Изпъкнали множества

Едно множество в $\mathbb{R}^n$ се нарича изпъкнало, ако заедно с всеки две свои точки съдържа и отсечката, която ги свързва. Повечето от множествата, разглеждани в класическата математика (триъгълник, кръг, квадрат, елипса и т. н.), са изпъкнали. Изучаването обаче на изпъкналите множества започва едва около началото на XX век. В работите си Минковски, Радон, Хели, Каратеодори са намирали интересни свойства на изпъкналите множества, които ги превръщат в привлекателен обект за теоретични изследвания, съчетаващ геометрична нагледност и аналитични свойства. През 50-те години на нашия век интересът към изпъкналите множества нараства още по-силно, защото се оказва, че те с успех могат да се използват в изследването на редица важни за практиката оптимизационни задачи. Днес без преувеличение може да се твърди, че изпъкналите множества и свързаните с тях понятия, твърдения, методи за разсъждение са съществен дял от техническия арсенал на теорията и практиката на математическото оптимиране.

В тази глава са дадени някои от основните свойства на изпъкналите множества. Ударението е поставено на т. нар. теореми за отделимост, които се оказват полезни при изследването на широк кръг задачи (излизащ далеч извън разглежданите в тази книга въпроси). Грубо казано, от теоремите за отделимост следва, че ако две изпъкнали множества в пространството $\mathbb{R}^n$ нямат общи точки, то съществува (хипер)равнина, която ги „разделя“, т. е. едното множество е изцяло от едната страна на (хипер)равнината, а второто — изцяло от другата ѝ страна. Съществуването на разделяща (хипер)равнина в различни ситуации води до характеризации на решенията на оптимизационни задачи. Теоремите за отделимост играят важна роля при разкриване на структурата на изпъкналите множества, при изследване на двойнствеността в математическото оптимиране и при обосноваването на числените методи за оптимиране.

1. Изпъкнали множества

Определение 1. Нека x и y са две точки в $\mathbb{R}^n$. Под *отсечка* $\overline{xy}$ с краища x и y ще разбираме множеството от точки z от вида $z = x + \lambda(y - x)$, където

λ е число между 0 и 1, т. е.

$$\overline{xy} = \{z \in \mathbb{R}^n : z = x + \lambda(y - x),\ 0 \le \lambda \le 1\}.$$

От фиг. 6 се вижда, че това определение съответства на интуитивните ни представи за отсечка с краища x и y.

Не е трудно да се убедим, че

$$\overline{xy} = \{z : z = (1 - \lambda)x + \lambda y,\ 0 \le \lambda \le 1\} = \{z : z = \mu x + \lambda y,\ \lambda + \mu = 1,\ \lambda, \mu \ge 0\}.$$

Определение 2. Едно множество X се нарича *изпъкнало*, ако заедно с всеки две свои точки x и y съдържа и отсечката $\overline{xy}$.

За удобство приемаме, че празното множество $\emptyset$ е изпъкнало.

На фиг. 7,а е изобразено изпъкнало множество, а на фиг. 7,б — множество, което не е изпъкнало, защото не съдържа отсечката с краища x и y.

Упражнение 1. Докажете, че кълбото $B(x, r)$ с център в x и радиус r е изпъкнало множество.

Упражнение 2. Нека A и B са матрици с размери $n \times p$ и $n \times q$ и $a \in \mathbb{R}^p$, $b \in \mathbb{R}^q$. Докажете, че множеството $X = \{x \in \mathbb{R}^n : Ax \ge a, Bx \ge b\}$ е изпъкнало.

Определение 3. Ще казваме, че точката x е *изпъкнала комбинация* на точките $x_1, x_2, \dots, x_k$, ако съществуват такива неотрицателни числа $\lambda_1, \lambda_2, \dots, \lambda_k$, че $\sum\limits_{i=1}^{k} \lambda_i = 1$ и $x = \sum\limits_{i=1}^{k} \lambda_i x_i$.

Очевидно отсечката $\overline{x_1 x_2}$ се състои от всички изпъкнали комбинации на точките x_1 и x_2.

Твърдение 1. Изпъкналото множество X съдържа всяка изпъкнала комбинация на краен брой свои точки.

Доказателство. Ще направим индукция по броя k на точките в изпъкналата комбинация (дължината на изпъкналата комбинация) $\sum\limits_{i=1}^{k} \lambda_i x_i$. При $k = 2$ твърдението е тривиално — то съвпада с определението на изпъкнало множество. Нека твърдението е вярно за произволна изпъкнала комбинация с дължина $k - 1$ и нека $\sum\limits_{i=1}^{k} \lambda_i x_i$ е изпъкнала комбинация на точките $x_i \in X$, $i = 1, 2, \dots, k$. Можем да предполагаме, че всички числа λ_i, $i = 1, 2, \dots, k$, са положителни. Полагаме $\lambda = \sum\limits_{i=1}^{k-1} \lambda_i = 1 - \lambda_k$ и тогава

$$z = \sum_{i=1}^{k} \lambda_i x_i = \lambda \sum_{i=1}^{k-1} \frac{\lambda_i}{\lambda} x_i + \lambda_k x_k.$$

Тъй като $\sum_{i=1}^{k-1} \dfrac{\lambda_i}{\lambda} = 1$, изразът $\sum_{i=1}^{k-1} \dfrac{\lambda_i}{\lambda} x_i$ е изпъкнала комбинация на точките x_i, $i = 1, 2, \dots, k-1$. От индуктивното предположение следва, че точката $z_1 = \sum_{i=1}^{k-1} \dfrac{\lambda_i}{\lambda} x_i$ е от X. От представянето $z = \lambda z_1 + \lambda_k x_k$ и от равенството $\lambda + \lambda_k = \sum_{i=1}^{k-1} \lambda_i + \lambda_k = 1$ се вижда,че z е точка от отсечката $\overline{z_1 x_k}$, чиито краища z_1 и x_k са от X. От изпъкналостта на X следва, че $z \in X$. Доказателството е завършено. □

Упражнение 3. Нека X_t, $t \in T$, е семейство от изпъкнали множества. Да се докаже, че $\bigcap\{X_t : t \in T\}$ също е изпъкнало множество.

Нека X е множество в $\mathbb{R}^n$. Чрез $\overline{X}$ ще означаваме *затворената обвивка* на множеството X, т. е. $\overline{X}$ е най-малкото затворено множество, съдържащо X. То се състои от границите на всички сходящи редици, чиито елементи са от X.

Упражнение 4. Нека X е изпъкнало множество. Докажете, че затворената обвивка $\overline{X}$ също е затворено множество.

Нека $X \subset \mathbb{R}^n$. Точката $x_0 \in X$ се нарича *вътрешна* за X, ако съществува кълбо $B(x_0, \varepsilon)$ с център x_0 и радиус ε, което се съдържа в X. Обединението на всички вътрешни за X точки ще бележим с $\operatorname{int} X$ и ще наричаме *вътрешност* на X. $\operatorname{int} X$ е най-голямото отворено множество, което се съдържа в X. Съществуват множества без вътрешни точки (т. е. $\operatorname{int} X = \emptyset$).

Твърдение 2. Нека множеството X е изпъкнало, $x_1 \in \overline{X}$ и $x_2 \in \operatorname{int} X$, $x_1 \neq x_2$. Тогава всяка точка y, $y \neq x_1$ от отсечката $\overline{x_1 x_2}$ е вътрешна за X.

Доказателство. Нека $\lambda \in (0, 1)$ и $y = \lambda x_1 + (1-\lambda) x_2$. Понеже $x_2 \in \operatorname{int} X$, то съществува $\varepsilon > 0$ такова, че $B(x_2, \varepsilon) \subset X$. Ще покажем, че $B(y, (1-\lambda)\varepsilon) \subset X$. Да изберем произволно $z \in B(y, (1-\lambda)\varepsilon)$. Тъй като $x_1 \in \overline{X}$, множеството

$$B(x_1, ((1-\lambda)\varepsilon - \|z - y\|)/\lambda) \cap X$$

е непразно (фиг. 8). В частност съществува $z_1 \in X$ такова, че

$$\|z_1 - x_1\| < \frac{(1-\lambda)\varepsilon - \|z - y\|}{\lambda}. \tag{1}$$

Нека $z_2 = \dfrac{z - \lambda z_1}{1-\lambda}$. Като използваме равенството $x_2 = \dfrac{y - \lambda x_1}{1-\lambda}$ и неравенството на триъгълника, получаваме

$$\begin{aligned}
\|z_2 - x_2\| &= \left\| \frac{z - \lambda z_1}{1-\lambda} - \frac{y - \lambda x_1}{1-\lambda} \right\| = \left\| \frac{(z - \lambda z_1) - (y - \lambda x_1)}{1-\lambda} \right\| \\
&= \frac{1}{1-\lambda} \|z - y + \lambda(x_1 - z_1)\| \leq \frac{1}{1-\lambda} (\|z - y\| + \lambda \|x_1 - z_1\|).
\end{aligned}$$

От последното неравенство и от (1) следва, че $\|z_2 - x_2\| < \varepsilon$, т. е. $z_2 \in X$. Поради специалния избор на точката z_2 в сила е равенството $z = \lambda z_1 + (1-\lambda)z_2$. Следователно $z \in X$. Показахме, че всяка точка z, за която $\|z - y\| < (1-\lambda)\varepsilon$, принадлежи на X, т. е. $y \in \operatorname{int} X$. □

Следствие 1. Ако X е изпъкнало множество, то $\operatorname{int} X$ също е изпъкнало множество.

Точката $x_0 \in \mathbb{R}^n$ се нарича *контурна* за множеството X, ако всяко кълбо $B(x_0, \varepsilon)$, $\varepsilon > 0$, има общи точки с множеството X, така и с допълнението му $\mathbb{R}^n \setminus X$. Множеството от всички контурни за дадено множество X точки ще бележим с ∂X. Очевидно $\partial X = \overline{X} \setminus \operatorname{int} X$.

Следствие 2. Нека X е изпъкнало множество и $\operatorname{int} X \neq \emptyset$. Тогава

1) $\overline{X} = \overline{\operatorname{int} X}$,

2) $\operatorname{int} X = \operatorname{int} \overline{X}$,

3) $\partial X = \partial \overline{X}$.

Доказателство. 1) Очевидно $\overline{\operatorname{int} X} \subset \overline{X}$. Ще докажем, че $X \subset \overline{\operatorname{int} X}$. Тогава ще е в сила и съотношението $\overline{X} \subset \overline{\operatorname{int} X}$. Нека $x \in X$ и $y \in \operatorname{int} X$. Според твърдение 2 при $\lambda \in (0,1)$ имаме $\lambda x + (1-\lambda)y \in \operatorname{int} X$. Това означава, че $x = \lim\limits_{\substack{\lambda \to 1 \\ \lambda < 1}} (\lambda x + (1-\lambda)y)$ е от $\overline{\operatorname{int} X}$.

2) Очевидно $\operatorname{int} X \subset \operatorname{int} \overline{X}$. Нека $x \in \operatorname{int} \overline{X}$ и $z \in \operatorname{int} X$, $z \neq x$. При достатъчно малко $\mu > 0$ точката $y = x + \mu(x - z) \in \operatorname{int} \overline{X} \subset \overline{X}$. Но тогава $x = \dfrac{1}{\mu+1} y + \dfrac{\mu}{\mu+1} z \in \overline{yz}$. Прилагаме твърдение 2 и получаваме $x \in \operatorname{int} X$.

3) $\partial X = \overline{X} \setminus \operatorname{int} X$, което от свойство 2) е равно на $\overline{X} \setminus \operatorname{int} \overline{X} = \partial \overline{X}$. □

Твърдение 3. Нека изпъкналото и затворено множество X съдържа лъча $\{x_0 + tb : t \geq 0\}$, $b \in \mathbb{R}^n$. Тогава за всяко $x \in X$ множеството X съдържа и лъча $\{x + tb : t \geq 0\}$.

Доказателство. Нека $x \in X$ и $t \geq 0$ са произволно избрани. От условието $x_0 + ktb \in X$, $k = 1, 2, \dots$, следва

$$x_k = \left(1 - \frac{1}{k}\right) x + \frac{1}{k}(x_0 + ktb) \in X, \ k = 1, 2, \dots$$

Но $x_k \to x + tb$, $k \to +\infty$, следователно $x + tb \in X$. □

Задача 1. Нека X е изпъкнало множество, $x_1 \in \operatorname{int} X$, $x_2 \in \partial X$. Докажете, че за всяко $\alpha > 0$ точката $y = x_2 + \alpha(x_2 - x_1)$ не принадлежи на $\overline{X}$.

2. Изпъкнала обвивка на множество

Нека X е произволно множество. Съвкупността от всички изпъкнали комбинации на негови точки се нарича *изпъкнала обвивка* на X и се бележи $\operatorname{co} X$, т. е.

$$\operatorname{co} X = \left\{ z \in \mathbb{R}^n : z = \sum_{i=1}^{k} \lambda_i x_i,\ x_i \in X,\ \sum_{i=1}^{k} \lambda_i = 1,\ \lambda_i \geq 0,\ i = 1, 2, \ldots, k \right\}.$$

Изрично ще отбележим, че в $\operatorname{co} X$ могат да участвуват изпъкнали комбинации с произволна дължина k.

На фиг. 9 са показани изпъкнали обвивки на различни множества (с плътна линия са начертани множествата, а защрихованата част дава изпъкналата обвивка). На фиг. 9, а множеството X е полуокръжност, а на фиг. 9, б X се състои от окръжност k и точка x_0 извън нея.

Твърдение 1. Множеството X е изпъкнало тогава и само тогава, когато $X = \operatorname{co} X$.

Доказателство. От дефиницията на $\operatorname{co} X$ и твърдение 1 от § 1 се вижда, че $\operatorname{co} X \subset X$, когато X е изпъкнало множество. За да докажем обратното, достатъчно е да забележим, че $\operatorname{co} X$ е изпъкнало множество, каквото и да е множеството X. Действително нека $z_1 = \sum_{i=1}^{k_1} \lambda_i x_i$ и $z_2 = \sum_{j=1}^{k_2} \mu_j y_j$ са точки от $\operatorname{co} X$. За $\alpha \in [0, 1]$ получаваме

$$z = \alpha z_1 + (1 - \alpha) z_2 = \alpha \sum_{i=1}^{k_1} \lambda_i x_i + (1 - \alpha) \sum_{j=1}^{k_2} \mu_j y_j.$$

Числата $\alpha\lambda_i$, $i = 1, 2, \ldots, k_1$, и $(1 - \alpha)\mu_j$, $j = 1, 2, \ldots, k_2$, са неотрицателни и тяхната сума е

$$\alpha \sum_{i=1}^{k_1} \lambda_i + (1 - \alpha) \sum_{j=1}^{k_2} \mu_j = \alpha + 1 - \alpha = 1.$$

Следователно z е изпъкнала комбинация на точките x_1, x_2, $\ldots$, x_{k_1}, y_1, y_2, $\ldots$, y_{k_2}, които принадлежат на X, т. е. $z \in \operatorname{co} X$. □

Упражнение 1. Докажете, че $\operatorname{co} X$ съвпада с обединението на изпъкналите обвивки на крайните подмножества K на X: $\operatorname{co} X = \cup\{\operatorname{co} K : K \subset X, K$ — крайно множество$\}$.

Упражнение 2. Докажете, че $\operatorname{co} X$ съвпада със сечението на всички изпъкнали множества, съдържащи X.

Упражнение 3. Нека $X_1, X_2, \ldots, X_m$ са изпъкнали множества. Докажете, че $\operatorname{co}\left(\bigcup_{i=1}^{m} X_i\right)$ се състои от изпъкнали комбинации от вида $z = \sum_{i=1}^{m} \lambda_i a_i$, където $a_i \in X_i$, $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, и $\sum_{i=1}^{m} \lambda_i = 1$.

Изрично ще отбележим, че дължината на изпъкналите комбинации в тази задача не надхвърля m, докато в определението на $\operatorname{co} X$ участват произволно дълги изпъкнали комбинации.

Теорема 1 (Каратеодори)**.** Ако X е непразно подмножество на $\mathbb{R}^n$, то всяка точка $z \in \operatorname{co} X$ може да се представи като изпъкнала комбинация на не повече от $n+1$ точки от X.

Доказателство. Нека за $z = \sum_{i=1}^{k} \lambda_i x_i$ е изпъкнала комбинация на точките $x_i \in X$, $i = 1, 2, \ldots, k$. Без ограничение на общността можем да считаме, че $k > n+1$. Ще докажем, че z може да се представи като изпъкнала комбинация на $k-1$ точки. Разглеждаме векторите $y_i \in \mathbb{R}^{n+1}$, $i = 1, 2, \ldots, k$, които се получават като към n-те координати на x_i добавим още една координата — $(n+1)$-ва, която е равна на 1: $y_i = (x_i, 1)$. Тъй като $k > n+1$, векторите $\{y_i\}_{i=1}^{k} \in \mathbb{R}^{n+1}$ са линейно зависими. Следователно съществуват такива числа β_i, $i = 1, 2, \ldots, k$, че поне едно от тях не е нула и освен това е в сила равенството $\sum_{i=1}^{k} \beta_i y_i = 0$. Това равенство е еквивалентно на следните равенства:

$$\sum_{i=1}^{k} \beta_i x_i = 0, \quad \sum_{i=1}^{k} \beta_i = 0. \tag{1}$$

Тъй като не всички числа β_i са нули, от първото равенство на (1) можем да определим една от точките x_{i_0} (за която $\beta_{i_0} \neq 0$), чрез останалите:

$$x_{i_0} = -\frac{1}{\beta_{i_0}} \sum_{\substack{i=1 \\ i \neq i_0}}^{k} \beta_i x_i. \tag{2}$$

Заместваме x_{i_0} в представянето $z = \sum_{i=1}^{k} \lambda_i x_i$ с получения в (2) израз

$$z = \sum_{\substack{i=1 \\ i \neq i_0}}^{k} \lambda_i x_i + \lambda_{i_0} x_{i_0} = \sum_{\substack{i=1 \\ i \neq i_0}}^{k} \left(\lambda_i - \frac{\lambda_{i_0}}{\beta_{i_0}} \beta_i\right) x_i. \tag{3}$$

Получихме представяне за z чрез $k-1$ точки (точката x_{i_0} отпада от представянето).

Не е трудно да се види, че сумата от коефициентите пред точките x_i, $i \neq i_0$, в това представяне е точно 1. Достатъчно е да вземем предвид, че $\sum\limits_{i=1}^{k} \beta_i = 0$ и $\sum\limits_{i=1}^{k} \lambda_i = 1$. За да бъде (3) изпъкнала комбинация, трябва още $\left(\lambda_i - \dfrac{\lambda_{i_0}}{\beta_{i_0}}\beta_i\right)$, $i \neq i_0$, да са неотрицателни числа. Произволният начин, по който избрахме индекса i_0 (искахме само $\beta_{i_0} \neq 0$), не ни гарантира неотрицателност на тези числа. За да заобиколим тази трудност, ще изберем индекса i_0 по специален начин. Тъй като не всички числа β_i са нули и тъй като сумата им е нула, то сред числата β_i, $i = 1, 2, \ldots, k$, има положителни. Сред всички индекси i, за които $\beta_i > 0$, означаваме с i_0 онзи, за който числото $\dfrac{\lambda_i}{\beta_i}$ е най-малко, т.е. i_0 е такъв индекс, че $\beta_{i_0} > 0$ и $\dfrac{\lambda_i}{\beta_i} \geq \dfrac{\lambda_{i_0}}{\beta_{i_0}}$ всеки път, когато $\beta_i > 0$. Тогава за индексите i, за които $\beta_i > 0$, е изпълнено неравенството $\lambda_i - \beta_i \dfrac{\lambda_{i_0}}{\beta_{i_0}} \geq 0$. За останалите индекси (за тях $\beta_i \leq 0$) неравенството $\lambda_i - \beta_i \dfrac{\lambda_{i_0}}{\beta_{i_0}} \geq 0$ е изпълнено по очевидни съображения ($\dfrac{\lambda_{i_0}}{\beta_{i_0}} \geq 0$ и $\lambda_i \geq 0$). Теоремата е доказана, защото можем чрез неколкократно повтаряне на описаните разсъждения да представим z като изпъкнала комбинация на не повече от $n + 1$ точки. □

Следствие 1. Нека X е компактно множество на $\mathbb{R}^n$. Тогава $\operatorname{co} X$ също е компактно множество на $\mathbb{R}^n$.

Доказателство. Достатъчно е да докажем, че всяка редица $\{z_m\}_{m \geq 1}$ от точки на $\operatorname{co} X$ съдържа сходяща подредица, която клони към точка от $\operatorname{co} X$.

По теоремата на Каратеодори $z_m = \sum\limits_{i=1}^{n+1} \lambda_{i,m} x_{i,m}$, където $\lambda_{i,m} \geq 0$, $\sum\limits_{i=1}^{n+1} \lambda_{i,m} = 1$ и $x_{i,m} \in X$. Всяка от редиците $\{\lambda_{i,m}\}_{m \geq 1}$, $\{x_{i,m}\}_{m \geq 1}$, $i = 1, 2, \ldots, n + 1$, лежи в ограниченото и затворено множество $\{0 \leq \lambda_{i,m} \leq 1,\ x_{i,m} \in X\}$ и затова от всяка от тях може да се избере сходяща подредица. Тъй като подредица на сходяща редица е също сходяща (и клони към същата граница), можем да твърдим, че съществува редица от индекси $m_1 < m_2 < \cdots$, за която редиците $\{x_{i,m_s}\}_{s \geq 1}$ и $\{\lambda_{i,m_s}\}_{s \geq 1}$ са сходящи и клонят съответно към $x_{i,0}$ и $\lambda_{i,0}$, $i = 1, 2, \ldots, n+1$. От равенствата $z_{m_s} = \sum\limits_{i=1}^{n+1} \lambda_{i,m_s} x_{i,m_s}$ и $\sum\limits_{i=1}^{n+1} \lambda_{i,m_s} = 1$ следва, че редицата $\{z_{m_s}\}$ клони към $z_0 = \sum\limits_{i=1}^{n+1} \lambda_{i,0} x_{i,0}$, като освен това $\sum\limits_{i=1}^{n+1} \lambda_{i,0} = 1$ и $\lambda_{i,0} \geq 0$, $i = 1, 2, \ldots, n + 1$. Следователно z_0 е изпъкнала комбинация на точките $x_{i,0}$, $i = 1, 2, \ldots, n + 1$. Тъй като тези точки са от X, то $z_0 \in \operatorname{co} X$. Твърдението е доказано. □

Пример 1. Множеството

$$X = \{(x_1, x_2),\ x_1 \geq 0,\ x_2 = \sqrt{x_1}\}$$

е затворено, но не е ограничено. Неговата изпъкнала обвивка

$$\operatorname{co} X = \{(x_1, x_2),\ x_1 \geq 0,\ 0 < x_2 \leq \sqrt{x_1}\}$$

не е затворено множество.

Задача 1. Да се докаже, че ако X е отворено, то $\operatorname{co} X$ също е отворено множество.

Нека $P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ е полином с реални коефициенти, $a_0 \neq 0$. Нулите на този полином са комплексни числа и могат да се разглеждат като точки от равнината $\mathbb{R}^2$.

Задача 2 (теорема на Гаус—Люка)**.** Нека $P(z) = a_0z^n + a_1z^{n-1} + \cdots + a_{n-1}z + a_n$ е полином от степен $n > 1$. Докажете, че нулите на полинома $P'(z) = na_0z^{n-1} + (n-1)a_1z^{n-2} + \cdots + a_{n-1}$ се съдържат в изпъкналата обвивка на нулите на полинома $P(z)$.

3. Теорема на Радон. Теорема на Хели за изпъкналите множества

Теорема 1 (Радон)**.** Нека множеството C се състои от краен брой различни точки $x_1, x_2, \ldots, x_k$ на $\mathbb{R}^n$ и нека $k \geq n + 2$. Тогава съществуват подмножества C_1 и C_2 на C, за които са изпълнени условията:

а) $C_1 \cup C_2 = C$; б) $C_1 \cap C_2 = \emptyset$; в) $\operatorname{co} C_1 \cap \operatorname{co} C_2 \neq \emptyset$.

Доказателство. Както и в доказателството на теоремата на Каратеодори разглеждаме точки y_i, $i = 1, 2, \ldots, k$, на $\mathbb{R}^{n+1}$, които се получават от x_i, като добавим още една координата, равна на 1: $y_i = (x_i, 1)$, където $x_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, k$. Тъй като $k > n + 1$, векторите $\{y_i\}_{i=1}^k$ са линейно зависими. Значи съществуват числа β_i (не всички равни на нула), $i = 1, 2, \ldots, k$, за които $\sum\limits_{i=1}^{k} \beta_i y_i = 0$. Това равенство е еквивалентно на следните две равенства:

$$\sum_{i=1}^{k} \beta_i x_i = 0 \quad \text{и} \quad \sum_{i=1}^{k} \beta_i = 0. \tag{1}$$

Тъй като не всички числа β_i са нули, а сумата им е нула, то сред тях има както отрицателни, така и положителни. За удобство ще смятаме, че най-напред в редицата $\beta_1, \beta_2, \ldots, \beta_k$ стоят неотрицателните числа, а след това отрицателните:

$$\beta_1 \geq 0,\ \beta_2 \geq 0,\ \ldots,\ \beta_r \geq 0,\ \beta_{r+1} < 0,\ \beta_{r+2} < 0,\ \ldots,\ \beta_k < 0.$$

Полагаме $\beta = \sum\limits_{i=1}^{r} \beta_i = -\sum\limits_{i=r+1}^{k} \beta_i$. Очевидно $\beta > 0$ и $\sum\limits_{i=1}^{r} \dfrac{\beta_i}{\beta} = -\sum\limits_{i=r+1}^{k} \dfrac{\beta_i}{\beta} = 1$. От първото равенство на (1) се вижда, че $\sum\limits_{i=1}^{r} \dfrac{\beta_i}{\beta} x_i = \sum\limits_{i=r+1}^{k} \left(-\dfrac{\beta_i}{\beta}\right) x_i$. Това показва, че точката $x = \sum\limits_{i=r+1}^{k} \left(-\dfrac{\beta_i}{\beta}\right) x_i$ е изпъкнала комбинация както на точките от множеството $C_1 = \{x_1, x_2, \dots, x_r\}$, така и на точките от множеството $C_2 = \{x_{r+1}, x_{r+2}, \dots, x_k\}$. Следователно $x \in \operatorname{co} C_1 \cap \operatorname{co} C_2$. Тъй като $C_1 \cap C_2 = \emptyset$, теоремата е доказана. □

Теоремата, която ще докажем сега, има многобройни приложения. Тя дава инструмент, с чиято помощ може да се установи кога семейство от изпъкнали множества има обща точка. Доказателството се опира съществено на доказаната по-горе теорема на Радон.

Теорема 2 (Хели)**.** Нека $B_1, B_2, \dots, B_k$ е семейство от изпъкнали подмножества на $\mathbb{R}^n$. Ако всеки $n+1$ члена на това семейство имат обща точка, то и всички множества от семейството имат обща точка.

Доказателство. Ако $k \le n+1$, съществуването на точка, обща за всички множества от семейството $B_1, B_2, \dots, B_k$, следва от условието на теоремата. Затова предполагаме, че $k \ge n+1$. Най-напред ще докажем, че първите $n+2$ члена на семейството $B_1, B_2, \dots, B_{n+1}, B_{n+2}$ имат обща точка. От разсъжденията ще проличи, че първите $n+3$, първите $n+4$ и т. н. всичките k члена на семейството имат обща точка. Да разгледаме семейството $B_1, B_2, \dots, B_{n+2}$. Нека допуснем, че тези множества нямат обща точка. За всяко $i = 1, 2, \dots, n+2$ образуваме подсемейството $B_1, B_2, \dots, B_{i-1}, B_{i+1}, \dots, B_{n+2}$ (т. е. цялото интересуващо ни семейство без B_i). Това подсемейство има $n+1$ члена и съгласно условието на теоремата съществува точка x_i, обща за B_m, $m = 1, 2, \dots, i-1, i+1, \dots, n+2$. Като намерим такава точка x_i за всяко $i = 1, 2, \dots, n+2$, получаваме множеството $C = \{x_1, x_2, \dots, x_{n+2}\}$. Съществено е да отбележим, че всяко множество B_i със сигурност съдържа различните от x_i точки на C и че точките на C са различни.

За множеството C може да приложим теоремата на Радон. Нека $x \in \operatorname{co} C_1 \cap \operatorname{co} C_2$, където $C_1 \cap C_2 = \emptyset$ и $C_1 \cup C_2 = C$. Ще докажем, че точката x е обща за семейството от множества $B_1, B_2, \dots, B_{n+2}$. Наистина нека B_p е произволен член на това семейство. Точката x_p принадлежи точно на едно от множествата C_1, C_2. Нека x_p е от C_1. Тогава B_p, съдържайки всички различни от x_p точки на C, ще съдържа и цялото C_2. Следователно изпъкналото множество B_p ще съдържа и точката $x \in \operatorname{co} C_2$. Теоремата е доказана. □

Упражнение 1. Формулирайте теоремата на Хели при $n = 1, 2, 3$.

Упражнение 2. Намерете пример, показващ че теоремата на Хели не е в сила за неизпъкнали множества.

Следствие 1. Нека $\{B_t\}$ е произволно (не обезателно крайно) семейство от затворени и изпъкнали подмножества на $\mathbb{R}^n$, всеки $n+1$ члена на което имат обща точка. Нека поне едно от множествата B_t е ограничено. Тогава съществува точка, която принадлежи на всички множества от семейството.

Доказателство. От теоремата на Хели виждаме, че всяко крайно подмножество на семейството $\{B_t\}$ има обща точка. Нека B_{t_0} е ограничено. Тогава семейството $\{B_t \cap B_{t_0}\}$ се състои от ограничени и затворени множества и освен това произволно крайно негово подсемейство има непразно сечение. От съображения за компактност следва, че има точка, лежаща във всички множества $B_t \cap B_{t_0}$, т. е. и в B_t. $\square$

Задача 1. В равнината $\mathbb{R}^2$ са дадени $k \geq 3$ точки, като всеки три от тях се съдържат в кръг с радиус 1. Докажете, че съществува кръг с радиус 1, съдържащ всички дадени точки.

Задача 2 (теорема на Юнг)**.** В равнината са дадени $k \geq 3$ точки, разстоянието между всеки две от които не надминава 1. Докажете, че съществува кръг с радиус $\frac{1}{\sqrt{3}}$, съдържащ всичките k точки.

Задача 3. В равнината е дадена права l и семейство от краен брой изпъкнали множества, всеки две от които имат обща точка. Да се докаже, че някоя успоредна на l права пресича всяко от дадените множества.

Задача 4. Дадени са равнинно изпъкнало множество A и семейство от $k > 3$ полуравнини, обединението на които покрива A. Докажете, че съществува подсемейство от три полуравнини, които също покриват A.

Нека C е множество в равнината. Ще казваме, че C' е *транслация* на C, ако съществува точка x такава, че $C' = x + C = \{z : z = x + c,\ c \in C\}$.

Задача 5. Нека $C, A_1, A_2, \ldots, A_k$, $k \geq 3$, са изпъкнали множества в равнината, като за всеки три от множествата $\{A_i\}_{i=1}^k$ съществува транслация на C, която се съдържа във всяко от трите множества. Докажете, че съществува транслация на C, която се съдържа във всички множества A_i, $i = 1, 2, \ldots, k$.

Задача 6. Нека изпъкналите равнинни множества $A_1, A_2, \ldots, A_k$, $k \geq 3$, са такива, че всеки три от тях се съдържат в транслация на изпъкналото множество C. Докажете, че съществува транслация на C, съдържаща всички множества A_i, $i = 1, 2, \ldots, k$.

Задача 7. Нека $A_1, A_2, \ldots, A_k$ са равнинни изпъкнали множества и нека за всеки три от тях съществува транслация на изпъкналото множество C, пресичаща всяко от трите множества. Тогава съществува транслация на C, пресичаща всички множества A_i, $i = 1, 2, \ldots, k$.

4. Сума на множество и умножение на множество с число

Определение 1. Нека X и Y са две (необезателно изпъкнали) множества. Под тяхна *сума* (в смисъл на Минковски) се разбира множеството от всички точки от вида $z = a + b$, където a е точка от X, а b — точка от Y, т. е.

$$X + Y = \{z : z = a + b, a \in X, b \in Y\}.$$

В случая, когато едното от двете множества се състои само от една точка — например $X = \{a\}$, за сумата $X + Y$ получаваме $X + Y = a + Y = \{z : z = a + b, b \in Y\}$. Това е транслацията на множеството Y, определена от вектора a. Следователно

$$X + Y = \cup\{b + X : b \in Y\} = \cup\{Y + a : a \in X\}.$$

На фиг. 10 е показана сумата на следните равнинни множества: две перпендикулярни отсечки (фиг. 10, а); квадрат и кръг с центрове в O (фиг. 10, б); триъгълник OMN и отсечка OP (фиг. 10, в).

Упражнение 1. Ако X и Y са изпъкнали множества, то тяхната сума $X + Y$ също е изпъкнало множество.

Определение 2. Нека X е множество, а λ — число. С λX ще означаваме множеството $\{z : z = \lambda a, a \in X\}$.

Ще отбележим, че ако X е изпъкнало множество, то λX също е изпъкнало множество. Не е трудно да се въведе и понятието *разлика на две множества*. С $X - Y$ ще означаваме множеството $X + (-1)Y$. Ясно е, че ако X и Y са изпъкнали множества, то $X - Y$ също е изпъкнало множество. Читателят сам ще се убеди, че свойствата на операциите сума на множества и умножение на множество с число наподобяват свойствата на съответните операции с вектори:

1) $X + Y = Y + X$;

2) $X + (Y + Z) = (X + Y) + Z$;

3) $(\lambda\mu)X = \lambda(\mu X) = \mu(\lambda X)$;

4) $\lambda(X + Y) = \lambda X + \lambda Z$.

Редно е обаче да обърнем внимание на някои особености при операциите с множества. Равенството $X = -X$ е изпълнено за всяко симетрично относно нулата множество, докато при вектори равенството $a = -a$ е изпълнено само за $a = 0$. Множеството $X - X$ в общия случай не се състои само от вектора нула. При използване на алгебрични операции с множества възникват и други затруднения, на които няма да се спираме.

Теорема 1. Сумата на две затворени множества, поне едното от които е ограничено, е затворено множество.

Доказателство. Нека $X + Y = Z$, X и Y са затворени и X е ограничено. Да разгледаме редицата $\{z_k\}$, $z_k \in Z$, $z_k \to z$. Ще покажем, че $z \in Z$.

По определение $z_k = x_k + y_k$, където $x_k \in X$, $y_k \in Y$, $k = 1, 2, \ldots$ Тъй като редицата $\{x_k\}$ е ограничена, от нея може да се избере сходяща подредица $\{x_{k_i}\}$, клоняща към x, където $x \in X$. Тогава $y_{k_i} = z_{k_i} - x_{k_i}$ ще клони към $y = z - x \in Y$. Получаваме $z = x + y$, $x \in X$, $y \in Y$, т. е. $z \in X + Y$. □

Следният пример показва, че сумата на две затворени множества не винаги е затворено множество:

Пример 1. Разглеждаме следните затворени множества в $\mathbb{R}^2$ (фиг. 11):

$$X = \left\{(x_1, x_2), x_2 \geq \frac{1}{x_1}, x_1 > 0\right\}, \qquad Y = \left\{(x_1, x_2), x_2 \leq -\frac{1}{x_1}, x_1 > 0\right\}.$$

Точката $\left(\frac{1}{k}, 0\right)$ принадлежи на $X + Y$ за всяко $k = 1, 2, \ldots$ Нейната граница $(0, 0)$ обаче не е точка от $X + Y$.

Задача 1. Докажете, че $\overline{X} + \overline{Y} \subset \overline{X + Y}$.

Задача 2. Да се докаже, че сумата на две отворени множества е отворено множество.

Задача 3. Да се докаже, че ако λ и μ са неотрицателни числа и X е изпъкнало множество, то $(\lambda + \mu)X = \lambda X + \mu X$. Вярно ли е това твърдение, ако X не е изпъкнало множество или ако числата λ и μ могат да бъдат произволни, а не само неотрицателни?

Задача 4. Нека X и Y са изпъкнали множества. Докажете, че множеството $\{x : (x + Y) \cap X \neq \emptyset\}$ е изпъкнало. Установете, че това множество съвпада с множеството $X - Y$.

Задача 5. Нека X_1 и X_2 са изпъкнали многоъгълници в равнината с периметри съответно l_1 и l_2 . Докажете, че $X_1 + X_2$ е многоъгълник е периметър $l_1 + l_2$.

Задача 6 (теорема на Брун—Минковски)**.** Да се докаже, че за всеки две изпъкнали равнинни множества X и Y е в сила неравенството $S_{\lambda X + \mu Y} \geq (\lambda\sqrt{S_X} + \mu\sqrt{S_Y})^2$, където с S_E е означено лицето на изпъкналото множество E.

Задача 7. Да се докаже, че всеки изпъкнал многоъгълник в равнината може да се представи като сума на триъгълници.

5. Проекция на точка върху множество

Определение 1. Точката $p \in X$ се нарича *проекция* на точката y върху множеството X, ако $\|y - p\| = \inf\{\|x - y\| : x \in X\}$.

Проекцията на точката y върху X е решение на минимизационната задача

$$\|y - x\| \to \min, \quad x \in X,$$

която ще наричаме *задача за проекцията*. Числото

$$d(y, X) = \inf\{\|y - x\| : x \in X\}$$

се нарича *разстояние* между точката y и множеството X. Ако за точката y съществува проекция p върху X, тогава $d(y, X) = \|y - p\|$. Ако $y \in X$, тогава $d(y, X) = 0$ и y съвпада с проекцията си върху X.

Следващите теореми дават отговор на основните въпроси, свързани със съществуването, единствеността и характеризацията на решението на задачата за проекцията. Първата от тях е следствие от теоремата на Вайерщрас (вж. следствие 1, § 1, гл. 1).

Теорема 1 (съществуване на решение)**.** Нека множеството X е непразно и затворено. Тогава всяка точка $y \in \mathbb{R}^n$ има проекция върху X.

Теорема 2 (единственост на решението)**.** Нека X е изпъкнало множество. Тогава всяка точка $y \in \mathbb{R}^n$ има не повече от една проекция върху X.

Доказателство. Да предположим че точката $y \notin X$ има две проекции p' и p'' върху изпъкнало множество X, $p' \neq p''$. Точката $z = \frac{1}{2}(p' + p'')$ също принадлежи на X. Разглеждаме триъгълника с върхове y, p' и p'' (фиг. 12). Получаваме

$$\|y - z\| = \left\|y - \frac{1}{2}(p' + p'')\right\| \leq \frac{1}{2}(\|y - p'\| + \|y - p''\|) = \|y - p'\| = \|y - p''\|.$$

От друга страна,

$$\|y - z\| \geq \|y - p'\| = \|y - p''\|.$$

От това следва, че по-горното неравенство (на триъгълника) е изпълнено като равенство, следователно векторите $y - p'$ и $y - p''$ са положително колинеарни (пропорционални), т. е. $y - p' = k(y - p'')$, $k > 0$. От $\|y - p'\| = \|y - p''\|$ следва, че $k = 1$, или $y - p' = y - p''$. Но това означава, че $p' = p''$. Полученото противоречие доказва теоремата. □

Следствие 1. Нека X е изпъкнало, затворено и непразно множество в $\mathbb{R}^n$. Тогава за всяко $y \in \mathbb{R}^n$ съществува единствена точка p, която е проекция на y върху X.

Теорема 3 (характеризиране на проекцията)**.** Нека $X \subset \mathbb{R}^n$ е изпъкнало множество и $y \in \mathbb{R}^n$. Точката $p \in X$ е проекция на y върху X тогава и само тогава, когато за всяко $x \in X$ е изпълнено

$$\langle x - p, y - p\rangle \leq 0. \tag{1}$$

Геометрично това условие означава, че ъгълът между векторите $y - p$ и $x - p$ за произволно $x \in X$ не трябва да е остър (фиг. 13).

Доказателство. *Необходимост*. Нека $p \in X$ е проекция на y върху X и нека $x \in X$. От изпъкналостта на X следва, че $z = \lambda x + (1 - \lambda)p \in X$ за всяко $\lambda \in (0, 1]$. Тогава

$$\|z - y\|^2 = \|\lambda(x - p) + p - y\|^2 = \lambda^2\|x - p\|^2 + \|p - y\|^2 + 2\lambda\langle x - p, p - y\rangle.$$

От дефиницията на проекция имаме, че $\|z - y\| \geq \|p - y\|$, следователно

$$\lambda^2\|x - p\|^2 + 2\lambda\langle x - p, p - y\rangle \geq 0$$

или

$$\lambda\|x - p\|^2 \geq 2\langle x - p, y - p\rangle$$

за всяко достатъчно малко число $\lambda > 0$. Но това означава, че е изпълнено условието (1)

Достатъчност. Нека за точката $p \in X$ е изпълнено условието (1) за всяко $x \in X$. Тогава

$$\|x - y\|^2 = \|x - p + p - y\|^2 = \|x - p\|^2 + \|p - y\|^2 + 2\langle x - p, p - y\rangle \geq \|p - y\|^2.$$

Полученото неравенство означава, че p е проекция на y върху X. □

Задача 1. Нека X е изпъкнало и затворено множество и p е проекция на y върху X. Да се докаже, че за всяко $x \in X$ са изпълнени неравенствата

$$\langle x - y, x - p\rangle \geq \|x - p\|^2, \qquad \|x - y\| \geq \|x - p\|.$$

Задача 2. Нека X е множество в $\mathbb{R}^n$. Докажете, че за всеки две точки $y, z \in \mathbb{R}^n$ е в сила неравенството

$$|d(y, X) - d(z, X)| \leq \|y - z\|.$$

Задача 3. Нека x_1 и x_2 са проекции на точките y_1 и y_2 върху изпъкнало множество X. Да се докаже, че

$$\langle y_1 - y_2, x_1 - x_2\rangle \geq 0.$$

6. Отделимост на изпъкнали множества

Нека X и Y са две множества в равнината $\mathbb{R}^2$. Интересува ни въпросът, кога между тези две множества може да се прекара такава права линия l, че множеството X да лежи изцяло в една полуравнина, определена от правата l, а множеството Y — в другата полуравнина (фиг. 14).

Ако за множествата X и Y съществува права линия l с описаното свойство, казваме, че X и Y са отделими.

Ако X и Y са подмножества на пространството $\mathbb{R}^3$, естествено е да се разгледа въпросът за съществуване на равнина, която да разделя множествата X и Y, т. е. X да лежи в едното полупространство, определено от полуравнината, а Y — в другото полупространство.

В най-обща форма за подмножества на $\mathbb{R}^n$ постановката на въпроса за отделимост на множества се дава чрез следната дефиниция:

Определение 1. Ще казваме,че подмножествата X и Y на $\mathbb{R}^n$ са отделими, ако съществуват ненулев вектор $v \in \mathbb{R}^n$ и числото t, за които неравенствата $\langle x, v\rangle \leq t \leq \langle y, v\rangle$ са изпълнени за всяка точка $x \in X$ и всяка точка $y \in Y$.

Ще припомним, че *хиперравнина* е множество от вида $M = \{x \in \mathbb{R}^n : \langle v, x\rangle = t\}$, където v е ненулев вектор.

Забележка. Без загуба на общност можем да приемем, че $||v|| = 1$. Действително, ако $||v|| \neq 1$, то можем да заместим v с $\frac{v}{||v||}$ и t с $\frac{t}{||v||}$.

Всяка хиперравнина определя две отворени полупространства:

$$H_+ = \{x \in \mathbb{R}^n : \langle v, x\rangle \geq t\} \quad \text{и} \quad H_- = \{x \in \mathbb{R}^n : \langle v, x\rangle \leq t\}.$$

Следователно две множества X и Y са отделими тогава и само тогава, когато съществува хиперравнина H, за която $X \subset H_-$ и $Y \subset H_+$. В такъв случай казваме, че *хиперравнината H разделя множествата X и Y*.

Пример 1. Ако X и Y са различни едноточкови множества, $X = \{x_0\}$, $Y = \{y_0\}$, то те са отделими. Достатъчно е да вземем хиперравнината, която минава през средата на отсечката с краища x_0 и y_0 и е перпендикулярна на тази отсечка. Същото може да се изрази и така: полагаме $v = y_0 - x_0$ и $t = \frac{1}{2}\langle v, x_0 + y_0\rangle$. Тогава $\langle v, y_0\rangle - t = \frac{1}{2}\langle v, v\rangle > 0$ и $t - \langle v, x_0\rangle = \frac{1}{2}\langle v, v\rangle > 0$. Следователно $\langle v, x_0\rangle < t < \langle v, y_0\rangle$.

Пример 2. Нека X е кръг в $\mathbb{R}^2$ с център $O = (0, 0)$ и радиус 1. Нека Y също е кръг в $\mathbb{R}^2$ с радиус 1, но с център в точката $b_0 = (\sqrt{2}, \sqrt{2})$ (фиг. 15). От геометрични съображения е ясно, че тези множества могат да бъдат отделени с хиперравнина (в $\mathbb{R}^2$ хиперравнините са прави линии). Съответната разделяща линия може да се намери например със следните разсъждения:

Нека $v = \frac{1}{2}b_0 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ и $t = 1$. Като вземем предвид, че $\|v\| = 1$ и че за всяко $x \in X$ $\|x\| \le 1$, получаваме $|\langle v, x\rangle| \le \|v\|\|x\| \le 1$, т. е. $-1 \le \langle v, x\rangle \le 1$ за $x \in X$. Да разгледаме множеството, което се получава от X чрез транслация, определена от вектора $b_0 : Y = b_0 + X$. Тогава за всяко $y \in Y$ ще се намери $x \in X$, за което $y = b_0 + x$. В такъв случай $\langle v, y\rangle = \langle v, b_0\rangle + \langle v, x\rangle = 2 + \langle v, x\rangle \ge 1$, или $\langle v, x\rangle \le 1 \le \langle v, y\rangle$, когато $x \in X$ и $y \in Y$. Следователно разделящата права е $\{z : \langle v, z\rangle = t\}$.

Примери на неотделими множества могат да се дадат много лесно. Очевидно е, че два концентрични кръга не могат да се отделят с права линия. По-общо казано, ако две множества имат „голяма обща част“ (например, ако съдържат един и същ кръг), то те не могат да бъдат отделени с права линия. Може да се случи обаче двете множества да нямат общи точки и въпреки това да са неотделими. Такива са множествата от фиг. 16.

Целта на по-нататъшните разглеждания в този параграф е да се убедим, че за изпъкнали множества тази аномалия не възниква, т. е. ще докажем, че всеки две непресичащи се изпъкнали множества са отделими.

Определение 2. Ще казваме, че множествата X и Y от $\mathbb{R}^n$ са *силно отделими*, ако съществуват вектор v и числата t_1, t_2 такива, че за всяко $x \in X$ и $y \in Y$ са изпълнени неравенствата

$$\langle v, x\rangle \le t_1 < t_2 \le \langle v, y\rangle.$$

На фиг. 17 са дадени множества, които са силно отделими (фиг. 17, а) и отделими, но не силно отделими (фиг. 17, б).

Теорема 1. Нека X е затворено, изпъкнало и непразно множество и $y \notin X$. Тогава X и $\{y\}$ са силно отделими.

Доказателство. Нека p е проекцията на y върху X. Полагаме $v = y - p$. Очевидно $\|v\| = \|y - p\| > 0$. Нека $t_1 = \langle v, p\rangle$ и $t_2 = \langle v, y\rangle$. Тогава

$$t_2 - t_1 = \langle v, y - p\rangle = \|y - p\|^2 > 0.$$

Съгласно теорема 3 от § 5, характеризираща проекцията, твърдението „p е проекция на y върху X“ е еквивалентно с неравенството

$$\langle v, x\rangle = \langle y - p, x\rangle \le \langle y - p, p\rangle = t_1$$

за всяко $x \in X$. Следователно $\langle v, x\rangle \le t_1 < t_2 = \langle v, y\rangle$ за всяко $x \in X$. Това неравенство показва, че множеството X и точката y са силно отделими. □

Хиперравнините $H_1 = \{x \in \mathbb{R}^n : \langle v, x\rangle = t_1\}$ и $H_2 = \{x \in \mathbb{R}^n : \langle v, x\rangle = t_1\}$, изобразени на фиг. 18, „силно отделят“ множеството X от точката y.

Лема 1. Множествата X и Y са отделими (силно отделими) тогава и само тогава, когато множеството $X - Y$ и нулата $\{0\}$ са отделими (силно отделими).

Доказателство. Ако множествата X и Y са отделими (силно отделими), то са в сила неравенствата

$$\langle v, x\rangle \le t \le \langle v, y\rangle,\ \ x \in X,\ \ y \in Y$$
$$(\langle v, x\rangle \le t_1 < t_2 \le \langle v, y\rangle,\ \ x \in X,\ \ y \in Y).$$

Тогава за всяко $z = x - y \in X - Y$, $x \in X$, $y \in Y$, са изпълнени неравенствата

$$\langle v, z\rangle = \langle v, x - y\rangle = \langle v, x\rangle - \langle v, y\rangle \le 0 \tag{1}$$
$$(\langle v, x\rangle - \langle v, y\rangle \le t_1 - \langle v, y\rangle \le t_1 - t_2 < 0),$$

които показват, че $X - Y$ и $\{0\}$ са отделими (силно отделими) множества.

Обратно, ако $X - Y$ и $\{0\}$ са отделими (силно отделими) множества, то неравенствата (1) са в сила. Не е трудно да забележим, че в този случай са изпълнени и неравенствата

$$\sup\{(v, x) : x \in X\} - \inf\{(v, y) : y \in Y\} \le 0$$
$$(\sup\{(v, x) : x \in X\} - \inf\{(v, y) : y \in Y\} \le t_1 - t_2 < 0).$$

Оттук непосредствено се вижда, че множествата X и Y са отделими (силно отделими). □

Теорема 2. Нека X и Y са непразни, затворени и изпъкнали множества, които нямат обща точка. Нека Y е ограничено множество. Тогава X и Y са силно отделими.

Доказателство. Множеството $X - Y$ е изпъкнало и затворено (вж. § 4). Освен това и $0 \notin X - Y$, защото би излязло, че X и Y имат обща точка. Тогава силната отделимост следва от лема 1 и теорема 1. □

Условието за ограниченост на едното от двете множества X и Y е съществено за верността на теоремата, тъй като в противен случай разликата $X - Y$ може да не е затворено множество (вж. теорема 1 от § 4). В това се убеждаваме с помощта на фиг. 19, където са изобразени две неограничени и непресичащи се подмножества на равнината, които не са строго отделими:

$$X = \left\{(x_1, x_2) : x_2 \ge \frac{1}{x_1},\ x_1 > 0\right\}, \qquad Y = \{(x_1, x_2) : x_2 \le 0\}.$$

Упражнение 1. Нека $\langle v, x\rangle \ge 0$ за всяко x от кълбо с център O и радиус $\delta > 0$. Докажете, че $v = 0$.

Упражнение 2. Нека $v \neq 0$ и $\langle v, x\rangle \geq 0$ за всяко $x \in X$. Докажете, че ако кълбо с център x_0 и радиус $\delta > 0$ се съдържа в X, то $\langle v, x_0\rangle > 0$.

Задача 1 (Теорема на Хели за равнината)**.** Като използвате теоремата за отделимост, докажете следното твърдение: Ако всеки три от изпъкналите, затворени и ограничени подмножества $A_1, A_2, \dots, A_k$ на равнината имат обща точка, то сечението на тези множества не е празно.

Задача 2. Нека X и Y са непразни, затворени и ограничени подмножества на равнината $\mathbb{R}^2$. Нека е известно, че произволно множество T, състоящо се от три или по-малко точки на Y, е силно отделимо от X. Докажете, че X и Y са силно отделими.

Задача 3. Посочете двойка затворени множества X и Y в равнината такива, че: а) Y е ограничено; б) $\operatorname{co} X \cap \operatorname{co} Y = \emptyset$; в) X и Y не са силно отделими.

Задача 4. Нека X е подмножество на $\mathbb{R}^2$, за което да е изпълнено следното свойство: за всяка точка x_0 извън X съществува права линия, която силно отделя x_0 от X. Докажете, че X е изпъкнало множество.

Задача 5 (Теорема на Дубовски—Милютин)**.** Нека $X_1, X_2, \dots, X_m$ са изпъкнали и затворени множества такива, че $\bigcap\limits_{i=1}^{m} X_i = \emptyset$. Тогава съществуват вектори $c_1, c_2, \dots, c_m$, не всички равни на нула, и числа $\lambda_1, \lambda_2, \dots, \lambda_m$ такива, че

$$\langle c_i, x\rangle \leq \lambda_i \text{ за } x \in X_i,\ i = 1, 2, \dots, m,$$

и

$$\sum_{i=1}^{m} c_i = 0, \qquad \sum_{\lambda=1}^{m} c_i = 0.$$

Задача 6 (Теорема на Киршбергер)**.** Нека в $\mathbb{R}^n$ са дадени две крайни множества от точки X и Y. Нека е известно, че за всяко множество $Z \subset X \cup Y$, състоящо се от най-много $n + 2$ точки, множествата $Z \cap X$ и $Z \cap Y$ са силно отделими. Докажете, че самите множества X и Y са силно отделими.

7. Опорни хиперравнини

Нека X е изпъкнало множество и $\operatorname{int} X \neq \emptyset$. Ще докажем, че през всяка точка $x_0 \in \overline{X} \setminus \operatorname{int} X = \partial X$ минава такава хиперравнина H, че множеството X лежи изцяло от едната ѝ страна. Можем да си мислим, че H „се допира“ до множеството X в точката x_0. Както ще видим по-нататък, съществуването на допирателна за X хиперравнина през всяка точка $x \in \partial X$ е от голямо значение за теорията на оптимирането.

Теорема 1. Нека X е изпъкнало множество с непразна вътрешност и нека $x_0 \in \partial X$. Тогава съществува такъв вектор $v_0 \neq 0$, че $\langle v_0, x\rangle \leq \langle v_0, x_0\rangle$ за всяко $x \in \overline{X}$, при това $\langle v_0, x\rangle < \langle v_0, x_0\rangle$ за всяко $x \in \operatorname{int} X$.

Доказателство. Тъй като $\partial X = \partial \overline{X}$ (следствие 2 от § 1, то $x_0 \in \partial X$. Следователно съществува редица y_n, всеки член на която не принадлежи на $\overline{X}$ и която клони към x_0 при $n \to \infty$. Съгласно теоремата за отделимост (вж. теорема 1, § 6) точката y_n може да се отдели (силно) от затвореното и изпъкнало множество $\overline{X}$ с хиперравнина, т. е. съществува вектор v_n, $\|v_n\| = 1$, такъв, че

$$(1) \qquad \langle v_n, x\rangle < \langle v_n, y_n\rangle \text{ за всяко } x \in \overline{X}.$$

Без ограничение на общността можем да си мислим, че редицата $\{v_n\}_{n=1}^{\infty}$ е сходяща и клони към v_0, $\|v_0\| = 1$. С граничен преход в (1) (при фиксирано $x \in \overline{X}$) получаваме неравенството

$$(2) \qquad \langle v_0, x\rangle \leq \langle v_0, x_0\rangle.$$

То е вярно за всяко $x \in \overline{X}$. Остава да докажем, че за всяка вътрешна за X точка неравенството е изпълнено строго. Нека $u_1 \in \operatorname{int} X$ и нека $\varepsilon > 0$ е такова число, че от $\|x - u_1\| < \varepsilon$ да следва $x \in X$. Ако допуснем, че $\langle v_0, u_1\rangle = \langle v_0, x_0\rangle$, то за точката $u_1 + \frac{\varepsilon}{2} v_0 \in X$ бихме имали $\left\langle v_0, u_1 + \frac{\varepsilon}{2} v_0\right\rangle = \langle v_0, u_1\rangle + \frac{\varepsilon}{2} = \langle v_0, x_0\rangle + \frac{\varepsilon}{2} > \langle v_0, x_0\rangle$, което противоречи на (2). Теоремата е доказана. □

Определение 1. Нека X е множество, $v \neq 0$ е вектор от $\mathbb{R}^n$ и $\sup\limits_{z\in X}\langle v, z\rangle < \infty$. Множеството $H = \left\{x \in \mathbb{R}^n : \langle v, x\rangle = \sup\limits_{z\in X}\langle v, z\rangle\right\}$ се нарича *опорна хиперравнина* за множеството X, определена от посоката v. Множеството $H_- = \{x \in \mathbb{R}^n : \langle v, x\rangle \leq \sup\limits_{z\in X}\langle v, z\rangle\}$ ще наричаме *опорно за X полупространство*, съответстващо на (определено от) посоката v.

Следствие 1. Нека X е изпъкнало множество с непразна вътрешност и нека $x_0 \in \partial X$. Тогава съществува опорна за X хиперравнина, която съдържа точката x_0.

Доказателство. От теорема 1 знаем, че съществува вектор v_0, $\|v_0\| = 1$, за който $\langle v_0, x\rangle \leq \langle v_0, x_0\rangle$ при всеки избор на $x \in \overline{X}$. От $x_0 \in \overline{X}$ следва, че x_0 е граница на редица $\{x_n\}_{n\geq 1} \subset X$. Тогава

$$\langle v_0, x_0\rangle = \lim_{n\to\infty}\langle v_0, x_n\rangle \leq \sup_{x\in X}\langle v_0, x\rangle \leq \langle v_0, x_0\rangle.$$

Следствието е доказано, защото хиперравнината $H = \{x \in \mathbb{R}^n : \langle v_0, x\rangle = \sup\limits_{z\in X}\langle v_0, z\rangle = \langle v_0, x_0\rangle\}$ съдържа x_0 и е опорна за X. □

Забележка. В следващия параграф ще се убедим, че за да съществува опорна хиперравнина на X, не е необходимо X да има непразна вътрешност.

Теорема 2. Всяко затворено и изпъкнало подмножество X на $\mathbb{R}^n$ може да се представи като сечение на своите опорни полупространства.

Доказателство. Нека M е множеството от всички опорни за X полупространства. Очевидно $X \subset \bigcap\limits_{H_- \in M} H_-$. Ако $x_0 \notin X$, то съществуват v_0, $\|v_0\| = 1$, и $\varepsilon > 0$, за които $\langle v_0, x\rangle < \langle v_0, x_0\rangle - \varepsilon$ за всяко $x \in X$. Тогава $H_-(v_0) = \{x \in \mathbb{R}^n : \langle v_0, x\rangle \leq \sup\limits_{z \in X}\langle v_0, z\rangle\}$ е опорно полупространство, което съдържа X, но не съдържа x_0. Теоремата е доказана. □

8. Размерност на изпъкнали множества

Определение 1. Нека L е линейно подпространство в $\mathbb{R}^n$ и нека x_0 е точка в $\mathbb{R}^n$. Множество от вида $x_0 + L$ се нарича *афинно многообразие*.

Упражнение 1. Да се докаже, че ако $x_1 \in x_0 + L$, то $x_0 + L = x_1 + L$.

Определение 2. Нека k е цяло положително число. Ще казваме, че афинното многообразие $x+L$, където L е подпространство на $\mathbb{R}^n$, има *размерност* k (означаваме $\dim(x + L) = k$), ако в L съществува система от k линейно независими вектора и всеки $k + 1$ (или повече) вектора на L са линейно независими.

С други думи, размерността на афинното многообразие $x+L$ се определя от размерността на линейното пространство L.

Определение 3. *Афинна обвивка* $\operatorname{aff} X$ на множеството $X \subset \mathbb{R}^n$ е афинното многообразие с най-малка размерност, съдържащо X, т. е. $\operatorname{aff} X = \cap\{A : A$ — афинно многообразие, $X \subset A\}$.

Определение 4. *Размерност* на множеството X наричаме числото $\dim \operatorname{aff} X$ и означаваме $\dim X$.

Пример 1. Единичното кълбо

$$B(0, 1) = \{x \in \mathbb{R}^n : \|x\| < 1\}$$

има размерност n и неговата афинна обвивка е $\mathbb{R}^n$. Множеството

$$B_2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 < 1, x_3 = 0\right\}$$

има размерност 2 и всяка негова точка е контурна, $\operatorname{aff} B_2 = \mathbb{R}^2$. Ако разгледаме B_2 като кълбо в $\mathbb{R}^2$:

$$B_2 = B_2(0, 1) = \left\{(x_1, x_2) \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} < 1\right\},$$

това множество ще има вътрешни точки (фиг. 20).

Упражнение 2. Да се докаже, че ако $x \in X$, то $\operatorname{aff}(X - x)$ е линейно подпространство.

Упражнение 3. Да се докаже, че

$$\operatorname{aff} X = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{s} \lambda_i x_i,\ x_i \in X,\ \sum_{i=1}^{s} \lambda_i = 1,\ s = 1, 2, \dots \right\}.$$

Упражнение 4. Да се докаже, че ако $X \subset Y$, то $\operatorname{aff} X \subset \operatorname{aff} Y$.

Упражнение 5. Да се докаже, че $\operatorname{aff}(X + Y) = \operatorname{aff} X + \operatorname{aff} Y$.

Упражнение 6. Да се докаже, че ако $\dim X = m$ и $x \in X$, то съществуват вектори $x^1, x^2, \dots, x^m \in X$ такива, че $x^1 - x, x^2 - x, \dots, x^m - x$ са линейно независими. Да припомним, че точката $x \in X$ се нарича *относително вътрешна* за множеството X, ако съществува кълбо $B(x, \varepsilon)$ такова, че $B(x, \varepsilon) \cap \operatorname{aff} X \subset X$. Множеството от относително вътрешни точки за X се нарича *относителна вътрешност* на X и се означава с $\operatorname{ri} X$. Точката $x \in \mathbb{R}^n$ се нарича *истински контурна* за X, ако тя е контурна за X и не е относително вътрешна. Множеството от всички истински контурни точки ще бележим с $\operatorname{r}\partial X$.

В пример 1 кълбото $B_2(0, 1)$ съвпада с $\operatorname{ri} B_2(0, 1)$, а точката $(1, 0, 0)$ е истински контурна точка на $B_2(0, 1)$.

Теорема 1. Всяка непразно изпъкнало множество X има непразна относителна вътрешност.

Доказателство. Нека $\dim X = m$. Без ограничение на общността можем да считаме, че X съдържа нулата. Нека $x^1, x^2, \dots, x^m$ са линейно независими вектори на X. Да разгледаме точката $x^* = \frac{1}{2m} \sum\limits_{i=1}^{m} x^i$. Ще докажем, че $x^* \in \operatorname{ri} X$. Най-напред да отбележим, че всяка точка x от $\operatorname{aff} X$ се представя във вида $x = \sum\limits_{i=1}^{m} \alpha_i x^i$ по единствен начин. С $\Delta(x)$ ще означаваме числото $\sum\limits_{i=1}^{m} |\alpha_i|$.

Нека положим $\mu = \min \left\{ \left\| \sum\limits_{i=1}^{m} \alpha_i x^i \right\| : \sum\limits_{i=1}^{m} |\alpha_i| = 1 \right\}$. Не е трудно да се убедим, че $\mu > 0$. Ще докажем, че $B\left(x^*, \frac{\mu}{2m}\right) \cap \operatorname{aff} X$ се съдържа в X. Наистина нека $t \in B\left(x^*, \frac{\mu}{2m}\right) \cup \operatorname{aff} X$. Тогава $y = x^* + x$, където $\|x\| < \frac{\mu}{2m}$. Очевидно $x = y - x^* \in \operatorname{aff} X - \operatorname{aff} X = \operatorname{aff} X$. Следователно $x = \sum\limits_{i=1}^{m} \alpha_i x^i$.

От дефиницията на числото μ се вижда, че

$$\frac{\|x\|}{\Delta(x)} = \frac{\left\| \sum\limits_{i=1}^{m} \alpha_i x^i \right\|}{\Delta(x)} = \left\| \sum_{i=1}^{m} \frac{\alpha_i}{\Delta(x)} x^i \right\| \geq \mu.$$

Тогава $\Delta(x) \le \frac{1}{\mu}\|x\| < \frac{1}{2m}$. Това означава, че в представянето $y = x* + \sum_{i=1}^{m} \alpha_i x^i = \sum_{i=1}^{m} \left(\frac{1}{2m} + \alpha_i\right) x^i$ коефициентите $\frac{1}{2m} + \alpha_i$ са положителни. Освен това $\sum_{i=1}^{m} \left(\frac{1}{2m} + \alpha_i\right) \le \frac{1}{2} + \Delta(x) \le \frac{1}{2} + \frac{1}{2m} < 1$. Следователно y се представя като изпъкнала комбинация на точките $0, x^1, x^2, \dots, x^m$: $y = \lambda 0 + \sum_{i=1}^{m} \left(\frac{1}{2m} + \alpha_i\right) x^i$, където $\lambda = 1 - \sum_{i=1}^{m} \left(\frac{1}{2m} + \alpha_i\right)$. Следователно y е от изпъкналото множество X. Теоремата е доказана. □

За $\operatorname{ri} X$ и $\operatorname{r\partial} X$ са в сила твърдения, аналогични на твърденията за $\operatorname{int} X$ и ∂X.

Твърдение 1. Нека X е изпъкнало множество, $x_1 \in \overline{X}$, $x_2 \in \operatorname{ri} \overline{X}$ и $x_1 \ne x_2$. Тогава всяка точка y, $y \ne x_1$, от отсечката $\overline{x_1 x_2}$ е относително вътрешна за X.

Твърдение 2. Нека X е изпъкнало множество. Тогава

1) $\overline{X} = \operatorname{ri} \overline{X}$;

2) $\operatorname{ri} \overline{X} = \operatorname{ri} X$;

3) $\operatorname{r\partial} X = \operatorname{r\partial} \overline{X}$.

Твърдение 3. През всяка контурна точка на изпъкналото множество X минава опорна хиперравнина.

Доказателство. Ако $\operatorname{int} X \ne \emptyset$, прилагаме следствие 1 от § 7. В противен случай $\dim X < n$ и X се съдържа в хиперравнина, която е опорна за X. □

В по-нататъшните разглеждания ще използваме следната

Теорема 2. През всяка истински контурна точка на изпъкналото множество X минава опорна хиперравнина, която не съдържа цялото множество X.

Доказателство. Без ограничение на общността можем да считаме, че $\operatorname{aff} X$ съдържа 0. Тогава $\operatorname{aff} X = \mathbb{R}^k \subset \mathbb{R}^n$ и x^0 е контурна точка за изпъкналото множество X, което има вътрешна (относно $\mathbb{R}^k$) точка.

Прилагаме теорема 1 от § 7, от която следва, че съществува вектор $v_0 \in \operatorname{aff} X$ такъв, че $\langle v_0, x\rangle \le \langle v_0, x^0\rangle$ за всяко $x \in X$, като за точките от $\operatorname{ri} X$ неравенството е строго. Следователно хиперравнината (в $\mathbb{R}^n$) $\{x \in \mathbb{R}^n : \langle x, v_0\rangle = \langle x^0, v_0\rangle\}$ е опорна за X, минава през x^0 и не съдържа $\operatorname{ri} X \ne \emptyset$. □

Теорема 3. Нека X и Y са непразни изпъкнали и непресичащи се множества. Тогава те са отделими с хиперравнина.

Доказателство. Според лема 1 от § 6 достатъчно е да покажем, че $Z = X - Y$ и $\{0\}$ са отделими. Очевидно $0 \notin Z$. Възможни са два случая:

$0 \notin \overline{Z}$ и $0 \in \overline{Z} \setminus \operatorname{ri} Z$. В първия случай $\{0\}$ е ограничено, затворено и изпъкнало множество, непресичащо се с изпъкналото и затворено множество $\overline{Z}$, следователно според теорема 2 от § 6 X и Y са силно отделими. Във втория случай е достатъчно да приложим теорема 2. $\square$

Задача 1. Докажете твърденията:

1) Ако X е непразно изпъкнало множество, то $z \in \operatorname{ri} X$ тогава и само тогава, когато за всяко $x \in X$ съществува $\mu > 1$ такова, че $(1-\mu)x + \mu z \in X$.

2) Ако X и Y са изпъкнали множества, то $\overline{X} = \overline{Y}$ тогава и само тогава, когато $\operatorname{ri} X = \operatorname{ri} Y$.

3) Ако X е изпъкнало множество и $\lambda \in \mathbb{R}^1$, то $ri(\lambda X) = \lambda \operatorname{ri} X$.

4) Ако X и Y са изпъкнали множества, то $\operatorname{ri}(X+Y) = \operatorname{ri} X + \operatorname{ri} Y$.

5) Докажете, че ако X и Y са непразни изпъкнали множества, удовлетворяващи $\operatorname{ri} X \cap \operatorname{ri} Y = \emptyset$, то те са отделими с хиперравнина, несъдържаща поне едното от множествата.

Упътване. Покажете, че от $\operatorname{ri} X \cap \operatorname{ri} Y = \emptyset$ следва $0 \notin \operatorname{ri}(X - Y)$.

Глава 3

Представяне на изпъкналите множества

Нека X е компактно изпъкнало подмножество на $\mathbb{R}^n$. Един от въпросите, които се разглеждат в тази глава, е следният: за кои подмножества M на X е изпълнено $\operatorname{co} M = X$? С други думи, за кои подмножества $M \subset X$ множеството X се „възстановява“ чрез операцията „изпъкнала обвивка“. Такива множества M съществуват. Например за множеството M от всички контурни точки на X е в сила съотношението $\operatorname{co} M = X$. Интересен е обаче въпросът за намиране на най-малкото множество M, за което все още е в сила равенството $\operatorname{co} M = X$. В тази връзка ще отбележим, че ако точката x_0 е от X и множеството $X \setminus \{x_0\}$ е изпъкнало, то x_0 обезателно трябва да лежи в множеството M, ако искаме $\operatorname{co} M$ да съвпада с X. Наистина, ако M не съдържа x_0, то $M \subset X \setminus \{x_0\}$ и $\operatorname{co} M \subset \operatorname{co}(X \setminus \{x_0\}) = X \setminus \{x_0\}$. Следователно $\operatorname{co} M \neq X$. Всяка точка $x \in X$, за която $X \setminus \{x\}$ е изпъкнало множество, се нарича крайна (екстремна) точка за X. От казаното дотук личи, че ако $\operatorname{co} M = X$, то M съдържа всички крайни точки на X. Основният резултат в § 1 на тази глава е, че множеството $\widehat{X}$ от крайните точки на X е най-малкото множество, чиято изпъкнала обвивка съвпада с X.

Оказва се, че по подобен начин изглеждат нещата и при неограничените изпъкнали множества. Всеки заострен конус съвпада с изпъкналата обвивка на „крайните“ си лъчи (§ 3), а всяко затворено и изпъкнало множество с крайни точки е сума на $\operatorname{co} \widehat{X}$ и един заострен конус $K(X)$ (§ 4). Особен интерес както от теоретична, така и от практическа гледна точка представляват многостенните множества, тъй като имат краен брой крайни точки и крайни лъчи (§ 5). Това е твърде важно при решаването на редица задачи от специален тип — линейно оптимиране, квадратично оптимиране, хиперболично оптимиране и т. н.

1. Крайни точки. Теорема на Минковски

Определение 1. Нека X е изпъкнало и затворено множество в $\mathbb{R}^n$ и $x_0 \in X$. Ще казваме, че x_0 *не е крайна* за X, ако съществуват точки x', $x'' \in X$, $x' \neq x''$, и число λ, $0 < \lambda < 1$, за които $x_0 = \lambda x' + (1-\lambda)x''$. В противен случай (когато не съществуват такива точки x', x'' и число λ) ще казваме, че точката x_0 е *крайна* за множеството X.

Крайните точки се наричат още ъглови, екстремни или неразложими точки. Ако сме означили с X дадено множество, с $\widehat{X}$ ще означаваме множеството от крайните му точки.

Пример 1. Ако X е отсечка, т. е. $X = \overline{x_1 x_2}$, то $\widehat{X} = \{x_1, x_2\}$.

Пример 2. Единствените крайни точки на триъгълника са неговите върхове.

Пример 3. Ако $A = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$, то $\widehat{A} = \{x \in \mathbb{R}^n : \|x\| = 1\}$.

Упражнение 1. Докажете, че точката x_0 е крайна за изпъкналото множество X тогава и само тогава, когато множеството $X \setminus \{x_0\}$ е изпъкнало.

Упражнение 2. Нека x_0 е крайна точка за множеството X и нека a е вектор в $\mathbb{R}^n$. Докажете, че $x_0 + a$ е крайна за множеството $X + a$.

Теорема 1. Ако H е опорна хиперравнина за изпъкналото и затворено множество X, то крайните точки на множеството $X_1 = X \cap H$ (ако има такива) са крайни и за множеството X, т. е. $\widehat{X_1} \subset \widehat{X}$.

Доказателство. Нека $\widehat{x} \in \widehat{X_1}$, $\langle a, x\rangle = \alpha$ е уравнението на опорната хиперравнина H и $\langle a, x\rangle \leq \alpha$ за $x \in X$. Допускаме, че $\widehat{x}$ не е крайна точка за X. Тогава ще съществуват две различни точки x' и x'' от множеството X, такива че $\widehat{x} = (1-\lambda)x' + \lambda x''$ за някое $0 < \lambda < 1$. Но от

$$\alpha = \langle a, \widehat{x}\rangle = \langle a, (1-\lambda)x' + \lambda x''\rangle = (1-\lambda)\langle a, x'\rangle + \lambda\langle a, x''\rangle \leq \alpha$$

се вижда, че $\langle a, x'\rangle = \alpha$ и $\langle a, x''\rangle = \alpha$, т. е. $x' \in H$, $x'' \in H$ или $x', x'' \in X_1$, което е невъзможно, тъй като $\widehat{x}$ е крайна точка за X_1. Следователно $\widehat{x}$ е крайна за множеството X. □

Теорема 2. Непразно, изпъкнало и затворено множество X има крайна точка тогава и само тогава, когато не съдържа права линия.

Доказателство. *Необходимост.* Ако $\widehat{x} \in X$ е крайна точка за X и допуснем, че в X се съдържа правата $\{x_0 + bt : t \in \mathbb{R}^1\}$, то в X ще се съдържа и правата $\{\widehat{x} + bt : t \in \mathbb{R}^1\}$ (вж. твърдение 3 от § 1 на гл. 2), което противоречи на това, че $\widehat{x} \in \widehat{X}$, тъй като

$$x' = \widehat{x} + b \in X, \quad x'' = \widehat{x} - b \in X, \quad \text{а} \quad \widehat{x} = \frac{1}{2}(x' + x'').$$

Достатъчност. Доказателството ще извършим с индукция по размерността на множеството X. Едномерните изпъкнали множества, несъдържащи права линия, са отсечката и лъчът. И двете имат крайни точки. Нека твърдението е вярно за изпъкнали множества с размерност, по-малка от k. Ще покажем, че е вярно и за множества с размерност k. И така, нека $\dim X = k$. През произволна истинска контурна точка $\overline{x}$ на X (такава сигурно съществува!) прекарваме опорна хиперравнина H, несъдържаща изцяло множеството X. Множеството $X_1 = X \cap H$ е изпъкнало, затворено, не съдържа права линия и $\dim X_1 < k$. Съгласно индуктивното предположение X_1 има крайна точка, т. е. $\widehat{X_1} \neq \emptyset$. Оттук и от теорема 1 следва, че $\widehat{X} \neq \emptyset$. □

Следствие 1. Всяко непразно изпъкнало компактно множество има поне една крайна точка.

Следствие 2. Ако X е изпъкнало затворено множество, $\widehat{X} \neq \emptyset$, H е опорна хиперравнина за X и $Z = X \cap H$, то $\widehat{Z} \neq \emptyset$.

Доказателство. От $\widehat{X} \neq \emptyset$ следва, че X не съдържа права линия, а оттук и $Z = X \cap H \neq \emptyset$ не съдържа права и следователно $\widehat{Z} \neq \emptyset$. □

Лема 1. Всяка относително вътрешна точка на изпъкнало компактно множество X се представя като изпъкнала комбинация на две истински контурни точки.

Доказателство. Нека $x \in \operatorname{ri} X$. През точката x прекарваме права $\rho : \{x + bt : t \in \mathbb{R}^1\}$ такава, че $\rho \subset \operatorname{aff} X$. Нека $t_1 = \max\{t : x+bt \in X\}$, $t_2 = \max\{t : x-bt \in X\}$. Очевидно $0 < t_1, t_2 < \infty$, точките $x_1 = x + bt_1 \in \mathrm{r}\partial X$, $x_2 = x - bt_2 \in \mathrm{r}\partial X$ и

$$x = \frac{t_2}{t_1 + t_2} x_1 + \frac{t_1}{t_1 + t_2} x_2,$$

т. е. x е изпъкнала комбинация на две истински контурни точки. □

Теорема 3 (Минковски). Ако X е изпъкнало компактно множество, то $X = \operatorname{co} \widehat{X}$.

Доказателство. Очевидно $\operatorname{co} \widehat{X} \subset X$. Обратното включване ще докажем с индукция по размерността на множеството. За едномерните множества (отсечките) твърдението е очевидно. Нека твърдението е вярно за множество с размерност по-малка от k и $\dim X = k$. Нека $\overline{x}$ е истинска контурна точка на X. Прекарваме през нея опорна хиперравнина H, несъдържаща X. Множеството $X_1 = X \cap H$ е изпъкнало и компактно, $\overline{x} \in X_1$ и $\dim X_1 < k$. Следователно

$$\overline{x} = \sum_{i=1}^{s} \alpha_i \widehat{x_i}, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{s} \alpha_i = 1, \quad \widehat{x_i} \in \widehat{X_1}.$$

Но $\widehat{X}_1 \subset \widehat{X}$. Следователно $\overline{x}$ е изпъкнала комбинация на крайни точки на X. Нека $x \in \operatorname{ri} X$. Тогава съгласно лема 1

$$x = (1 - \lambda)\overline{x}_1 + \lambda\overline{x}_2,$$

където $0 \le \lambda \le 1$ и $\overline{x}_1, \overline{x}_2 \in \mathrm{r}\partial X$. Но

$$\overline{x}_1 = \sum_{i=1}^{s_1} \alpha_i \widehat{x_i}, \quad \alpha_i \ge 0, \quad \sum_{i=1}^{s_1} \alpha_i = 1, \quad \widehat{x_i} \in \widehat{X},$$
$$\overline{x}_2 = \sum_{i=1}^{s_2} \beta_i \widehat{y_i}, \quad \beta_i \ge 0, \quad \sum_{i=1}^{s_2} \beta_i = 1, \quad \widehat{y_i} \in \widehat{X},$$

така че

$$x = (1 - \lambda) \sum_{i=1}^{s_1} \alpha_i \widehat{x_i} + \lambda \sum_{i=1}^{s_2} \beta_i \widehat{y_i}$$

е изпъкнала комбинация на крайни точки на X. Следователно $x \in \operatorname{co} \widehat{X}$ или $X \subset \operatorname{co} \widehat{X}$, а оттук $X = \operatorname{co} \widehat{X}$. □

Твърдението, което доказахме, е особено полезно за характеризация на множества с краен брой крайни точки (фиг. 21, а). То престава да бъде вярно за изпъкнали затворени, но неограничени множества (фиг. 21, б).

Задача 1. Докажете, че ако X е изпъкнало компактно множество и $\dim X = k$, то в X има поне $k + 1$ крайни точки.

2. Конус

Определение 1. Множеството $K \subset \mathbb{R}^n$ наричаме *конус* (с връх в нулата), ако от $z \in K$ и $\lambda \ge 0$ следва $\lambda z \in K$, т. е. ако $z \in K$, то K съдържа лъча $\Lambda = \{\lambda z : \lambda \ge 0\}$.

С други думи, K е конус, ако $\lambda K \subset K$ за всяко $\lambda \ge 0$.

Всеки лъч Λ с начало върха 0 на конуса K и съдържащ се в K наричаме *лъч* на конуса K. Един конус ще наричаме *изпъкнал*, ако е изпъкнало множество. Очевидно е, че всяка опорна хиперравнина на конус минава през върха му, т. е. през точката 0. Оттук следва, че ако една опорна хиперравнина съдържа точката z, тя съдържа и лъча $\{\lambda z : \lambda \ge 0\}$.

Пример 1. Нека A е матрица с размери $m \times n$. Всяко от множествата

$$K = \{x \in \mathbb{R}^n : Ax \le 0\},$$
$$K = \{x \in \mathbb{R}^n : Ax = 0, \ x \ge 0\}$$

е изпъкнал и затворен конус.

Примери на конуси са дадени на фиг. 22 (фиг. 22, а — изпъкнал затворен конус, фиг. 22, б — неизпъкнал незатворен конус).

Твърдение 1. Необходимо и достатъчно условие множеството $K \subset \mathbb{R}^n$ да бъде изпъкнал конус е с всеки два свои елемента да съдържа и произволна тяхна неотрицателна линейна комбинация, т. е. ако $z_1, z_2 \in K$ и $\alpha, \beta \geq 0$, то $\alpha z_1 + \beta z_2 \in K$.

Доказателство. *Необходимост.* Нека K е изпъкнал конус и $z_1, z_2 \in K$, $\alpha \geq 0$, $\beta \geq 0$. Ако $\alpha = 0$, $\beta = 0$, то $\alpha z_1 + \beta z_2 = 0 \in K$. Ако $\alpha + \beta > 0$, то $\alpha z_1 + \beta z_2 = (\alpha + \beta)\left(\dfrac{\alpha}{\alpha + \beta} z_1 + \dfrac{\beta}{\alpha + \beta} z_2\right) \in K$.

Достатъчност. Нека за всеки $z_1, z_2 \in K$ и $\alpha \geq 0, \beta \geq 0$ имаме $\alpha z_1 + \beta z_2 \in K$. В частност K съдържа и изпъкналите комбинации на всеки две свои точки, т. е. K е изпъкнало множество. Множеството K съдържа и произведението на всяка своя точка с неотрицателно число, т. е. K е конус или изпъкнал конус. □

Упражнение 1. Докажете, че ако K е конус, то $\overline{K}$ е също конус.

Упражнение 2. Докажете, че сечението на изпъкнали затворени конуси е изпъкнал затворен конус.

Упражнение 3. Докажете, че ако K е конус и $0 \in \operatorname{ri} K$, то $K = \operatorname{aff} K$.

Упражнение 4. Нека K_1 и K_2 са изпъкнали конуси в $\mathbb{R}^n$. Докажете че $K_1 + K_2$ е изпъкнал конус.

Определение 2. Нека X е непразно множество в $\mathbb{R}^n$. Множеството

$$X^* = \{y \in \mathbb{R}^n : \langle y, x \rangle \leq 0 \text{ за всяко } x \in X\}$$

се нарича *спрегнат* или *полярен конус* на множеството X.

С X^{**} ще означаваме спрегнатия на спрегнатия конус на X.

Лема 1. Нека X и Y са непразни множества в $\mathbb{R}^n$. Тогава:

1) X^* е затворен изпъкнал конус;

2) $X \subset X^{**}$;

3) ако $X \subset Y$, то $Y^* \subset X^*$.

Доказателството за тези свойства оставяме за упражнение на читателя.

Пример 2. а) Ако X е подпространство на $\mathbb{R}^n$, то X^* е ортогоналното му допълнение. В частност $\{0\}^* = \mathbb{R}^n$ и $\{\mathbb{R}^n\}^* = \{0\}$.

б) Да разгледаме множеството

$$X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq x_2,\ x_2 \geq 0\}.$$

Това множество е изпъкнал затворен конус (фиг. 23). Неговият спрегнат конус е

$$X^* = \{x \in \mathbb{R}^2 : x_1 \leq 0,\ x_1 \leq -x_2\}.$$

Теорема 1. Нека K е непразен изпъкнал конус. Тогава $K^{**} = \overline{K}$.

Доказателство. Знаем, че $K \subset K^{**}$ и K^{**} е затворено множество (лема 1), следователно $\overline{K} \subset K^{**}$. Нека $x \in K^{**}$ и $x \notin \overline{K}$. Точката x и множеството $\overline{K}$ са силно отделими, следователно съществува такъв ненулев вектор v, че:

$$\langle v, x \rangle > \langle v, y \rangle \tag{1}$$

за всяко $y \in \overline{K}$. Тъй като $\overline{K}$ е конус (упражнение 1), то за всяко $\lambda > 0$ и $y \in \overline{K}$

$$\langle v, x \rangle > \lambda \langle v, y \rangle, \text{ т. е. } \frac{1}{\lambda}\langle v, x \rangle > \langle v, y \rangle,$$

откъдето следва $\langle v, y \rangle \leq 0$ за всяко $y \in \overline{K}$. Следователно $v \in \overline{K}^* \subset K^*$. Но от (1) и от $0 \in \overline{K}$ следва $\langle v, x \rangle > 0$, което противоречи на условието $x \in X^{**}$. □

Задача 1. Намерете спрегнатите конуси на следните множества:

а) $X = \{x \in \mathbb{R}^2 : x_1 \geq x_2,\ x_1 \leq 2x_2\}$;

б) $X = \{x \in \mathbb{R}^2 : x_2 \geq -|x_1|\}$.

Задача 2. Нека K е непразен изпъкнал и затворен конус. Докажете, че $v \in K^*$ тогава и само тогава, когато проекцията на v върху K е нулевият вектор.

Задача 3. Нека K е непразен изпъкнал и затворен конус в $\mathbb{R}^n$. Докажете, че $K + K^* = \mathbb{R}^n$.

Задача 4. Докажете, че една и само една от следните две системи е разрешима:

1) $Ax = b$, $x \geq 0$;

2) $A^T y \leq 0$, $\langle b, y \rangle > 0$.

3. Представяне на изпъкнал конус

Определение 1. Изпъкнал затворен конус наричаме *заострен*, ако има крайна точка.

Очевидно, ако даден изпъкнал конус има крайна точка, тя е единствена и е точно неговият връх (нулата). Лесно се вижда, че заостреният конус не съдържа едновременно точките z и $-z$, ако $z \neq 0$. От доказаната в § 1 теорема 2 следва, че даден конус е заострен тогава и само тогава, когато не съдържа права линия.

Твърдение 1. Изпъкнал затворен конус е заострен тогава и само тогава, когато съществува опорна хиперравнина през върха му, която няма други общи точки с конуса.

Доказателство. Нека K е заострен конус. Разглеждаме множеството $X = \{x \in K : \|x\| = 1\}$. То е компактно, следователно $\operatorname{co} X$ е също компактно и $0 \notin \operatorname{co} X$, защото нулата не е от X и е крайна за K. Тогава съществува хиперравнина H с уравнение $\langle v, x\rangle = 0$ такава, че $\langle v, x\rangle > 0$ за всяко $x \in X \subset \operatorname{co} X$. Ще покажем, че H няма други общи точки с конуса. Нека $y \neq 0$ е произволна точка на K. Тогава $z = \dfrac{y}{\|y\|} \in K$ и $\|z\| = 1$, т. е. $z \in X$. Следователно $\langle v, z\rangle = \dfrac{1}{\|y\|}\langle v, y\rangle > 0$, откъдето следва $\langle v, y\rangle > 0$, т. е. $y \notin H$. Обратно, ако съществува опорна хиперравнина $H : \langle v, x\rangle = 0$ през върха на K, която да няма други общи точки с него, то K не съдържа права и следователно е заострен. □

Определение 2. Лъч на изпъкнал затворен конус K ще наричаме *краен* (*екстремен*) за K, ако на всяка негова точка, различна от нулата, не може да се представи като положителна линейна комбинация на две различни точки на конуса, нележащи едновременно на този лъч.

Лесно се вижда, че ако една точка (различна от 0) от даден лъч има това свойство, то всички точки на лъча също го имат.

Множеството от точките на крайните лъчи на конуса K ще означаваме с $\overrightarrow{K}$. Ако Λ е краен лъч за K, ще означаваме това с $\Lambda \subset \overrightarrow{K}$.

Твърдение 2. Ако K е изпъкнал затворен конус, $H = \{z : \langle v, z\rangle = 0\}$ е неговата опорна хиперравнина и лъчът $\Lambda \subset \overrightarrow{K \cap H}$, то $\Lambda \subset \overrightarrow{K}$.

Доказателство. Нека $\langle v, z\rangle \leq 0$ за всяко $z \in K$. Ако $\Lambda \subset \overrightarrow{X \cap H}$ и допуснем, че $\Lambda \not\subset \overrightarrow{K}$, то за някоя точка $\bar{z} \in \Lambda$ съществуват точки $z_1 \in K$, $z_2 \in K$, нележащи едновременно на Λ, такива че $\bar{z} = \alpha z_1 + \beta z_2$, $\alpha > 0$, $\beta > 0$. Но $0 = \langle v, \bar{z}\rangle = \alpha\langle v, z_1\rangle + \beta\langle v, z_2\rangle \leq 0$, следователно $\langle v, z_1\rangle = \langle v, z_2\rangle = 0$, т. е. $z_1 \in K \cap H$, $z_2 \in K \cap H$, което означава, че $\Lambda \not\subset \overrightarrow{K \cap H}$. Полученото противоречие показва, че $\Lambda \subset \overrightarrow{K}$. □

Теорема 1. Затворен изпъкнал конус $K \neq \{0\}$ има краен лъч тогава и само тогава, когато е заострен.

Доказателство. *Необходимост*. Нека $z \neq 0$ е точка от краен лъч на K. Допускаме, че K не е заострен. Тогава той съдържа права линия, например $\{y : y = z + bt,\ t \in \mathbb{R}^1\}$.

1) Ако $\pm b \notin \{\lambda z : z \geq 0\}$, то $z' = z + b \notin \{\lambda z : \lambda \geq 0\}$ и $z'' = z - b \notin \{\lambda z : \lambda \geq 0\}$. Но $z = z' + z''$, което значи, че z не е точка от краен лъч — противоречие.

2) Ако $b \in \{\lambda z : \lambda \geq 0\}$, т. е. $z = t_0 b$, $t_0 > 0$, то $z'' = z - bt_1 \notin \{\lambda z : \lambda \geq 0\}$ при $t_1 > t_0$. Тогава $z = z' + z''$, където $z' = z + t_1 b$, и също стигаме до противоречие. Аналогичен е случаят, когато $-b \in \{\lambda z : \lambda \geq 0\}$. Следователно K е заострен конус.

Достатъчност. Нека K е заострен конус. Ще направим индукция по размерността на конуса. За конус с размерност единица твърдението е очевидно вярно, тъй като заострен едномерен конус е лъч. Нека твърдението е вярно за конус с размерност $p < k$. Ще покажем, че е вярно и за k-мерен заострен конус K. Ако $z \in K$, $z \neq 0$, е истинска контурна точка, то съществува опорна хиперравнина H през z такава, че $K \not\subset H$. Тогава $K \cap H$ е заострен конус, $\dim(K \cap H) < k$ и за него твърдението е вярно, т. е. $\overrightarrow{K \cap H} \neq \emptyset$. Но от $\overrightarrow{K \cap H} \subset \overrightarrow{K}$ следва $\overrightarrow{K} \neq \emptyset$. □

Твърдение 3. Всяка относително вътрешна точка на заострен конус K, $\dim K > 1$, се представя като изпъкнала комбинация на две истински контурни точки.

Доказателство. Нека $z \in \operatorname{ri} K$. Понеже K е заострен, съществува хиперравнина H през върха му, която няма други общи точки с K. През точката z прекарваме права с направляващ вектор $b \in (H \cap \operatorname{aff} K)$, $b \neq 0$. Тогава $\pm b \notin K$. Ще покажем, че

$$t_1 = \sup\{t : z + bt \in K\} < \infty.$$

Действително, ако $z + bt \in K$ за всяко $t > 0$, $t_0 \frac{1}{t} z + b \in K$ за всяко $t > 0$, следователно $b \in K$. Получихме противоречие. Аналогично

$$t_2 = \sup\{t : z - bt \in K\} < \infty.$$

Но $z_1 = z + t_1 b$ и $z_2 = z - t_2 b$ са истински контурни точки и

$$z = \frac{t_2}{t_1 + t_2}(z + t_1 b) + \frac{t_1}{t_1 + t_2}(z - t_2 b).$$ □

Теорема 2. Всяка точка на заострен конус се представя като изпъкнала комбинация на точки от крайни лъчи, т. е. $K = \operatorname{co} \overrightarrow{K}$.

Доказателство. Ще използваме индукция по размерността на конуса. За едномерен конус твърдението е очевидно вярно. Допускаме, че е вярно и за конуси с размерност, по-малка от k, и нека K е k-мерен конус. Ако $z \in K$ е истинска контурна точка, прекарваме през нея опорна хиперравнина H, $K \not\subset H$. Тогава $K \cap H$ е заострен конус с размерност, по-малка от k, и следователно

$$z = \sum_{i=1}^{s} \alpha_i \bar{z}_i, \quad \alpha_i \geq 0, \quad i = 1, \ldots, s, \quad \sum_{i=1}^{s} \alpha_i = 1, \quad \bar{z}_i \in \overrightarrow{K \cap H}.$$

Но $\overrightarrow{K \cap H} \subset \overrightarrow{K}$. Ако $z \in \operatorname{ri} K$, от твърдение 3 следва, че тя се представя във вида

$$z = \lambda z' + (1-\lambda) z'', \quad 0 \leq \lambda \leq 1, \quad z' \in \mathrm{r}\partial K, \quad z'' \in \mathrm{r}\partial K.$$

Но

$$z' = \sum_{i=1}^{s_1} \alpha_i \bar{y}_i, \quad z'' = \sum_{j=1}^{s_2} \beta_j \bar{z}_j, \quad \bar{y}_i \in \overrightarrow{K}, \quad \bar{z}_j \in \overrightarrow{K},$$

$$\alpha_i \geq 0, \quad \beta_j \geq 0, \quad i = 1, \ldots, s_1, \quad j = 1, \ldots, s_2, \quad \sum_{i=1}^{s_1} \alpha_i = 1, \quad \sum_{j=1}^{s_2} \beta_j = 1,$$

и понеже изпъкнала комбинация на изпъкнали комбинации също е изпъкнала комбинация, то z се представя като изпъкнала комбинация на точки от крайни лъчи, т. е. $K \subset \operatorname{co} \overrightarrow{K}$. Очевидно $\operatorname{co} \overrightarrow{K} \subset K$. Следователно $\operatorname{co} \overrightarrow{K} = K$. С това доказателството е завършено. □

Определение 3. Множеството $\operatorname{ray} X = \{z = \lambda x : x \in X, \ \lambda \geq 0\}$ наричаме *конична обвивка* на множеството $X \subset \mathbb{R}^n$.

Очевидно коничната обвивка е конус и ако K е конус, то $\operatorname{ray} K = K$.

Определение 4. Сечението на всички изпъкнали конуси, съдържащи X, наричаме *изпъкнала конична обвивка* на множеството $X \subset \mathbb{R}^n$ и означаваме $\operatorname{con} X$.

Лесно се вижда, че $\operatorname{con} X = \operatorname{co}(\operatorname{ray} X) = \operatorname{ray}(\operatorname{co} X)$, т. е.

$$\operatorname{con} X = \left\{ z : z = \sum_{i=1}^{n+1} \alpha_i z_i, \ \alpha_i \geq 0, \ z_i \in X, \ i = 1, \ldots, n+1 \right\}.$$

Ако $K = \operatorname{con} X$, то множеството X наричаме пораждащо множество за конуса K. Като използваме въведените понятия, теорема 2 може да се изкаже така: Ако K е изпъкнал затворен конус и $\operatorname{ray} X = \overrightarrow{K}$, то $K = \operatorname{con} X$.

4. Представяне на изпъкнали множества

Определение 1. Точката (вектора) $z \in \mathbb{R}^n$ ще наричаме *рецесивна посока* (*рецесия*) за (на) изпъкналото множество X, ако за всяко $x \in X$ и $\lambda \geq 0$ имаме $x + \lambda z \in X$.

От свойствата на затворените изпъкнали множества (гл. 2, § 1) следва, че ако за някое $x_0 \in X$ е изпълнено $x_0 + \lambda z \in X$ за всяко $\lambda \geq 0$, то $x + \lambda z \in X$ за всяко $x \in X$ и $\lambda \geq 0$. Тогава z е рецесивна посока на X, когато X съдържа всички лъчи с начало точка от X и посока z. Множеството от всички рецесивни посоки (рецесии) на изпъкналите затворени множества X ще означаваме с $K(X)$:

$$K(X) = \{z \in \mathbb{R}^n : x + \lambda z \in X,\ x \in X,\ \lambda \geq 0\}.$$

Теорема 1. Ако X е изпъкнало затворено множество, то:

1) $K(X)$ е изпъкнал затворен конус;

2) X е неограничено тогава и само тогава, когато $K(X) \neq \{0\}$;

3) множеството $\widehat{X}$ от крайните точки на X е непразно тогава и само тогава, когато $K(X)$ е заострен конус.

Доказателство. 1) Ако $z_1, z_2 \in K(X)$ и $\alpha \geq 0, \beta \geq 0$, то за кое да е $x \in X$ и $\lambda \geq 0$ имаме $y = x + \lambda\alpha z_1 \in X$ и $y + \lambda\beta z_2 \in X$. Тогава

$$x + \lambda(\alpha z_1 + \beta z_2) = y + \lambda\beta z_2 \in X, \text{ т.е. } \alpha z_1 + \beta z_2 \in K(X).$$

Следователно $K(X)$ е изпъкнал конус. Ще покажем, че $K(X)$ е затворено множество. Нека $\{z_k\}_{k=1}^{\infty} \to z^*$, $z_k \in K(X)$. Тогава за $x \in X$ и $\lambda \geq 0$ имаме $x + \lambda z_k \in X$, $k = 1, 2, \ldots$ Оттук получаваме $x + \lambda z^* \in X$ (X е затворено) за всяко $\lambda \geq 0$, или $z^* \in K(X)$, т.е. $K(X)$ е затворено множество.

2) Нека X е неограничено множество. Ще покажем, че съществува $z \neq 0$, $z \in K(X)$. Тъй като X е неограничено, можем да изберем редица $\{x_k\}$ от точки на X такава, че $\|x_k\| \to \infty$ при $k \to \infty$. Редицата $\left\{\dfrac{x_k}{\|x_k\|}\right\}$ е ограничена и от нея можем да изберем сходяща подредица $\left\{\dfrac{x_{k_i}}{\|x_{k_i}\|}\right\} \to z$, $\|z\| = 1$. Тогава за произволни фиксирани $x \in X$ и $\lambda \geq 0$ имаме

$$\begin{aligned} x + \lambda z &= \lim_{i\to\infty}\left(x + \lambda\frac{x_{k_i}}{\|x_{k_i}\|}\right) = \lim_{i\to\infty}\left[x + \frac{\lambda}{\|x_{k_i}\|}(x_{k_i} - x) + \frac{\lambda x}{\|x_{k_i}\|}\right] \\ &= \lim_{i\to\infty}\left[x + \frac{\lambda}{\|x_{k_i}\|}(x_{k_i} - x)\right]. \end{aligned}$$

За достатъчно големи k_i е изпълнено $0 < \dfrac{\lambda}{\|x_{k_i}\|} < 1$ и тогава $x + \dfrac{\lambda}{\|x_{k_i}\|}(x_{k_i} - x) \in X$, тъй като X е изпъкнало. Следователно $x + \lambda z \in X$ (X е затворено), т. е. $z \in K(X)$.

Обратното твърдение в 2) е очевидно.

3) Множеството X съдържа правата $\{x + bt : t \in \mathbb{R}^1\}$ тогава и само тогава, когато $K(X)$ съдържа правата $\{bt : t \in \mathbb{R}^1\}$. Оттук непосредствено следва, че $\widehat{X} \neq \emptyset$ тогава и само тогава, когато $\widehat{K}(X) \neq \emptyset$, т. е. когато $K(X)$ е заострен. □

Упражнение 1. Ако X и Y са изпъкнали затворени множества и $X \cap Y \neq \emptyset$, то $K(X \cap Y) = K(X) \cap K(Y)$. Покажете с пример, че условието $X \cap Y \neq \emptyset$ е съществено.

Лема 1. Ако X е изпъкнало затворено множество и $\widehat{X} \neq \emptyset$, то всяка относително вътрешна точка на X се представя като изпъкнала комбинация на две истински контурни точки или във вида $\overline{x}+\overline{z}$, където $\overline{x} \in \mathrm{r}\partial X$ и $\overline{z} \in K(X)$.

Доказателство. През произволна точка $x \in \mathrm{ri}\, X$ прекарваме правата $\{x + zt : t \in \mathbb{R}^1\} \subset \mathrm{aff}\, X$. Да означим

$$t_1 = \sup\{t : x + tz \in X\},$$
$$t_2 = \sup\{t : x - tz \in X\}.$$

Съществуват следните възможности:

1) И t_1, и $t_2 < \infty$. Тогава точките $x_1 = x + t_1 z$ и $x_2 = x - t_2 z$ са истински контурни и
$$x = \frac{t_2}{t_1 + t_2} x_1 + \frac{t_1}{t_1 + t_2} x_2.$$

2) Или t_1, или t_2 е крайно. Нека $t_2 < \infty$. Тогава $\overline{x} = x - zt_2 \in \mathrm{r}\partial X$, а $z \in K(X)$, т. е. $x = \overline{x} + t_2 z$ и $t_2 z \in K(X)$.

Случаят $t_1 = t_2 = \infty$ е невъзможен, тъй като $\widehat{X} \neq \emptyset$, т. е. X не съдържа права линия (теорема 2 от § 1). □

Упражнение 2. Докажете, че ако $x^0 \in X$, то $K(X) = \bigcap\limits_{k=1}^{\infty} \dfrac{1}{k}(X - x^0)$.

Теорема 2. Ако X е изпъкнало затворено множество и $\widehat{X} \neq \emptyset$, то $X = \mathrm{co}\, \widehat{X} + K(X)$.

Доказателство. Ясно е, че $\mathrm{co}\, \widehat{X} \subset X$ и ако $x \in X$ и $z \in K(X)$, то $x + z \in X$, следователно $\mathrm{co}\, \widehat{X} + K(X) \subset X$. Обратното включване ще докажем с индукция по размерността на множеството X. За едномерните множества (отсечки и лъчи) твърдението е очевидно. Нека твърдението е вярно за множества с размерност, по-малка от k, и $\dim X = k$. Като имаме предвид представянето на относително вътрешните точки, достатъчно е да докажем, че всяка истинска контурна точка на X се представя като изпъкнала комбинация на крайни точки плюс рецесивна посока. И така нека $\overline{x} \in \mathrm{r}\partial X$. Прекарваме през $\overline{x}$ опорна хиперравнина H такава, че $X \not\subset H$. Тогава $X_1 = X \cap H$ е изпъкнало затворено множество и $\widehat{X_1} \neq \emptyset$, $\dim X_1 < k$. От индуктивното допускане имаме

$$\overline{x} = \sum_{i=1}^{s} \alpha_i \widehat{x_i} + z, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{s} \alpha_i = 1, \quad \widehat{x_i} \in \widehat{X_1}, \quad z \in K(X_1).$$

Но $\widehat{X_1} \subset \widehat{X}$ и $K(X_1) = K(X \cap H) = K(X) \cap K(H) \subset K(X)$, т. е. $\overline{x} \in \mathrm{co}\,\widehat{X} + K(X)$ или $X \subset \mathrm{co}\,\widehat{X} + K(X)$. □

Забележка. Ако X е изпъкнало компактно множество, то $K(X) = \{0\}$ и $X = \mathrm{co}\,\widehat{X}$, т. е. получаваме теоремата на Минковски.

Ще припомним, че две подпространства S и T на $\mathbb{R}^n$ наричаме *допълнителни*, ако $S + T = \mathbb{R}^n$ и $S \cap T = \{0\}$. Всяко от тези подпространства е допълнение за другото.

Две подпространства S и T на $\mathbb{R}^n$ са допълнителни тогава и само тогава, когато $S \cap T = \{0\}$ и $\dim S + \dim T = n$. Всяка точка $x \in \mathbb{R}^n$ може да се представи като $x = x_1 + x_2$, $x_1 \in S$, $x_2 \in T$, и това представяне е единствено. Ако S и T са допълнителни подпространства и $x = x_1 + x_2$, $x_1 \in S$, $x_2 \in T$, точките x_1 и x_2 наричаме *проекции* на точката x върху подпространствата S и T и ще означаваме съответно

$$x_1 = \mathrm{пр}_S\, x, \quad x_2 = \mathrm{пр}_T\, x.$$

С $\mathrm{пр}_S X$ и $\mathrm{пр}_T X$ означаваме множествата от проекции на точките от X съответно върху S и T.

Определение 2. Нека X е изпъкнало затворено множество. Множеството $S(X) = K(X) \cap (-K(X))$ ще наричаме *рецесивно подпространство* на множеството X.

С други думи, $z \in S(X)$, ако $\{x + tz : t \in \mathbb{R}^1\} \subset X$ за всяко $x \in X$.

Следващите свойства се получават непосредствено от определението за рецесивно подпространство:

1) $S(X)$ е подпространство;

2) $s \in S(X)$ тогава и само тогава, когато $s \in K(X)$ и $-s \in K(X)$;

3) $\widehat{X} \neq \emptyset$ тогава и само тогава, когато $S(X) = \{0\}$;

4) $S(K(X)) = S(X)$.

Теорема 3. Ако X е изпъкнало затворено множество и T е допълнително подпространство на $S(X)$, то:

1) $K(X) \cap T$ е заострен конус;

2) множеството $X \cap T$ има крайна точка;

3) а) $\text{пр}_T X = X \cap T$ и б) $\text{пр}_S X = S(X)$.

Доказателство. 1) Ако допуснем, че $K(X) \cap T$ не е заострен конус, то съществува $z \neq 0$ такова, че $z \in K(X) \cap T$ и $-z \in K(X) \cap T$. Тогава

$$z \in (K(X) \cap T) \cap (-K(X) \cap T) = K(X) \cap (-K(X)) \cap T = S(X) \cap T.$$

Но $S(X) \cap T = \{0\}$, т. е. стигаме до противоречие.

2) Следва от това, че $K(X \cap T) = K(X) \cap T$ (упражнение 1) е заострен конус. Съгласно теорема 1 $X \cap T$ има крайна точка.

3) а) Очевидно $X \cap T \subset \text{пр}_T X$. Ще докажем и обратното включване. Нека $x_1 \in \text{пр}_T X$, т. е. $x_1 \in T$, и съществуват точки $x \in X$, $x_2 \in S(X)$ такива, че $x = x_1 + x_2$. От дефиницията на $S(X)$ получаваме $x_1 = x - x_2 \in X$. Следователно $x_1 \in X \cap T$, т. е. $\text{пр}_T X \subset X \cap T$, от което следва $\text{пр}_T X = X \cap T$.

б) Очевидно $\text{пр}_S X \subset S(X)$. Ще покажем обратното включване. Нека $x_2 \in S(X)$ и $x_1 \in X \cap T$. Тогава $x = x_1 + x_2 \in X$, т. е. $x_2 \in \text{пр}_S X$, и оттук $S(X) \subset \text{пр}_S X$ или $\text{пр}_S X = S(X)$. □

Теорема 4. Ако X е изпъкнало затворено множество и T е допълнение на $S(X)$, то $X = X \cap T + S(X)$.

Доказателство. От $X \cap T + S(X) \subset X$ и $X \subset \text{пр}_T X + \text{пр}_S X = X \cap T + S(X)$ следва $X = X \cap T + S(X)$. □

Теорема 5. Ако X е изпъкнало затворено множество, T е допълнение на $S(X)$ и $X_1 = X \cap T$, то

$$X = \operatorname{co} \widehat{X_1} + K(X_1) + S(X). \tag{1}$$

Доказателство. Тъй като $X = X_1 + S(X)$ и $\widehat{X_1} \neq \emptyset$, то от теорема 2 следва $X = \operatorname{co} \widehat{X_1} + K(X_1) + S(X)$. □

Забележка. Понеже $K(X_1) + S(X) = K(X \cap T) + S(X) = K(X) \cap T + S(K(X)) = K(X)$, то (1) може да се запише така:

$$X = \operatorname{co} \widehat{X_1} + K(X).$$

Задача 1. Докажете, че ако X е изпъкнало затворено множество с крайна точка и $\dim X > 1$, то всяка относителна вътрешна точка се представя като изпъкнала комбинация на две истински контурни точки.

5. Многостенни множества

В този параграф ще разгледаме някои специфични свойства на изпъкналите многостенни множества. Те са важен подклас на изпъкналите множества, тъй като много практически задачи се свеждат до оптимизация в многостенни области.

Определение 1. Множество от вида

$$X = \{x \in \mathbb{R}^n : Ax \le b\}, \tag{1}$$

където $A = (a_{ij})$ е матрица с реални коефициенти и размери $m \times n$, $b = (b_1, b_2, \dots, b_m) \in \mathbb{R}^m$, ще наричаме *многостенно множество*.

В определението на многостенно множество могат да участват и равенства, но това не променя същността на нещата, понеже всяко равенство е еквивалентно на две неравенства. Неравенствата и равенствата в определението на многостенно множество наричаме *условия* или *ограничения*. Едно условие е активно за точката $x \in X$, ако се удовлетворява от тази точка като равенство. Ясно е, че многостенните множества са изпъкнали и затворени множества.

Ако едно многостенно множество е конус, то се нарича *многостенен конус*. Лесно се вижда, че многостенен конус се представя във вида

$$K = \{z \in \mathbb{R}^n : Az \le 0\}.$$

Ако $X = \{x \in \mathbb{R} : Ax \le b\}$, то $K(X) = \{z : Az \le 0\}$ и $S(X) = \{s : As = 0\}$ (вж. § 4).

Теорема 1. Ако $X \ne \emptyset$ е изпъкнало многостенно множество в $\mathbb{R}^n$, то $\widehat{X} \ne \emptyset$ тогава и само тогава, когато $\operatorname{rank} A = n$.

Доказателство. *Необходимост.* Нека $\widehat{X} \ne \emptyset$. Тогава $K(X)$ е заострен конус и $S(X) = \{0\}$. Следователно $\operatorname{rank} A = n$.

Достатъчност. Нека $\operatorname{rank} A = n$. Тогава $S(X) = \{0\}$, $K(X)$ е заострен конус, следователно $\widehat{X} \ne \emptyset$. □

Теорема 2. Ако $X \ne \emptyset$ е изпъкнало многостенно множество в $\mathbb{R}^n$ и $\overline{x} \subset X$, то $\overline{x} \in \widehat{X}$ тогава и само тогава, когато $\operatorname{rank} A' = n$, където A' е матрицата от коефициентите на активните за точката $\overline{x}$ ограничения.

Доказателство. *Необходимост.* Нека $\overline{x} \in \widehat{X}$. Допускаме, че $\operatorname{rank} A' < n$. Тогава ще съществува $h \ne 0$ такова, че $A'h = 0$. Но при достатъчно малко $\delta > 0$ точките $\overline{x} \pm \delta h \in X$ и $\overline{x} = \frac{1}{2}[(\overline{x} + \delta h) + (\overline{x} - \delta h)]$, което противоречи на допускането, че $\overline{x} \in \widehat{X}$.

Достатъчност. Нека $\operatorname{rank} A' = n$. Допускаме, че $\overline{x} \notin \widehat{X}$. Тогава

$$\overline{x} = (1-\lambda)x_1 + \lambda x_2, \quad 0 < \lambda < 1, \quad x_1, x_2 \in X, \quad x_1 \neq x_2.$$

Но

$$b' = A'\overline{x} = A'[(1-\lambda)x_1 + \lambda x_2] = (1-\lambda)A'x_1 + \lambda A'x_2 \leq b'.$$

Следователно $A'x_1 = A'x_2 = b'$. Стигаме до противоречие с допускането, че $\operatorname{rank} A' = n$. □

Теорема 3. Ако $K = \{z : Az \leq 0\}$ и $\operatorname{rank} A = n$, то $\overline{z} \in \overrightarrow{K}$ тогава и само тогава, когато $\operatorname{rank} A' = n - 1$ (A' е матрицата на активните в точката $\overline{z}$ ограничения).

Доказателството не се различава съществено от това на теорема 2.

Следствие 1. Крайните точки на изпъкналото многостенно множество (ако има такива) са краен брой.

Следствие 2. Крайните лъчи на заострен многостенен конус са краен брой.

Както вече знаем, ако K е изпъкнал заострен конус и $\overrightarrow{K'} = \{z \in \overrightarrow{K} : \|z\| = 1\}$, т. е. $\overrightarrow{K'}$ е множество, което съдържа точно един елемент (различен от нула) от всеки краен лъч на K и не съдържа други елементи, то $K = \operatorname{con} \overrightarrow{K'}$.

За произволно изпъкнало многостенно множество X от теоремите на представянето получаваме $\widehat{X} = \operatorname{co} \widehat{X}_1 + K(X_1) + S(X)$. Оттук

$$X = \operatorname{co} \widehat{X}_1 + \operatorname{con} \overrightarrow{K'}(X) + S(X),$$

т. е. изпъкналите многостенни множества се представят като сума на изпъкналата обвивка на краен брой точки (крайните точки), коничната обвивка на краен брой точки и рецесивното подпространство.

Определение 2. Казваме, че конусът Y е *крайно породен*, ако съществуват вектори $y_1, y_2, \ldots, y_k$ от Y такива, че

$$Y = \left\{ y : y = \sum_{i=1}^{k} \alpha_i y_i,\ \alpha_i \geq 0,\ i = 1, \ldots, k \right\}.$$

Векторите y_i, $i = 1, 2, \ldots, k$, се наричат *образуващи* за Y. От казаното порано е ясно, че всеки многостенен конус $K = \{x : Ax \leq 0\}$ е крайно породен (вж. §§ 4, 5). Нашата цел е да покажем, че и обратното твърдение е вярно, т. е. всеки крайно породен конус е многостенен. В сила е следното просто

Твърдение 1. Нека Y е крайно породен конус и $y_1, y_2, \ldots, y_k$ са неговите образуващи. Тогава спрегнатият конус $Y^* = \{x : \langle x, y \rangle \leq 0$ за всяко $y \in Y\}$ е многостенен и се представя във вида $\{x : \langle y_i, x \rangle \leq 0, i = 1, 2, \ldots, k\}$.

Нека $A = (a_{ij})$ е матрица $m \times n$ и $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, m$. Очевидно конусът $K = \{x : Ax \leq 0\}$ може да се представи и във вида

$$K = \{x : \langle a_i, x \rangle \leq 0,\ i = 1, 2, \ldots, m\}.$$

За спрегнатия конус $K^* = \{y : \langle y, z \rangle \leq 0$ за всяко $z \in K\}$ е в сила следното

Твърдение 2. Ако $K = \{x : \langle a_i, x \rangle \leq 0,\ i = 1, 2, \ldots, m\}$, то

$$K^* = \left\{ y : y = \sum_{i=1}^{m} \alpha_i a_i,\ \alpha_i \geq 0,\ i = 1, 2, \ldots, m \right\}.$$

Доказателство. Нека $y = \sum_{i=1}^{m} \alpha_i a_i$, където $\alpha_i \geq 0$, $i = 1, 2, \ldots, m$. Тогава за $x \in K$ е в сила неравенството

$$\langle y, x \rangle = \sum_{i=1}^{m} \alpha_i \langle a_i, x \rangle \leq 0,$$

т. е. $y \in K^*$. Обратно, нека $y \in K^*$, т. е. $\langle y, x \rangle \leq 0$, когато $\langle a_i, x \rangle \leq 0$, $i = 1, 2, \ldots, m$. От теоремата на Фаркаш следва, че съществуват числа $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, за които $y = \sum_{i=1}^{m} \lambda_i a_i$. Твърдението е доказано. $\square$

Теорема 4. Един конус е крайно породен тогава и само тогава, когато е многостенен конус.

Доказателство. Вече се убедихме, че всеки многостенен конус е крайно породен. Нека Y е крайно породен конус, т. е.

$$Y = \left\{ y : y = \sum_{i=1}^{k} \alpha_i y_i,\ \alpha_i \geq 0,\ i = 1, 2, \ldots, k \right\}.$$

Доказахме, че Y^* е многостенен конус. Следователно Y^* е крайно породен. Но тогава Y^{**} е многостенен конус. Остава да докажем, че $Y = Y^{**}$. Според теорема 1 от § 2 достатъчно е да се убедим, че Y е затворено множество. Пак ще използваме теоремата на Фаркаш. Нека $z_0 = \lim_{j \to \infty} z_j$, $z_j \in Y$. Нека x е вектор, за който $\langle y_i, x \rangle \leq 0$, $i = 1, 2, \ldots, k$. Тогава $\langle z_j, x \rangle \leq 0$ за всяко $j = 1, 2, \ldots$. След граничен преход получаваме $\langle z_0, x \rangle \leq 0$. Съгласно теоремата на Фаркаш съществуват числа $\lambda_i \geq 0$, $i = 1, 2, \ldots, k$, за които $z_0 = \sum_{i=1}^{k} \lambda_i y_i$, т. е. $z_0 \in Y$. Теоремата е доказана. $\square$

Определение 3. Всяко ограничено многостенно множество се нарича *многостен*.

От теоремата на Минковски и от резултатите в § 4 следва, че всеки многостен е изпъкнала обвивка на краен брой точки. Сега ще докажем и обратното твърдение, че всяка изпъкнала обвивка на краен брой точки е многостен.

Теорема 5. Едно множество е многостен тогава и само тогава, когато е изпъкнала обвивка на краен брой точки.

Доказателство. Достатъчно е да докажем, че всяко множество от вида $X = \mathrm{co}(x_1, \ldots, x_p)$ е многостен, т. е. че съществуват вектори a_i и числа b_i, $i = 1, 2, \ldots, m$, такива, че

$$X = \mathrm{co}(x_1, x_2, \ldots, x_p) = \{z : \langle a_i, z\rangle \le b_i,\ i = 1, 2, \ldots, m\}.$$

Разглеждаме множеството

$$X' = \{(x, 1) \in \mathbb{R}^{n+1} : x \in X\} = \mathrm{co}\{(x_1, 1), \ldots, (x_p, 1)\}$$

и конуса $Y = \bigcup\limits_{t\ge 0} tX'$. Множеството Y е крайно породен конус с образуващи $y_j = (x_j, 1)$, $j = 1, 2, \ldots, p$, т. е.

$$Y = \left\{ y \in \mathbb{R}^{n+1} : y = \sum_{j=1}^{p} \alpha_j y_j,\ \alpha_j \ge 0 \right\}.$$

Съгласно теорема 4 Y е многостенен конус, т. е. съществуват вектори $z_i = (a_i, \beta_i) \in \mathbb{R}^n \times \mathbb{R}^1 = \mathbb{R}^{n+1}$, $i = 1, 2, \ldots, m$, такива, че

$$Y = \{y \in \mathbb{R}^{n+1} : \langle z_i, y\rangle \le 0,\ i = 1, 2, \ldots, m\}.$$

В частност за всяка точка $(x, 1) \in Y$, където $x \in X$, са в сила неравенствата

$$0 \ge \langle z_i, (x, 1)\rangle = \langle a_i, x\rangle + \beta_i, \quad i = 1, 2, \ldots, m,$$

т. е. $\langle a_i, x\rangle \le -\beta_i$. Обратно, ако за някое $x \in \mathbb{R}^n$ са в сила неравенствата $\langle a_i, x\rangle \le -\beta_i$, $i = 1, 2, \ldots, m$, то за $y = (x, 1)$ са в сила неравенствата $\langle z_i, y\rangle \le 0$, $i = 1, 2, \ldots, m$, и следователно y е точка от Y. Тъй като $Y \cap \{z = (x, 1) : x \in \mathbb{R}^n\} = X'$, то $x \in X$. Следователно

$$X = \{x \in \mathbb{R}^n : \langle a_i, x\rangle \le -\beta_i,\ i = 1, 2, \ldots, m\}.$$

Полагаме $b_i = -\beta_i$, $i = 1, 2, \ldots, m$. Твърдението е доказано. □

Глава 4

Изпъкнали функции

Изпъкналите функции имат интересни аналитични свойства. Оказва се например, че изпъкналите функции са непрекъснати във вътрешността на дефиниционната си област. С тях е свързано понятието субградиент, обобщаващо класическото понятие производна на функция. Свойствата на изпъкналите функции са основа за получаване на необходими и достатъчни условия за екстремум.

1. Основни свойства на изпъкналите функции

В тази глава ще разглеждаме функции, дефинирани в подмножества на $\mathbb{R}^n$ със стойности в $\mathbb{R}^1$.

Определение 1. Нека X е изпъкнало подмножество на $\mathbb{R}^n$. Функцията $f : X \to \mathbb{R}^1$ се нарича *изпъкнала*, ако

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2), \tag{1}$$

когато $x_1, x_2 \in X$ и числото t е от интервала $[0, 1]$.

Геометричното тълкуване на неравенството (1), наречено още неравенство на Йенсен, е показано фиг. 24: стойностите на функцията f в точката $x_t = tx_1 + (1-t)x_2$ не надминават стойностите на линейната функция (на t)

$$tf(x_1) + (1-t)f(x_2) = f(x_2) + t(f(x_1) - f(x_2)),$$

т. е. графиката на $f(x_t)$ е под графиката на линейната функцията $f(x_2) + t(f(x_1) - f(x_2))$.

Пример 1. Функцията $f(x) = \langle c, x \rangle$, където $c \in \mathbb{R}^n$ е фиксиран вектор, е изпъкнала. За нея неравенството на Йенсен се изпълнява като равенство.

Пример 2. Функцията $f(x) = \|x\|$ е изпъкнала, тъй като

$$f(tx_1 + (1-t)x_2) = \|tx_1 + (1-t)x_2\| \leq t\|x_1\| + (1-t)\|x_2\|$$

за всяко $0 \leq t \leq 1$.

Упражнение 1. Нека $f_1, \dots, f_k : X \to \mathbb{R}^1$ са изпъкнали функции и $\lambda_i > 0$, $i = 1, 2, \dots, k$, са реални числа. Докажете, че функцията $f(x) = \sum_{i=1}^{k} \lambda_i f_i(x)$ е изпъкнала.

Упражнение 2. Нека $f_1, \dots, f_k : X \to \mathbb{R}^1$ са изпъкнали функции. Докажете, че функцията $f(x) = \max_{i \in \{1,\dots,k\}} f_i(x)$ е изпъкнала.

Упражнение 3. Нека X е непразно множество в $\mathbb{R}^n$ и функцията $f : \mathbb{R}^n \to \mathbb{R}^1$ е дефинирана по следния начин:

$$f(x) = d(x, X) = \inf\{\|y - x\|, y \in X\}.$$

Докажете, че функцията f е изпъкнала.

Нека X е непразно множество в $\mathbb{R}^n$ и $f : X \to \mathbb{R}^1$. Ще припомним, че надграфика на функцията f наричаме множеството

$$\operatorname{epi} f = \{(x, r) \in \mathbb{R}^{n+1} : x \in X,\ r \geq f(x)\}.$$

Твърдение 1. Нека X е непразно изпъкнало множество. Функцията $f : X \to \mathbb{R}^1$ е изпъкнала тогава и само тогава, когато множеството $\operatorname{epi} f$ е изпъкнало.

Доказателство. Нека f е изпъкнала функция и (x_1, r_1), $(x_2, r_2) \in \operatorname{epi} f$. Ако $t \in (0, 1)$, то от $r_1 \geq f(x_1)$, $r_2 \geq f(x_2)$ следва $tr_1 + (1-t)r_2 \geq tf(x_1) + (1-t)f(x_2) \geq f(tx_1 + (1-t)x_2)$. Това означава, че $\operatorname{epi} f$ е изпъкнало множество.

Обратно, нека $\operatorname{epi} f$ е изпъкнало множество и $x_1, x_2 \in X$. Точките $(x_1, f(x_1))$ и $(x_2, f(x_2))$ са от $\operatorname{epi} f$, следователно за всяко $t \in (0, 1)$ е изпълнено

$$(tx_1 + (1-t)x_2, tf(x_1) + (1-t)f(x_2)) \in \operatorname{epi} f.$$

От това следва, че

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2),$$

т. е. f е изпъкнала функция. □

Упражнение 4. Докажете, че ако f е изпъкнала функция, то множеството на Лебег $L(f, \alpha) = \{x \in X : f(x) \leq \alpha\}$ е изпъкнало за всяко $\alpha \in \mathbb{R}^1$. Посочете пример за неизпъкнала функция, за която съответните множества на Лебег са изпъкнали.

Нека X е изпъкнало множество и $f : X \to \mathbb{R}^1$. По дадени $x \in X$ и $s \in \mathbb{R}^n$ дефинираме функцията на една променлива $\varphi_{x,s}(t) = f(x+ts)$ в множеството $\{t \in \mathbb{R}^1 : x + ts \in X\}$.

Твърдение 2. Функцията f е изпъкнала тогава и само тогава, когато за всяко $x \in X$ и $s \in \mathbb{R}^n$ функцията $\varphi_{x,s}(t)$ е изпъкнала (по t).

Доказателство. Ако f е изпъкнала функция, то при всяко $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ е изпълнено

$$\begin{aligned}\varphi_{x,s}(\alpha_1 t_1 + \alpha_2 t_2) &= f(x + (\alpha_1 t_1 + \alpha_2 t_2)s) = f(\alpha_1(x + t_1 s) + \alpha_2(x + t_2 s)) \\ &\leq \alpha_1 f(x + t_1 s) + \alpha_2 f(x + t_2 s) = \alpha_1 \varphi_{x,s}(t_1) + \alpha_2 \varphi_{x,s}(t_2).\end{aligned}$$

Обратно, нека $\varphi_{x,s}(t)$ е изпъкнала по t за произволни $x \in X$ и $s \in \mathbb{R}^n$. Тогава

$$\begin{aligned}f(\alpha_1 x_1 + \alpha_2 x_2) &= f(x_2 + (\alpha_1 \cdot 1 + \alpha_2 \cdot 0)(x_1 - x_2)) \\ &= \varphi_{x_2, x_1 - x_2}(\alpha_1 \cdot 1 + \alpha_2 \cdot 0) \\ &\leq \alpha_1 \varphi_{x_2, x_1 - x_2}(1) + \alpha_2 \varphi_{x_2, x_1 - x_2}(0) \\ &= \alpha_1 f(x_1) + \alpha_2 f(x_2).\end{aligned}$$ □

Твърдение 3. Нека $f : X \to \mathbb{R}^1$ е изпъкнала функция, $x_1, x_2, \ldots, x_k$ са точки от X, а $\lambda_1, \lambda_2, \ldots, \lambda_k$ — неотрицателни числа, за които $\sum_{i=1}^{k} \lambda_i = 1$. Тогава

$$f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i). \tag{2}$$

Доказателство. Ще направим индукция по дължината k на изпъкналата комбинация $x = \sum_{i=1}^{k} \lambda_i x_i$. При $k = 2$ неравенството (2) съвпада с (1) и е следствие от изпъкналостта на f. Нека допуснем, че неравенството (2) е в сила за $k = p - 1$ и нека $x = \sum_{i=1}^{p} \lambda_i x_i$, където $x_i \in X$, $\lambda_i \geq 0$, $i = 1, 2, \ldots, p$, и $\sum_{i=1}^{p} \lambda_i = 1$. За определеност ще смятаме, че $\lambda_i > 0$, $i = 1, 2, \ldots, p$ (в противен случай, т. е. ако $\lambda_i = 0$ за някое i, то доказваното неравенство ще е изпълнено съгласно индуктивното предположение). Тогава

$$f\left(\lambda_1 x_1 + \sum_{i=2}^{p} \lambda_i x_i\right) = f\left(\lambda_1 x_1 + \lambda \sum_{i=2}^{p} \frac{\lambda_i}{\lambda} x_i\right),$$

където $\lambda = \sum_{i=2}^{p} \lambda_i$. Точката $y = \sum_{i=2}^{p} \frac{\lambda_i}{\lambda} x_i$ е от изпъкналото множество X, защото $\sum_{i=2}^{p} \frac{\lambda_i}{\lambda} = 1$. Следователно поради изпъкналостта на f имаме

$$f(x) = f(\lambda_1 x_1 + \lambda y) \leq \lambda_1 f(x_1) + \lambda f(y).$$

От друга страна, според индуктивното предположение

$$f(y) = f\left(\sum_{i=2}^{p} \frac{\lambda_i}{\lambda} x_i\right) \leq \sum_{i=2}^{p} \frac{\lambda_i}{\lambda} f(x_i).$$

От двете неравенства получаваме

$$f\left(\sum_{i=1}^{p} \lambda_i x_i\right) \leq \lambda_1 f(x_1) + \lambda f(y) \leq \lambda_1 f(x_1) + \sum_{i=2}^{p} \lambda_i f(x_i) = \sum_{i=1}^{p} \lambda_i f(x_i).$$

Твърдението е доказано. □

Следствие 1. Нека $f : X \to \mathbb{R}^1$ е изпъкнала функция, $x_1, \ldots, x_k \in X$ и нека $P = \mathrm{co}(x_1, \ldots, x_k)$. Тогава за всяко $x \in P$

$$f(x) \leq \max\{f(x_i) : i = 1, 2, \ldots, k\}.$$

Доказателство. За всяко $x = \sum_{i=1}^{k} \lambda_i x_i \in P$ е в сила неравенството

$$f(x) \leq \sum_{i=1}^{k} \lambda_i f(x_i) \leq \max\{f(x_i) : i = 1, 2, \ldots, k\}.$$

В частност (вж. гл. 3, § 5) f е ограничена отгоре функция във всяко ограничено многостенно множество (многостен), съдържащо се в X. □

Твърдение 4. Нека $f : X \to \mathbb{R}^1$ е изпъкнала функция и $x_0 \in \mathrm{int}\, X$. Тогава съществува кълбо $B(x_0, \delta)$, $\delta > 0$, в което функцията f е ограничена отгоре.

Доказателство. Нека N е n-мерен куб с център в точката x_0, който изцяло се съдържа в X. Според следствие 1 f е ограничена в N. Но $x_0 \in \mathrm{int}\, N$, следователно съществува $\delta > 0$ такова, че $B(x_0, \delta) \subset N$. Твърдението е доказано. □

Теорема 1. Нека $f : X \to \mathbb{R}^1$ е изпъкнала функция. Тогава f е непрекъсната във всяка вътрешна точка на X.

Доказателство. Нека $x_0 \in \operatorname{int} X$ и нека $\delta > 0$ е такова число, че $f(x) \le r$ за всяко $x \in B(x_0, \delta)$. За произволна точка $y \in B(x_0, \delta)$, $y \ne x_0$, полагаме

$$w = x_0 + \frac{\delta(y - x_0)}{\|y - x_0\|}, \quad z = x_0 - \frac{\delta(y - x_0)}{\|y - x_0\|}.$$

Точката x_0 лежи на отсечката с краища y и z (фиг. 25) и може да се представи като изпъкнала комбинация на тези точки: $x_0 = \dfrac{1}{1+\lambda} z + \dfrac{\lambda}{1+\lambda} y$, където $\lambda = \dfrac{\delta}{\|y - x_0\|}$. Поради изпъкналостта на функцията f

$$f(x_0) \le \frac{1}{1+\lambda} f(z) + \frac{\lambda}{1+\lambda} f(y).$$

Следователно $\lambda(f(x_0) - f(y)) \le f(z) - f(x_0) \le r - f(x_0)$, т. е.

$$f(x_0) - f(y) \le \frac{1}{\lambda}(r - f(x_0)) = \frac{\|y - x_0\|}{\delta}(r - f(x_0)). \tag{3}$$

От друга страна, като вземем предвид, че y лежи на отсечката с краища x_0, w, получаваме $\lambda y = w - (1 - \lambda)x_0 = w + (\lambda - 1)x_0$, т. е. $y = \dfrac{1}{\lambda} w + \dfrac{\lambda - 1}{\lambda} x_0$. Тъй като $\lambda = \dfrac{\delta}{\|y - x_0\|} > 1$, то от изпъкналостта на f получаваме $f(y) \le \dfrac{1}{\lambda} f(w) + \dfrac{\lambda - 1}{\lambda} f(x_0)$. Следователно

$$\lambda(f(y) - f(x_0)) \le f(w) - f(x_0) \le r - f(x_0).$$

Оттук получаваме

$$f(y) - f(x_0) \le \frac{1}{\lambda}(r - f(x_0)) = \frac{\|y - x_0\|}{\delta}(r - f(x_0)). \tag{4}$$

От (3) и (4) се вижда, че

$$|f(x_0) - f(y)| \le \frac{1}{\delta}\|y - x_0\|(r - f(x_0)).$$

Теоремата е доказана. □

Следствие 2. Всяка изпъкнала функция $f : \mathbb{R}^n \to \mathbb{R}^1$ е непрекъсната.

Сега ще покажем, че изпъкналите функции имат и свойства, подобни на диференцуемостта.

Определение 2. Нека X е непразно множество в $\mathbb{R}^n$ и $f : X \to \mathbb{R}^1$. Нека $x_0 \in X$ и $s \in \mathbb{R}^n$, $s \neq 0$, е вектор, за който $x_0 + ts \in X$ за достатъчно малки числа $t > 0$. Казваме, че f е *диференцуема по посока s* в точката x_0, ако съществува границата

$$\lim_{\substack{t \to 0 \\ t > 0}} \frac{f(x_0 + ts) - f(x_0)}{t}.$$

Тази граница ще означаваме с $\dfrac{\partial f}{\partial s}(x_0)$ и ще наричаме *производна на f в точката x_0 по посока s.*

Теорема 2. Изпъкналата функция $f : X \to \mathbb{R}^1$ има във всяка точка $x_0 \in \operatorname{int} X$ производна по всяка посока $s \in \mathbb{R}^n$.

Доказателство. Нека $\varphi(t) = f(x_0 + ts)$ и $t \in \mathbb{R}^1$ е такова, че $x_0 + ts \in X$. Понеже x_0 е вътрешна за X, функцията $\varphi(t)$ е дефинирана в околност на нулата и е изпъкнала. Трябва да докажем, че $\displaystyle\lim_{t \to 0+} \frac{\varphi(t) - \varphi(0)}{t}$ съществува. Това ще направим, като най-напред се убедим, че частното $\dfrac{\varphi(t) - \varphi(0)}{t}$ намалява, когато t клони към 0, и след това, че е ограничено отдолу. Нека $0 < t_1 < t_2$. Тогава

$$t_1 = \frac{t_1}{t_2} t_2 + \left(1 - \frac{t_1}{t_2}\right) \cdot 0.$$

Поради изпъкналостта на $\varphi(t)$ имаме

$$\varphi(t_1) = \frac{t_1}{t_2} \varphi(t_2) + \left(1 - \frac{t_1}{t_2}\right) \varphi(0),$$

т. е.

$$\frac{\varphi(t_1) - \varphi(0)}{t_1} \leq \frac{\varphi(t_2) - \varphi(0)}{t_2}.$$

Нека сега $t_3 < 0 < t$. Тогава

$$0 = -\frac{t_3}{t - t_3} t + \frac{t}{t - t_3} t_3.$$

Пак поради изпъкналостта на $\varphi(t)$ получаваме

$$\varphi(0) \leq \frac{t}{t - t_3} \varphi(t_3) - \frac{t_3}{t - t_3} \varphi(t),$$

а след опростяване —

$$t_3(\varphi(t) - \varphi(0)) \leq t(\varphi(t_3) - \varphi(0)).$$

Тъй като $t_3 < 0$, следва, че

$$\frac{\varphi(t) - \varphi(0)}{t} \geq \frac{\varphi(t_3) - \varphi(0)}{t_3}, \quad t > 0,$$

т. е. частното $\dfrac{\varphi(t) - \varphi(0)}{t}$ е ограничено отдолу. Теоремата е доказана. $\square$

Задача 1. Нека $f : \mathbb{R}^1 \to \mathbb{R}^1$ е функция, която има лява и дясна производна във всяка точка (изпъкналите функции например са такива). Да се докаже, че f е диференцуема във всички точки с евентуално изключение на изброимо много (т. е. съществува редица $t_1, t_2, \ldots, t_n, \ldots$ такава, че от $t \neq t_i$, $i = 1, 2, \ldots$, следва, че f е диференцуема в t).

Задача 2. Нека $X \subset \mathbb{R}^n$ е изпъкнало множество и $f : X \to \mathbb{R}^1$ е непрекъсната функция, удовлетворяваща условието

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}$$

при всеки избор на x_1, x_2 от X. Докажете, че f е изпъкнала функция.

Задача 3. Запазва ли се твърдението, съдържащо се в задача 2, ако заменим изискването за непрекъснатост с по-слабото изискване: за всяка точка $x_0 \in X$ съществува околност, в която f е ограничена функция?

Задача 4. Докажете, че не съществува функция $f : \mathbb{R}^1 \to \mathbb{R}^1$, която е изпъкнала, ограничена и различна от константа в $\mathbb{R}^1$.

2. Диференцуеми изпъкнали функции

Нека X е непразно множество в $\mathbb{R}^n$ и $f : X \to \mathbb{R}^1$. Ще припомним, че функцията f е диференцуема в точката $x_0 \in \operatorname{int} X$, ако съществува вектор $f'(x_0)$, наричан *производна* или *градиент* на функцията f, такъв, че

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0) - \langle f'(x_0), x - x_0 \rangle}{\|x - x_0\|} = 0.$$

Като използваме това определение и дефиницията на производна по посока, получаваме

Твърдение 1. Нека $f : X \to \mathbb{R}^1$ е диференцуема в точката $x \in \operatorname{int} X$. Тогава тя е диференцуема в x по всяка посока $s \in \mathbb{R}^n$ и

$$\frac{\partial f}{\partial s}(x) = \langle f'(x), s \rangle.$$

В частност от това твърдение следва, че

$$f'(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right),$$

където $\dfrac{\partial f}{\partial x_i}(x)$ е частната производна на f по x_i в точката x.

Обратно, ако $f : \mathbb{R}^n \to \mathbb{R}^1$ има частни производни, които са непрекъснати в точката x, то f е диференцуема в x.

Твърдение 2. Нека X е непразно, изпъкнало и отворено множество и $f : X \to \mathbb{R}^1$ е изпъкнала в X и диференцуема във всяка точка на X. Тогава за всяко $h \in \mathbb{R}^n$ производната по посока $\dfrac{\partial f}{\partial h} : X \to \mathbb{R}^1$ е непрекъсната функция в X.

Доказателство. В доказателството на теорема 2 от предишния параграф показахме, че

$$\frac{f(z + \lambda h) - f(z)}{\lambda} \geq \frac{\partial f}{\partial h}(z) = \langle f'(z), h\rangle$$

за всяко $z \in X$, $h \in \mathbb{R}^n$, $\lambda > 0$. Нека редицата $z_n \to x \in X$. Лявата част на неравенството е непрекъсната функция на z, следователно

$$\frac{f(x + \lambda h) - f(x)}{\lambda} \geq \overline{\lim_{z_n \to x}} \langle f'(z_n), h\rangle$$

за всяко $\lambda > 0$. При $\lambda \to 0+$ получаваме

$$\langle f'(x), h\rangle \geq \overline{\lim_{z_n \to x}} \langle f'(z_n), h\rangle,$$

т. е. $\dfrac{\partial f}{\partial h}$ е полунепрекъсната отгоре в точката x за всяко $h \in \mathbb{R}^n$. Като заменим h с $-h$, ще следва

$$-\langle f'(x), h\rangle \geq \overline{\lim_{z_n \to x}} \langle -f'(z_n), h\rangle,$$

откъдето

$$\langle f'(x), h\rangle \leq -\overline{\lim_{z_n \to x}} (-\langle f'(z_n), h\rangle) = \underline{\lim_{z_n \to x}} \langle f'(z_n), h\rangle,$$

т. е. $\dfrac{\partial f}{\partial h}(z)$ е полунепрекъсната отдолу в X. Следователно $\dfrac{\partial f}{\partial h}(z)$ е непрекъсната по z в X. □

Следствие 1. Ако $f : X \to \mathbb{R}^1$ е изпъкнала и диференцуема функция в изпъкналото и отворено множество X, то частните ѝ производни $\dfrac{\partial f}{\partial x_i}$, $i = 1, 2, \dots, n$, са непрекъснати в X, т. е. нейната производна $f' : X \to \mathbb{R}^n$ е непрекъсната функция в X.

Теорема 1. Нека X е непразно отворено и изпъкнало множество и $f : X \to \mathbb{R}^1$ е диференцуема във всяка точка на X. Тогава следните три твърдения са еквивалентни:

1) f е изпъкнала в X;

2) за всяко $x \in X$ и за всяко $y \in X$

$$f(y) \geq f(x) + \langle f'(x), y - x \rangle;$$

3) за всяко $x \in X$ и за всяко $y \in X$

$$\langle f'(x) - f'(y), x - y \rangle \geq 0.$$

Доказателство. Най-напред ще докажем, че от

1) следва 2). Нека f е изпъкнала и $x, y \in X$, $x \neq y$. Тогава за всяко $t \in (0, 1]$ е изпълнено

$$f\left(y + t(x - y)\right) - f(y) \leq t\left(f(x) - f(y)\right).$$

Тъй като f има производна по посоката $s = x - y$, то

$$\frac{\partial f}{\partial s}(y) = \lim_{t \to 0} \frac{f(y + ts) - f(y)}{t} \leq f(x) - f(y).$$

С помощта на твърдение 1 получаваме

$$\langle f'(y), x - y \rangle = \frac{\partial f}{\partial s}(y) \leq f(x) - f(y),$$

откъдето следва 2).

От 2) следва 3): от 2) следва, че за всяка двойка точки $x, y \in X$ са в сила неравенствата

$$f(x) \geq f(y) + \langle f'(y), x - y \rangle,$$
$$f(y) \geq f(x) + \langle f'(x), y - x \rangle.$$

Като ги съберем почленно, получаваме 3).

От 3) следва 1): от 3) следва, че функцията $\varphi(t) = f(x + t(y - x)) = \varphi_{x,y-x}(t)$ е дефинирана в отворен интервал, съдържащ $[0, 1]$, и има непрекъсната и растяща производна φ' в този интервал. Действително, ако $1 \geq t_2 \geq t_1 \geq 0$, то

$$\varphi'(t_2) - \varphi'(t_1) = \langle f'(x + t_2(y - x)) - f'(x + t_1(y - x)), y - x \rangle \geq 0.$$

От монотонността на φ' следва съществуване на лява и дясна граница във всяка точка $t \in [0,1]$, т. е.

$$+\infty > \lim_{\varepsilon\to 0+} \varphi'(t+\varepsilon) \geq \lim_{\varepsilon\to 0+} \varphi'(t-\varepsilon) > -\infty, \quad t \in [0,1].$$

Нека $t_0 \in [0,1]$. От теоремата за средните стойности имаме

$$\frac{\varphi(t)-\varphi(t_0)}{t-t_0} = \varphi'(t_0+\theta(t-t_0)), \quad 0 \leq \theta \leq 1.$$

Ако $t \to t_0$, лявата страна на това равенство клони към производната $\varphi'(t_0)$ независимо от това, дали t клони към t_0 отляво или отдясно. Това означава, че φ' е непрекъсната в t_0, т. е. φ има непрекъсната производна в $[0,1]$.

Нека $0 < \mu < 1$. Тогава

$$\begin{aligned} 0 &\leq \mu(t_2-t_1)\int_0^1 \left[\varphi'(t_1+\Gamma(t_2-t_1)) - \varphi'(t_1+\mu\Gamma(t_2-t_1))\right] d\Gamma \\ &= (1-\mu)\varphi(t_1) + \mu\varphi(t_2) - \varphi((1-\mu)t_1+\mu t_2), \end{aligned}$$

т. е. φ е изпъкнала функция. Но тогава и f е изпъкнала, както доказахме в предишния параграф. □

Определение 1. Нека X е отворено множество в $\mathbb{R}^n$ и $f : X \to \mathbb{R}^1$. Казваме, че функцията f е *двукратно диференцуема* в точката $x_0 \in \operatorname{int} X$, ако съществуват вектор $f'(x_0) \in \mathbb{R}^n$ и симетрична матрица $f''(x_0)$ с размери $n \times n$ такива, че

$$\lim_{\substack{x\to x_0\\ x\neq x_0}} \frac{f(x)-f(x_0)-\langle f'(x_0), x-x_0\rangle - \frac{1}{2}\langle x-x_0, f''(x_0)(x-x_0)\rangle}{\|x-x_0\|^2} = 0.$$

Елементите a_{ij} на матрицата $f''(x_0)$ са вторите частни производни $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$.

Ще припомним, че симетричната матрица A се нарича *неотрицателно* (*положително*) *дефинитна*, ако

$$\langle x, Ax\rangle \geq 0$$

за всяко $x \in \mathbb{R}^n$. Ако това неравенство е строго за всяко $x \neq 0$, матрицата A се нарича *положително* (*строго положително*) *дефинитна*.

Теорема 2. Нека X е изпъкнало отворено множество в $\mathbb{R}^n$, $f : X \to \mathbb{R}^1$ е функция с непрекъснати втори частни производни. Функцията f е изпъкнала в X тогава и само тогава, когато втората производна $f''(x)$ е неотрицателно дефинитна матрица за всяко $x \in X$.

Доказателство. *Необходимост.* Нека $x_0 \in X$, $x \in \mathbb{R}^n$ и $t > 0$ е толкова малко, че $x_0 + tx \in X$. От изпъкналостта на f следва

$$f(x_0 + tx) \geq f(x_0) + t\langle f'(x_0), x\rangle,$$

а от двукратната диференцуемост получаваме

$$f(x_0 + tx) = f(x_0) + t\langle f'(x_0), x\rangle + \frac{1}{2}t^2\langle x, f''(x_0)x\rangle + t^2\|x\|^2 O_{x_0}(tx).$$

Тук O_{x_0} е функция, удовлетворяваща условието $\lim\limits_{x\to 0} O_{x_0}(x) = 0$. От тези две съотношения следва

$$\frac{1}{2}t^2\langle f''(x_0)x, x\rangle + t^2\|x\|^2 O_{x_0}(tx) \geq 0.$$

Като разделим на t^2, след граничния преход $t \to 0$ получаваме, че $f''(x_0)$ е неотрицателно дефинитна.

Достатъчност. Разглеждаме функцията $\varphi(t) = f(x_0 + tx)$, където $x_0 \in X$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}^1$. Лесно се вижда, че

$$\varphi''(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(x_0 + tx)}{\partial x_i \partial x_j} x_i x_j = \langle f''(x_0 + tx)x, x\rangle,$$

т. е. $\varphi''(t) \geq 0$. Тогава $\varphi(t)$ е изпъкнала функция и от твърдение 2 на § 1 следва, че f е изпъкнала функция. Теоремата е доказана. □

Пример 1. Нека A е симетрична матрица с размери $n \times n$ и $b \in \mathbb{R}^n$. Разглеждаме функцията

$$f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle b, x\rangle.$$

Нека $s \in \mathbb{R}^n$. Тогава

$$f(x + s) = f(x) + \langle Ax + b, s\rangle + \frac{1}{2}\langle As, s\rangle.$$

Следователно

$$f'(x) = Ax + b, \quad f''(x) = A$$

и $f(x)$ е изпъкнала тогава и само тогава, когато матрицата A е неотрицателно дефинитна.

Упражнение 1. Намерете пример на функция $f : \mathbb{R}^n \to \mathbb{R}^1$, която е изпъкнала в изпъкналото множество $X \subset \mathbb{R}^n$, но чиято втора производна не е неотрицателно дефинитна.

3. Субградиенти на изпъкнали функции

Определение 1. Нека X е подмножество на $\mathbb{R}^n$, $f : X \to \mathbb{R}^1$ е функция и $x_0 \in X$. Векторът $x^* \in \mathbb{R}^n$ се нарича *субградиент* на f в точката x_0, ако за всяко $x \in X$ е изпълнено

$$f(x) - f(x_0) \geq \langle x^*, x - x_0 \rangle.$$

Това неравенство може да се запише и по следния начин:

$$(-1)(r - f(x_0)) + \langle x^*, x - x_0 \rangle \leq 0$$

за всяко $(r, x) \in \operatorname{epi} f$. Оттук следва геометричната интерпретация на субградиента (фиг. 26): векторът $(-1, x^*)$ е нормален вектор на хиперравнината, опорна към надграфиката на f в точката $(f(x_0), x_0)$.

Ясно е, че в дадена точка f може да има повече от един субградиент.

Определение 2. Множеството от всички субградиенти на f в точката x_0 се нарича *субдиференциал* на f в точката x_0 и се бележи с $\partial_f(x_0)$:

$$\partial_f(x_0) = \{x^* \in \mathbb{R}^n : f(x) - f(x_0) \geq \langle x^*, x - x_0 \rangle \text{ за всяко } x \in X\}.$$

Пример 1. Функцията $f(x) = \|x\|$ удовлетворява неравенството

$$f(x) - f(0) \geq \langle c, x \rangle = \langle c, x - 0 \rangle$$

за всяко $x \in \mathbb{R}^n$ при произволно c, $\|c\| \leq 1$. Следователно $\partial_f(0) = \{c \in \mathbb{R}^n : \|c\| \leq 1\}$.

Упражнение 1. Докажете съотношенията

1) $\partial_{\lambda f}(x) = \lambda \partial_f(x)$, $\lambda > 0$;

2) $\partial_f(\lambda x) = \lambda \partial_f(x)$, $\lambda > 0$.

Упражнение 2. Докажете, че ако $x_1^* \in \partial_f(x_1)$ и $x_2^* \in \partial_f(x_2)$, то

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq 0.$$

Твърдение 1. Нека X е отворено множество, $x_0 \in X$ и функцията $f : X \to \mathbb{R}^1$ е непрекъсната в точката x_0. Тогава множеството $\partial_f(x_0)$ е изпъкнало и компактно.

Доказателство. Изпъкналостта и затвореността на $\partial_f(x_0)$ следват непосредствено от определението на субдиференциала. Ще докажем, че $\partial_f(x_0)$ е ограничено множество. Нека $\varepsilon > 0$ и нека $B(x_0, t)$, $t > 0$, е такова кълбо, че $f(x) - f(x_0) < \varepsilon$ при $x \in B(x_0, t)$. Тогава за всяко $x^* \in \partial_f(x_0)$ и $x \in B(x_0, t)$ е изпълнено

$$\varepsilon > f(x) - f(x_0) \geq \langle x^*, x - x_0 \rangle.$$

При достатъчно малко $t > 0$ $x = x_0 + tx^* \in B(x_0, t)$. Тогава $\varepsilon > \langle x^*, tx^* \rangle = t\|x^*\|^2$, т. е. $\|x^*\| < \sqrt{\dfrac{\varepsilon}{t}}$, с което твърдението е доказано. □

Твърдение 2. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ и $\partial_f(x) \neq \emptyset$ във всяка точка $x \in \mathbb{R}^n$. Тогава f е изпъкнала в $\mathbb{R}^n$.

Доказателство. Нека $x_1, x_2 \in \mathbb{R}^n$ и $\lambda \in (0, 1)$. По предположение съществува субградиент x^* в точката $x = \lambda x_1 + (1 - \lambda)x_2$, следователно са изпълнени неравенствата

$$f(x_1) \geq f(\lambda x_1 + (1 - \lambda)x_2) + (1 - \lambda)\langle x^*, x_1 - x_2 \rangle,$$
$$f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2) + \lambda\langle x^*, x_2 - x_1 \rangle.$$

Като умножим първото неравенство с λ, а второто с $1 - \lambda$ и съберем двете неравенства, получаваме дефиницията за изпъкналост на функция. □

Упражнение 3. Нека X е напразно изпъкнало множество, $f : X \to \mathbb{R}^1$ и нека $\partial_f(x) \neq \emptyset$ за всяко $x \in \operatorname{int} X$. Докажете, че f е изпъкнала в $\operatorname{int} X$.

Твърдение 3. Нека X е отворено изпъкнало множество, $f : X \to \mathbb{R}^1$ е изпъкнала функция, $x_0 \in X$ и $s \in \mathbb{R}^n$. Тогава съществува субградиент $y^* \in \partial_f(x_0)$, за който

$$\frac{\partial f}{\partial s}(x_0) = \langle y^*, s \rangle = \max_{x^* \in \partial_f(x_0)} \langle x^*, s \rangle,$$

където $\dfrac{\partial f}{\partial s}(x_0)$ е производната на f в точката x_0 по посока s.

Доказателство. Съществуването на производна по посока за изпъкнали функции доказахме в § 1. Съгласно определението на субградиент

$$f(x_0 + \lambda s) - f(x_0) \geq \lambda\langle x^*, s \rangle,$$

откъдето $\dfrac{\partial f}{\partial s}(x_0) \geq \langle x^*, s \rangle$ за всяко $x^* \in \partial_f(x_0)$. Сега ще намерим субградиент y^*, за който $\dfrac{\partial f}{\partial s}(x_0) = \langle y^*, s \rangle$.

Да разгледаме множествата

$$Y_1 = \left\{ \left(x_0 + ts, f(x_0) + t\frac{\partial f}{\partial s}(x_0) \right) \in \mathbb{R}^{n+1} : t \geq 0 \right\},$$

$$Y_2 = \{(x, r) \in \mathbb{R}^{n+1} : r > f(x),\ x \in X\}.$$

Лесно се проверява, че множествата Y_1 и Y_2 са изпъкнали и $Y_1 \cap Y_2 = \emptyset$. Можем да приложим теоремата за отделимост (вж. теорема 3, § 8, гл. 2): съществува такъв ненулев вектор, $(x^*, r^*) \in \mathbb{R}^{n+1}$, че

$$\langle (x^*, r^*), (x, r)\rangle \geq \left\langle (x^*, r^*), \left(x_0 + ts, f(x_0) + t\frac{\partial f}{\partial s}(x_0)\right)\right\rangle$$

за всяко $t \geq 0$ и за всяко (x, r), за което $r > f(x)$, $x \in X$. Това е еквивалентно на

(1) $$\langle x^*, x\rangle + rr^* \geq \langle x^*, x_0 + ts\rangle + r^*\left(f(x_0) + \frac{\partial f}{\partial s}(x_0)t\right).$$

Ако допуснем, че $r^* = 0$, ще получим

(2) $$\langle x^*, x\rangle \geq \langle x^*, x_0 + ts\rangle \text{ за всяко } x \in X \text{ и } t \geq 0.$$

Векторът $x = x_0 + ts - tx^*$ принадлежи на отвореното множество X, когато $t > 0$ е достатъчно малко число. Ако заместим това x в (2), получаваме неравенството $-t\langle x^*, x^*\rangle \geq 0$, което е възможно само ако $x^* = 0$. Но това противоречи на факта, че (r^*, x^*) е ненулев вектор. И така $r^* \neq 0$. Ако $r^* < 0$, лявата страна на (1) може да става произволно малка (отрицателна), следователно $r^* > 0$. Като разделим двете страни на (1) на r^* и извършим граничен преход $r \to f(x)$, получаваме неравенството

(3) $$f(x) - f(x_0) \geq \left\langle -\frac{x^*}{r^*}, x - x_0\right\rangle + \left\langle \frac{x^*}{r^*}, ts\right\rangle + t\frac{\partial f}{\partial s}(x_0),$$

което е изпълнено за всяко $t \geq 0$ и $x \in X$. При $t = 0$ получаваме

$$f(x) - f(x_0) \geq \left\langle -\frac{x^*}{r^*}, x - x_0\right\rangle,$$

следователно $y^* = -\dfrac{x^*}{r^*} \in \partial_f(x_0)$. От друга страна, при $x = x_0$ и $t > 0$ от (3) следва $0 \geq t\left(\dfrac{\partial f}{\partial s}(x_0) - \langle y^*, s\rangle\right)$ или $\dfrac{\partial f}{\partial s}(x_0) \leq \langle y^*, s\rangle$. С това доказателството е завършено. □

Следствие 1. Функцията $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала тогава и само тогава, когато за всяко $x \in \mathbb{R}^n$ имаме $\partial_f(x) \neq \emptyset$.

Твърдение 4. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала функция и $x_0 \in \mathbb{R}^n$. За всяко $\delta > 0$ съществува такова $c > 0$, че за всяко $x \in B(x_0, \delta)$ субдиференциалът $\partial_f(x)$ се съдържа в кълбото $B(0, c)$.

Доказателство. В § 1 доказахме, че ако $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала функция, тя е непрекъсната. Следователно за всяко $x_0 \in \mathbb{R}^n$ и за всяко $\delta > 0$ съществува константа $c_1 > 0$ такава, че $|f(x)| \leq c_1$ за $x \in B(x_0, \delta)$. Да изберем $x' \in B(x_0, \delta/2)$. Тогава

$$2c_1 \geq f(x'') - f(x') \geq \langle x^*, x'' - x' \rangle$$

за всяко $x'' \in B(x_0, \delta)$ и $x^* \in \partial_f(x')$. Нека $x^* \in \partial_f(x')$, $x^* \neq 0$ и нека $x'' = x' + \dfrac{\delta x^*}{2\|x^*\|}$. Тогава $x'' \in B(x_0, \delta)$ и от горното неравенство следва $\|x^*\| \leq \dfrac{4c_1}{\delta}$. Като изберем $c = \dfrac{4c_1}{\delta}$, завършваме доказателството. □

Твърдение 5. Нека X е изпъкнало множество и $f : X \to \mathbb{R}^1$ е изпъкнала функция, която е диференцуема в точката $x_0 \in \operatorname{int} X$. Тогава $\partial_f(x_0) = f'(x_0)$.

Доказателство. От следствие 1 имаме $\partial_f(x_0) \neq \emptyset$. Нека $x^* \in \partial_f(x_0)$. Тогава за всяко $s \in \mathbb{R}^n$ и за достатъчно малко $t > 0$

$$\begin{aligned} &f(x_0 + ts) \geq f(x_0) + t\langle x^*, s \rangle, \\ &f(x_0 + ts) = f(x_0) + t\langle f'(x_0), s \rangle + t\|s\| O_{x_0}(ts), \end{aligned}$$

където $\lim\limits_{t \to 0} O_{x_0}(ts) = 0$. Това означава, че

$$0 \geq \langle x^* - f'(x_0), s \rangle - \|s\| O_{x_0}(ts)$$

за всяко достатъчно малко t, т. е. $\langle x^* - f'(x_0), s \rangle \leq 0$ за всяко $s \in \mathbb{R}^n$. Ако $s = x^* - f'(x_0)$, ще получим че $x^* = f'(x_0)$. □

Теорема 1. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала функция и X е непразно изпъкнало множество в $\mathbb{R}^n$. Разглеждаме задачата

$$f(x) \to \min, \quad x \in X.$$

Точката x_0 е решение на тази задача тогава и само тогава, когато съществува такъв субградиент $x^* \in \partial_f(x_0)$, че за всяко $x \in X$

$$\langle x^*, x - x_0 \rangle \geq 0. \tag{4}$$

Доказателство. Нека x^* е субградиент, за който е изпълнено (4) за всяко $x \in X$. По дефиниция

$$f(x) \geq f(x_0) + \langle x^*, x - x_0 \rangle \geq f(x_0),$$

следователно x_0 е решение на задачата.

Обратно, нека x_0 е решение на задачата. Да построим следните множества в $\mathbb{R}^{n+1}$:

$$\begin{aligned} Y_1 &= \{(x - x_0, r) : x \in \mathbb{R}^n,\ r > f(x) - f(x_0)\}, \\ Y_2 &= \{(x - x_0, r) : x \in X,\ r \leq 0\}. \end{aligned}$$

Лесно се вижда, че Y_1 и Y_2 са изпъкнали множества. Освен това $Y_1 \cap Y_2 = \emptyset$. Действително , в противен случай ще съществува $x \in X$ със свойството $f(x) - f(x_0) < 0$, което противоречи на условието, че x_0 е решение. Според теорема 3 в § 8 от гл. 2 съществуват такъв нулев вектор $(\xi, \mu) \in \mathbb{R}^{n+1}$ и такова число α, че

(5) $\quad \langle \xi, x - x_0 \rangle + \mu r \leq \alpha$ за всяко (x, r), $x \in \mathbb{R}^n$, $r \geq f(x) - f(x_0)$,

(6) $\quad \langle \xi, x - x_0 \rangle + \mu r \geq \alpha$ за всяко (x, r), $x \in X$, $r \leq 0$.

Ако в (6) положим $x = x_0$ и $r = 0$, ще видим, че $\alpha \leq 0$. Да положим $x = x_0$ в (5). Тогава $\mu r \leq \alpha$ за всяко $r \geq f(x_0) - f(x_0) = 0$. Следователно $\mu \leq 0$ и $\alpha = 0$. Ако $\mu = 0$, от (5) следва $\langle \xi, x - x_0 \rangle \leq 0$ за всяко $x \in \mathbb{R}^n$. От това следва, че $\xi = 0 \in \mathbb{R}^n$. Това е противоречие, защото $(\xi, \mu) \neq 0$. Следователно $\mu < 0$. Като разделим (5) и (6) на $-\mu$ и означим $\dfrac{\xi}{\mu} = x^*$, получаваме

(7) $\quad r \geq \langle x^*, x - x_0 \rangle$ за всяко (x, r), $x \in \mathbb{R}^n$, $r > f(x) - f(x_0)$,

(8) $\quad \langle x^*, x - x_0 \rangle \geq r$ за всяко (x, r), $x \in X$, $r \leq 0$.

Ако положим $r = 0$ в (8), получаваме, че $\langle x^*, x - x_0 \rangle \geq 0$ за всяко $x \in X$. От (7) очевидно следва, че $f(x) - f(x_0) \geq \langle x^*, x - x_0 \rangle$ за всяко $x \in \mathbb{R}^n$. Следователно x^* е субградиент на f в x_0, за който е изпълнено (4). □

Следствие 2. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала функцията. Точката x_0 е точка на минимум на f в $\mathbb{R}^n$ тогава и само тогава, когато $0 \in \partial_f(x_0)$.

Задача 1. Намерете точките на минимум на функцията

$$f(x_1, x_2) = x_1^2 + x_2^2 + 2|x_1 + x_2 - 1|.$$

Задача 2. Докажете, че ако $f_i : \mathbb{R}^n \to \mathbb{R}^1$, $i = 1, 2, \ldots, m$, са изпъкнали функции, то

$$\partial_{\sum\limits_{i=1}^{m} f_i}(x) = \sum_{i=1}^{m} \partial_{f_i}(x).$$

Задача 3. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала. Намерете субградиента на функцията $g(x) = \max\{0, f(x)\}$.

Задача 4. Докажете, че ако $f : X \to \mathbb{R}^1$ е изпъкнала функция, то f е липшицова във всяко ограничено множество, съдържащо се в $\operatorname{int} X$.

Глава 5

Теорема на Кун и Такър

В тази глава ще разглеждаме задачи за минимум на изпъкнала функция f в изпъкнало подмножество на X на $\mathbb{R}^n$. Задачите от този клас имат редица интересни свойства: локалните минимуми са и глобални минимуми, необходимите условия за минимум стават необходими и достатъчни условия. Когато множеството X е зададено със система от неравенства

$$x \in X = \{x \in \mathbb{R}^n : g_j(x) \leq 0, \ \ j = 1, 2, \ldots, m\}$$

и функциите f и g_j, $j = 1, 2, \ldots, m$, са изпъкнали в $\mathbb{R}^n$, задачата

$$f(x) \to \min, \quad x \in X,$$

се нарича задача на *изпъкналото оптимиране*. В § 2 и 3 ще представим в различни варианти един от централните резултати на математическото оптимиране — теоремата на Кун и Такър, даваща необходимо и достатъчно условие за минимум. Чрез тази теорема се обобщава идеята за множители на Лагранж и се дава метод за решаване на конкретни задачи.

1. Екстремални свойства на изпъкналите функции

Нека f е реална функция, дефинирана в $\mathbb{R}^n$, и нека X е подмножество на $\mathbb{R}^n$.

Определение 1. Точката $x_0 \in X$ се нарича *точка на локален минимум* (или просто локален минимум) за f в X, ако съществува $\varepsilon > 0$, за което от $\|x - x_0\| \leq \varepsilon$ и $x \in X$ следва $f(x) \geq f(x_0)$.

Определение 2. Точката $x_0 \in X$ се нарича *точка на глобален минимум* (глобален минимум) за f в X, ако $f(x) \geq f(x_0)$ за всяко $x \in X$.

Теорема 1. Ако функцията f е изпъкнала и X е изпъкнало множество, локалните минимуми на f в X са глобални максимуми.

Доказателство. Нека x_0 е локален минимум на f в X. Тогава съществува такова $\varepsilon > 0$, че от $x \in X$ и $\|x - x_0\| \le \varepsilon$ следва $f(x) \ge f(x_0)$. Нека $y \in X$ и $\|y - x_0\| > \varepsilon$. Ще докажем, че $f(y) \ge f(x_0)$.

Да разгледаме точката z със свойството $\|z - x_0\| = \varepsilon$ и лежаща на отсечка с краища x_0 и y (фиг. 27). Тази точка може да се представи по следния начин:

$$z = \frac{\varepsilon}{\|y - x_0\|} y + \left(1 - \frac{\varepsilon}{\|y - x_0\|}\right) x_0 = x_0 + \frac{\varepsilon}{\|y - x_0\|}(y - x_0).$$

Поради изпъкналостта на X точката z е от X. Освен това $\|z - x\| = \varepsilon$. Следователно $f(z) \ge f(x_0)$. Понеже f е изпъкнала функция, получаваме

$$f(x_0) \le f(z) \le \frac{\varepsilon}{\|y - x_0\|} f(y) + \left(1 - \frac{\epsilon}{\|y - x_0\|}\right) f(x_0),$$

т. е.

$$0 \le \frac{\varepsilon}{\|y - x_0\|}(f(y) - f(x_0)).$$

Следователно $f(y) \ge f(x_0)$. Теоремата е доказана. $\square$

Определение 3. Функцията $f : X \to \mathbb{R}^1$, $X \subset \mathbb{R}^n$, се нарича *строго изпъкнала*, ако $f(tx_1 + (1 - t)x_2) < tf(x_1) + (1 - t)f(x_2)$, където $x_1, x_2 \in X$, $x_1 \ne x_2$ и $0 < t < 1$.

Твърдение 1. Ако функцията f е строго изпъкнала в изпъкналото множество X и x_0 е локален минимум на f в X, то x_0 е единственият глобален минимум на f в X.

Доказателство. Според предишната теорема x_0 е глобален минимум. Да допуснем, че съществува точка $x \in X$, $x \ne x_0$, за която $f(x) = f(x_0)$. От строгата изпъкналост следва, че

$$f\left(\frac{x + x_0}{2}\right) < \frac{1}{2} f(x) + \frac{1}{2} f(x_0) = f(x_0).$$

Тъй като X е изпъкнало множество, то $\frac{1}{2}(x + x_0) \in X$ и полученото неравенство е в противоречие с това, че x_0 е глобален минимум. $\square$

Упражнение 1. Нека $f : X \to \mathbb{R}^1$ е изпъкнала функция и $X \subset \mathbb{R}^n$ е изпъкнало множество. Докажете, че множеството от точки на минимум

$$X_0 = \left\{x_0 \in X : f(x_0) = \inf_X f(x)\right\}$$

е изпъкнало.

Теорема 2. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ е диференцуема функция в $\mathbb{R}^n$, която е изпъкнала в изпъкналото множество X. Точката x_0 е (глобален) минимум на f в X тогава и само тогава, когато за всяко $x \in X$

$$\langle f'(x_0), x - x_0 \rangle \geq 0. \tag{1}$$

Доказателство. Ако x_0 е точка на минимум, то за всяко $x \in X$ и $0 < t \leq 1$

$$\frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} \geq 0.$$

След граничен преход $t \to 0$ получаваме (1). Обратно, ако е изпълнено (1), то от теорема 1 на § 2, гл. 4 имаме $f(x) - f(x_0) \geq \langle f'(x_0), x - x_0 \rangle \geq 0$, т. е. x_0 е точка на минимум. Ще отбележим, че доказателството следва непосредствено и от твърдение 5 и теорема 1 от § 3 на гл. 4. □

От предишната теорема и теоремата на Ферма (вж. следствие 1, § 3, гл. 1) получаваме

Следствие 1. Нека множеството X е отворено и изпъкнало и функцията $f : X \to \mathbb{R}^1$ е диференцуема и изпъкнала в X. Точката x_0 е точка на минимум на f в X тогава и само тогава, когато $f'(x_0) = 0$.

Задача 1. Нека X е интервалът $[a, b]$ и $f : X \to \mathbb{R}^1$ е непрекъсната функция, за която всяка точка x, $a < x < b$, е точка на локален минимум. Докажете, че $f(x)$ приема само една стойност, т. е. f е константа.

Задача 2. Нека f е както в предишната задача, но има локален екстремум във всяка точка x, $a < x < b$. Докажете, че f е константа.

2. Афинни ограничения

Да разгледаме задачата

$$\begin{aligned} & f(x) \to \min, \\ & \langle a_i, x \rangle \leq b_i, \quad i = 1, 2, \ldots, l. \end{aligned} \tag{1}$$

където $f : \mathbb{R}^n \to \mathbb{R}^1$, $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}^1$, $i = 1, 2, \ldots, l$.

В § 3 на гл. 1 при предположение за диференцуемост на f доказахме следното: ако x^0 е решение на задачата (1), то съществуват такива неотрицателни числа λ_i^0, $i = 1, 2, \ldots, l$, че са изпълнени условията

$$f'(x^0) + \sum_{j=1}^{l} \lambda_j^0 a = 0, \tag{2}$$

(3)
$$\sum_{j=1}^{l} \lambda_j^0(\langle a_j, x^0\rangle - b_j) = 0.$$

Да съставим функцията на Лагранж за задачата (1):

$$L(x,\lambda) = f(x) + \sum_{j=1}^{l} \lambda_j(\langle a_j, x\rangle - b_j),$$

където $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l) \in \mathbb{R}^l$. Условието (2) означава, че $L'_x(x^0, \lambda^0) = 0$, т.е. x^0 е стационарна точка в $\mathbb{R}^n$ на функцията $L(x, \lambda^0)$. Ако допълнително предположим, че f е изпъкнала в $\mathbb{R}^n$, то $L(x, \lambda^0)$ е изпъкнала в $\mathbb{R}^n$ и условието (2) е еквивалентно на следното (вж. следствие 1 от § 1):

(4)
$$L(x^0, \lambda^0) \le L(x, \lambda^0), \quad x \in \mathbb{R}^n.$$

От друга страна, от (3) и от неравенството

$$\sum_{j=1}^{l} \lambda_j(\langle a_j, x^0\rangle - b_j) \le 0$$

за всяко $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l)$, $\lambda_j \ge 0$, $j = 1, 2, \dots, l$, следва

(5)
$$L(x^0, \lambda) \leqq L(x^0, \lambda^0).$$

Но (4) и (5) означават, че двойката (x^0, λ^0) е седлова точка на функцията на Лагранж в областта $\mathbb{R}^n \times \mathbb{R}^l_+$ ($\mathbb{R}^l_+ = \{y \in \mathbb{R}^l : y_j \ge 0,\ j = 1, \dots, l\}$).

И така при предположение за изпъкналост необходимите условия и достатъчните условия за оптималност за задачата (1), с които се запознахме в § 3 и 6 на гл. 1, стават едновременно *необходими и достатъчни условия*. Това твърдение ще докажем за различни изпъкнали задачи на математическото оптимиране без предположение за диференцуемост на функцията $f(x)$.

Теорема 1. Нека функцията $f : \mathbb{R}^n \to \mathbb{R}^1$ е изпъкнала в $\mathbb{R}^n$ и x^0 е решение на задачата (1). Тогава съществува такъв неотрицателен вектор $\lambda^0 = (\lambda_1^0, \lambda_2^0, \dots, \lambda_l^0)$, че двойката (x^0, λ^0) е седлова точка на функцията на Лагранж в областта $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^l_+$, т.е.

$$L(x^0, \lambda) \le L(x^0, \lambda^0) \le L(x, \lambda^0)$$

за всяко $x \in \mathbb{R}^n$ и за всяко $\lambda \in \mathbb{R}^l_+$.

Доказателство. Да означим с I множеството от индекси на активните ограничения в точката x^0, т. е.

$$I = \{j \in \{1, \ldots, l\} : \langle a_j, x^0 \rangle = b_j\}.$$

Множествата

$$Y_1 = \{(r, x) \in \mathbb{R}^{1+n} : r > f(x),\ \ x \in \mathbb{R}^n\},$$
$$Y_2 = \{(f(x^0), x) \in \mathbb{R}^{1+n} : \langle a_j, x - x^0 \rangle \leq 0,\ \ j \in I\}$$

са непразни и изпъкнали (докажете) и

$$Y_1 \cap Y_2 = \emptyset. \tag{6}$$

За да докажем (6), да допуснем, че съществува такова $x^* \in \mathbb{R}^n$, че $(f(x^0, x^*) \in Y_1 \cap Y_2)$. Тогава $f(x^0) > f(x^*)$ и $\langle a_j, x^* - x^0 \rangle \leq 0,\ j \in I$. За всяко $t > 0$ и за всяко $j \in I$ точката $x_t = x^0 + t(x^* - x^0)$ удовлетворява неравенството

$$\langle a_j, x_t \rangle = \langle a_j, x^0 \rangle + t\langle a_j, x^* - x^0 \rangle \leq b_j.$$

За $j \notin I$

$$\langle a_j, x_t \rangle = \langle a_j, x^0 \rangle + t\langle a_j, x^* - x^0 \rangle < b_j + t\langle a_j, x^* - x^0 \rangle \leq b_j$$

за достатъчно малко $t > 0$. Следователно x_t е допустима точка за задачата (1) при достатъчно малко $t > 0$.

От друга страна, от изпъкналостта на $f(x)$ имаме

$$f(x_t) \leq tf(x^*) + (1 - t)f(x^0)$$

и от $f(x^0) \leq f(x_t)$ следва

$$f(x^0) \leq tf(x^*) + (1 - t)f(x^0),$$

т. е. $f(x^0) \leq f(x^*)$. Но това противоречи на условието, че $(f(x^0), x^*) \in Y_1$. С това доказахме (6).

Като приложим теорема 3 (за отделимост) от § 8 на гл. 2, отделяме Y_1 и Y_2 с хиперравнина, т. е. намираме ненулев вектор $(\alpha, u) \in \mathbb{R}^{1+n}$, за който е изпълнено

$$\langle (\alpha, u), (r, y) \rangle \geq \langle (\alpha, u), (f(x^0), x) \rangle,$$

т. е.

$$\alpha r + \langle u, y \rangle \geq \alpha f(x^0) + \langle u, x \rangle \tag{7}$$

за всеки $y \in \mathbb{R}^n$, $r \in \mathbb{R}^1$ и $x \in \mathbb{R}^n$, за които $r > f(y)$ и $\langle a_j, x - x^0 \rangle \le 0$, $j \in I$.

Ясно е, че $\alpha \ge 0$ (в противен случай при $r \to +\infty$ ще получим противоречие със (7)). Ако допуснем, че $\alpha = 0$, от (7) следва

$$\langle u, y \rangle \ge \langle u, x^0 \rangle$$

за всяко $y \in \mathbb{R}^n$, което е възможно само ако $u = 0$. Но тогава и $(\alpha, u) = 0$, което е невъзможно, защото (α, u) е нулев вектор. Полагаме в (7) $v = \frac{1}{\alpha} u$ и получаваме

$$r + \langle v, y \rangle \ge f(x^0) + \langle v, x \rangle \tag{8}$$

за всяко $y \in \mathbb{R}^n$, всяко $r \in \mathbb{R}^1$, $r > f(y)$, и всяко x, за което $\langle a_j, x - x^0 \rangle \le 0$, $j \in I$. След граничния преход $r \to f(y)$ неравенството (8) очевидно ще остане в сила. И така

$$f(y) + \langle v, y \rangle \ge f(x^0) + \langle v, x \rangle \tag{9}$$

за всяко $y \in \mathbb{R}^n$ и $x \in \mathbb{R}^n$, $\langle a_j, x - x^0 \rangle \le 0$, $j \in I$. При $y = x^0$ от (9) получаваме $\langle v, x - x^0 \rangle \le 0$, когато $\langle a_j, x - x^0 \rangle \le 0$, $j \in I$. Прилагаме теоремата на Фаркаш и получаваме, че съществуват неотрицателни числа λ_j^0, $j \in I$, за които

$$v = \sum_{j \in I} \lambda_j^0 a_j.$$

Тогава (9) се записва по следния начин:

$$f(x^0) + \sum_{j \in I} \lambda_j^0 \langle a_j, x \rangle \le f(y) + \sum_{j \in I} \lambda_j^0 \langle a_j, y \rangle$$

за всяко $y \in \mathbb{R}^n$.

Като извадим от двете страни на това неравенство $\sum_{j \in I} \lambda_j^0 b_j$, ще получим при $x = x^0$

$$f(x^0) \le f(y) + \sum_{j \in I} \lambda_j^0 (\langle a_j, y \rangle - b_j), \quad y \in \mathbb{R}^n.$$

Полагаме $\lambda_j^0 = 0$, $j \notin I$, и получаваме

$$\sum_{j \in I} \lambda_j^0 (\langle a_j, y \rangle - b_j) = 0$$

и

$$L(x^0, \lambda^0) = f(x^0) \le L(x, \lambda^0).$$

Това неравенство и неравенството (5) означават, че (x^0, λ^0) е седлова точка. С това доказателството е завършено. □

3. Теорема на Кун и Такър

В този параграф ще докажем твърдения, аналогични на теорема 1 от предишния параграф, за поредица от последователно усложняващи се задачи на изпъкналото оптимиране.

1. Нека X е изпъкнало подмножество на $\mathbb{R}^n$ и функциите $f : X \to \mathbb{R}^1$, $g_j : X \to \mathbb{R}^1$, $j = 1, 2, \dots, m$, са изпъкнали в X. Разглеждаме следната задача на изпъкналото оптимиране:

$$(1) \qquad \begin{gathered} f(x) \to \min, \\ g_j(x) \le 0, \quad j = 1, 2, \dots, m, \quad x \in X. \end{gathered}$$

Въвеждаме функцията на Лагранж

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

Както вече знаем, двойката (x^0, λ^0) е седлова точка на функцията на Лагранж в областта $x \in X$, $\lambda \in \mathbb{R}^m_+$, когато

$$(2) \qquad L(x^0, \lambda) \le L(x^0, \lambda^0) \le L(x, \lambda^0)$$

за всяко $x \in X$, $\lambda \in \mathbb{R}^m_+$.

Ще използваме следното

Условие на Слейтър. Съществува точка $\overline{x} \in X$, за която $g_j(\overline{x}) < 0$, $j = 1, 2, \dots, m$.

Упражнение 1. Докажете, че условието на Слейтър е еквивалентно на условието: за всяко $j \in \{1, \dots, m\}$ съществува такова $\overline{x}^j \in X$, че $g_j(\overline{x}^j) < 0$.

Следващата теорема, формулирана през 1951 г., съдържа един от основните резултати на математическото оптимиране:

Теорема 1 (Кун—Такър)**.** Нека x^0 е решение на задачата (1) и нека условието на Слейтър е изпълнено. Тогава съществува такъв неотрицателен вектор $\lambda^0 = (\lambda^0_1, \dots, \lambda^0_m)$, че двойката (x^0, λ^0) е седлова точка на функцията на Лагранж в областта $x \in X$, $\lambda \in \mathbb{R}^m_+$.

Доказателство. Основната идея на доказателството отново се състои в прилагането на теорема за отделимост — в случая на теорема 1 (за опорна хиперравнина) от § 7 на гл. 2. Разглеждаме множеството

$$Y = \{y = (y_0, \dots, y_m) \in \mathbb{R}^{1+m} : \text{съществува } x \in X, \\ \text{за което } y_0 \ge f(x),\ y_j \ge g_j(x),\ j = 1, \dots, m\}.$$

Не е трудно да се види,че множеството Y е изпъкнало и има непразна вътрешност (докажете!). Точката $z^0 = (f(x^0), 0, \dots, 0)$ е контурна точка за Y. Действително, за всяко $\varepsilon > 0$ имаме $z_{+\varepsilon} = (f(x^0) + \varepsilon, 0, \dots, 0) \in Y$, но $z_{-\varepsilon} = (f(x^0) - \varepsilon, 0, \dots, 0) \notin Y$, защото x^0 е решение на (1). Следователно съществува ненулев вектор $v = (v_0, v_1, \dots, v_m)$, за който неравенството

(3) $$\langle v, y\rangle \geq \langle v, z^0\rangle \quad (= v_0 f(x^0))$$

е в сила за всяко $y \in Y$. Като вземем предвид, че координатите на всяко $y \in Y$ са неограничени отгоре (т. е. могат да бъдат произволно големи положителни числа), получаваме $v \geq 0$. Ще покажем, че $v_0 > 0$. Да допуснем противното, т. е. $v_0 = 0$, и да вземем точката $\overline{y} = (f(\overline{x}), g_1(\overline{x}), \dots, g_m(\overline{x}))$, където $\overline{x}$ е точката от условието на Слейтър. От (3) следва, че

$$\langle v, \overline{y}\rangle = \sum_{j=1}^{m} v_j g_j(\overline{x}) \geq \langle v, z^0\rangle = 0.$$

От друга страна, $v_j \geq 0$ и $g_j(\overline{x}) < 0$, $j = 1, 2, \dots, m$, следователно

$$\sum_{j=1}^{m} v_j g_j(\overline{x}) \leq 0.$$

Това е възможно само ако $v_j g_j(\overline{x}) = 0$ за всяко $j = 1, 2, \dots, m$, което от своя страна означава, че $v_j = 0$, $j = 1, 2, \dots, m$. Получаваме $v = 0$, което е противоречие. Следователно $v_0 > 0$.

Да положим $\lambda_i^0 = \dfrac{v_i}{v_0} \geq 0$. От (3) получаваме, че за всяко $x \in X$

(4) $$f(x) + \sum_{j=1}^{m} \lambda_j^0 g_j(x) \geq f(x^0).$$

При $x = x^0$ от това неравенство следва неравенството

$$\sum_{j=1}^{m} \lambda_j^0 g_j(x^0) \geq 0.$$

Но $\lambda_j^0 \geq 0$ и $g_j(x^0) \leq 0$, $j = 1, 2, \dots, m$, следователно

(5) $$\sum_{j=1}^{m} \lambda_j^0 g_j(x) = 0.$$

Условията (4) и (5) означават, че

$$f(x^0) = L(x^0, \lambda^0) \le L(x, \lambda^0), \quad x \in X.$$

От друга страна, тъй като $g_j(x^0) \le 0$, $j = 1, 2, \ldots, m$, то за всяко $\lambda \in \mathbb{R}^m_+$ е изпълнено

$$\sum_{j=1}^{m} \lambda_j g_j(x^0) \le 0.$$

Следователно

$$L(x^0, \lambda) \le L(x^0, \lambda^0) = f(x^0), \quad \lambda \in \mathbb{R}^m_+$$

С това доказателството е завършено. □

Пример 1. Ако условието на Слейтър не е изпълнено, теоремата на Кун и Такър може и да не е в сила. Да разгледаме задачата

$$f(x) = x \to \min, \quad g_1(x) = x^2 \le 0.$$

Решението на задачата е $x^0 = 0$ и условието на Слейтър не е изпълнено. Ако съществуваше $\lambda^0 \ge 0$, за което $(0, \lambda^0)$ да е седлова точка на функцията на Лагранж, то за всяко $x \in \mathbb{R}^1$ е в сила неравенството

$$L(x, \lambda) = x + \lambda^0 x^2 \ge 0.$$

Следователно $x \ge -\dfrac{1}{\lambda^0}$ за всяко $x \in \mathbb{R}^1$. Това е противоречие.

2. Да разгледаме отново задачата (1), предполагайки, че изпъкналото множество X е зададено със система от афинни неравенства

$$X = \{x \in \mathbb{R}^n : \langle a_j, x\rangle \le b_j, j = 1, 2, \ldots, l\},$$

а функциите f и g_j, $j = 1, 2, \ldots, m$, са дефинирани и изпъкнали в X. Тогава условието

$$L(x^0, \lambda^0) \le L(x, \lambda^0), \quad x \in X,$$

означава, че решението x^0 на (1) е едновременно решение на задачата за минимум с афинни ограничения

$$\begin{aligned} &L(x, \lambda^0) \to \min, \\ &\langle a_j, x\rangle \le b_j, \quad j = 1, \ldots, l. \end{aligned} \tag{6}$$

Тази задача разгледахме в предишния параграф и там доказахме: съществува такъв неотрицателен вектор $\mu^0 = (\mu^0_1, \ldots, \mu^0_l) \in \mathbb{R}^l_+$, че двойката (x^0, μ^0) е седлова точка на функцията на Лагранж

$$\mathcal{L}(x, \mu) = L(x, \lambda^0) + \sum_{j=1}^{l} \mu_j(\langle a_j, x\rangle - b_j),$$

при това

$$\sum_{j=1}^{l} \mu_j^0(\langle a_j, x^0\rangle - b_j) = 0. \tag{7}$$

Следователно за всеки $x \in \mathbb{R}^n$ и $\mu \in \mathbb{R}^l_+$

$$\mathcal{L}(x^0, \mu) \leq L(x^0, \lambda^0) = \mathcal{L}(x^0, \mu^0) \leq \mathcal{L}(x, \mu^0).$$

Тогава за задачата

$$\begin{aligned} &f(x) \to \min, \\ &g_i(x) \leq 0, \qquad i = 1, \ldots, m, \\ &\langle a_j, x\rangle \leq b_j, \qquad j = 1, \ldots, l, \end{aligned} \tag{8}$$

можем да формулираме следната

Теорема 2. Нека x^0 е решение на задачата (8), за която е изпълнено условието на Слейтър, т. е. съществува $\overline{x} \in \mathbb{R}^n$, удовлетворяващ неравенствата $g_i(\overline{x}) < 0$, $i = 1, 2, \ldots, m$, и $\langle a_j, \overline{x}\rangle \leq b_j$, $j = 1, 2, \ldots, l$. Тогава съществува такъв неотрицателен вектор $(\lambda^0, \mu^0) \in \mathbb{R}^{m+l}_+$, че (x^0, λ^0, μ^0) е седлова точка (спрямо x и (λ, μ)) на функцията на Лагранж

$$\Lambda(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{l} \mu_j(\langle a_j, x\rangle - b_j)$$

в областта $\mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^l_+$.

Доказателство. Нека λ^0 е множител на Лагранж за задачата (1), като

$$X = \{x \in \mathbb{R}^n : \langle a_j, x\rangle \leq b_j,\ \ j = 1, \ldots, l\}$$

и μ^0 е множител на Лагранж за задачата (6). Тогава от (5) и (7) имаме

$$\Lambda(x^0, \lambda^0, \mu^0) = f(x^0),$$

следователно

$$\Lambda(x^0, \lambda, \mu) \leq \Lambda(x^0, \lambda^0, \mu^0)$$

за всяко $\lambda \in \mathbb{R}^m_+$ и за всяко $\mu \in \mathbb{R}^l_+$.

От друга страна, за всяко $x \in \mathbb{R}^n$

$$\begin{aligned}\Lambda(x^0,\lambda^0,\mu^0) &= \mathcal{L}(x^0,\mu^0) \le \mathcal{L}(x,\mu^0)\\ &= L(x,\lambda^0) + \sum_{j=1}^{l}\mu_j^0(\langle a_j,x\rangle - b_j)\\ &= f(x) + \sum_{i=1}^{m}\lambda_i^0 g_i(x) + \sum_{j=1}^{l}\mu_j^0(\langle a_j,x\rangle - b_j)\\ &= \Lambda(x,\lambda^0,\mu^0)\end{aligned}$$

т. е. (x^0,λ^0,μ^0) е седлова точка. □

3. Ако множеството X е зададено със система неравенства и равенства

$$X = \{x \in \mathbb{R}^n : \langle a_j,x\rangle \le b_j,\ j=1,\dots,l,\ \langle c_k,x\rangle = d_k,\ k=1,\dots,p\},$$

то X може да се запише и със система неравенства:

$$\begin{aligned}X = \{x \in \mathbb{R}^n : \langle a_j,x\rangle &\le b_j, & j&=1,\dots,l,\\ \langle c_k,x\rangle &\le d_k, & k&=1,\dots,p,\\ \langle -c_k,x\rangle &\le -d_k, & k&=1,\dots,p\}.\end{aligned}$$

Функцията на Лагранж ще има вида

$$\begin{aligned}\Lambda(x,\lambda,\mu,\xi^1,\xi^2) &= f(x) + \sum_{i=1}^{m}\lambda_i g_i(x) + \sum_{j=1}^{l}\mu_j(\langle a_j,x\rangle - b_j)\\ &\quad + \sum_{k=1}^{p}\xi_k^2(\langle -c_k,x\rangle + d_k) + \sum_{k=1}^{p}\xi_k^1(\langle c_k,x\rangle - d_k)\\ &= f(x) + \sum_{i=1}^{m}\lambda_i g_i(x) + \sum_{j=1}^{l}\mu_j(\langle a_j,x\rangle - b_j)\\ &\quad + \sum_{k=1}^{p}(\xi_k^1 - \xi_k^2)(\langle c_k,x\rangle - d_k).\end{aligned}$$

Означаваме $\xi = \xi^1 - \xi^2$ и тогава записваме функцията на Лагранж така:

$$\Lambda(x,\lambda,\mu,\xi) = f(x) + \sum_{i=1}^{m}\lambda_i g_i(x) + \sum_{j=1}^{l}\mu_j(\langle a_j,x\rangle - b_j) + \sum_{k=1}^{p}\xi_k(\langle c_k,x\rangle - d_k). \tag{9}$$

Ако положим $\xi = \xi^1 - \xi^2$, то за ξ няма ограничение за знак.

И така за задачата

$$
\begin{aligned}
&f(x) \to \min, && \\
&g_i(x) \le 0, && i = 1, \dots, m, \\
&\langle a_j, x\rangle \le b_j, && j = 1, \dots, l, \\
&\langle c_k, x\rangle = d_k, && k = 1, \dots, p,
\end{aligned}
\tag{10}
$$

като предположим, че f и g_i са изпъкнали и условието на Слейтър е изпълнено (т. е. съществува $\overline{x} \in \mathbb{R}^n$, удовлетворяващо неравенството $g_i(\overline{x}) < 0$, $i = 1, \dots, m$, $\langle a_j, \overline{x}\rangle \le b_j$, $j = 1, \dots, l$, $\langle c_k, \overline{x}\rangle = d_k$, $k = 1, \dots, p$), в сила е следното твърдение:

Следствие 1. Нека x^0 е решение на (10). Тогава съществуват такива вектори $\lambda^0 \in \mathbb{R}^m_+$, $\mu^0 \in \mathbb{R}^l_+$ и $\xi^0 \in \mathbb{R}^p$, че $(x^0, \lambda^0, \mu^0, \xi^0)$ е седлова точка (спрямо x и (λ, μ, ξ)) на функцията на Лагранж (9) в областта $\mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^l_+ \times \mathbb{R}^p$.

4. Видяхме, че за една задача от типа (10) можем да формулираме различни функции на Лагранж и в съответствие с това — различни варианти на теоремата на Кун—Такър.

Да разгледаме отново задачата (10) при условията на следствие 1. Нека I, J и K са подмножества (евентуално празни) съответно на множествата от индекси $\{1, \dots, m\}$, $\{1, \dots, l\}$ и $\{1, \dots, k\}$. Означаваме с $\overline{I}$, $\overline{J}$ и $\overline{K}$ допълненията на I, J и K. Въвеждаме множеството

$$X_1 = \{x \in \mathbb{R}^n : g_i(x) \le 0,\ i \in I,\ \langle a_j, x\rangle \le b_j,\ j \in J,\ \langle c_k, x\rangle = d_k,\ k \in K\}$$

и функция на Лагранж във вида

$$\Lambda_1(x, \overline{\lambda}, \overline{\mu}, \overline{\xi}) = f(x) + \sum_{i \in I} \overline{\lambda}_i g_i(x) + \sum_{j \in J} \overline{\mu}_j (\langle a_j, x\rangle - b_j) + \sum_{k \in K} \overline{\xi}_k (\langle c_k, x\rangle - d_k). \tag{11}$$

Нека x^0 е решение на задачата (10) и $(\lambda^0, \mu^0, \xi^0)$ е съответният вектор на Лагранж. Тогава за всяко $x \in \mathbb{R}^n$ е изпълнено

$$
\begin{aligned}
f(x^0) \le \Lambda_1(x, \overline{\lambda}^0, \overline{\mu}^0, \overline{\xi}^0) &+ \sum_{i \in \overline{I}} \lambda_i^0 g_i(x) \\
&+ \sum_{j \in \overline{J}} \mu_j^0 (\langle a_j, x\rangle - b_j) + \sum_{k \in \overline{K}} \xi_k^0 (\langle c_k, x\rangle - d_k),
\end{aligned}
$$

където сме означили

$$\overline{\lambda}^0 = (\lambda_i^0)_{i \in I}, \quad \overline{\mu}^0 = (\mu_j^0)_{j \in J}, \quad \overline{\xi}^0 = (\xi_k^0)_{k \in K}.$$

Да означим с X_2 множеството

$$X_2 = \{x \in \mathbb{R}^n : g_i(x) \le 0, \quad i \in \overline{I}, \\ \langle a_j, x\rangle \le b_j, \quad j \in \overline{J}, \\ \langle c_k, x\rangle = d_k, \quad k \in \overline{K}\}.$$

Тогава за всяко $x \in X_2$ имаме

$$\sum_{i\in\overline{I}} \lambda_i^0 g_i(x) \le 0,$$
$$\sum_{j\in\overline{J}} \mu_j^0(\langle a_j, x\rangle - b_j) \le 0,$$
$$\sum_{k\in\overline{K}} \xi_k^0(\langle c_k, x\rangle - d_k) = 0.$$

Следователно

$$f(x^0) = \Lambda_1(x^0, \overline{\lambda}^0, \overline{\mu}^0, \overline{\xi}^0) \le \Lambda_1(x, \overline{\lambda}^0, \overline{\mu}^0, \overline{\xi}^0) \text{ за всяко } x \in X_2.$$

Очевидно за всяко $\overline{\lambda}$, $\overline{\mu}$, $\overline{\xi}$ със съответния знак и размерност е изпълнено

$$\Lambda_1(x^0, \overline{\lambda}, \overline{\mu}, \overline{\xi}) \le \Lambda_1(x, \overline{\lambda}^0, \overline{\mu}^0, \overline{\xi}^0).$$

Така получаваме

Следствие 2. Нека x^0 е решение на (10) и I, J и K са подмножества съответно на множествата $\{1, \dots, m\}$, $\{1, \dots, l\}$, $\{1, \dots, k\}$. Тогава съществуват такива вектори $\overline{\lambda}^0 \in \mathbb{R}_+^{|I|}$, $\overline{\mu}^0 \in \mathbb{R}_+^{|J|}$ и $\overline{\xi}^0 \in \mathbb{R}^{|K|}$ (с $|A|$ означаваме броя на елементите на множеството A), че $(x^0, \overline{\lambda}^0, \overline{\mu}^0, \overline{\xi}^0)$ е седлова точка на функцията на Лагранж (11) в областта $X_2 \times \mathbb{R}_+^{|I|} \times \mathbb{R}_+^{|J|} \times \mathbb{R}^{|K|}$, където $X_2 = \{x \in \mathbb{R}^n : g_i(x) \le 0, i \in \overline{I}, \langle a_j, x\rangle \le b_j, j \in \overline{J}, \langle c_k, x\rangle = d_k, k \in \overline{K}\}$.

Пример 2. Разглеждаме задачата $f(x) \to \min$, $\langle a_j, x\rangle \le b_j$, $j = 1, 2, \dots, l$, $x_j \ge 0$, $j = 1, 2, \dots, n$, където f е изпъкнала в $\mathbb{R}^n$ функция. Въвеждаме функция на Лагранж

$$L(x, \lambda) = f(x) + \sum_{j=1}^{l} \lambda_j(\langle a_j, x\rangle - b_j).$$

Съгласно следствие 2 съществува такъв вектор $\lambda^0 = (\lambda_1^0, \dots, \lambda_l^0) \in \mathbb{R}_+^l$, че за всяко $x \in \mathbb{R}_+^n$

$$f(x^0) \le f(x) + \sum_{j=1}^{l} \lambda_j^0(\langle a_j, x\rangle - b_j).$$

5. Ще докажем още веднъж, но по друг начин теоремата на Кун и Такър за общата задача на изпъкналото оптимиране.

Разглеждаме задачата

$$f(x) \to \min, \quad x \in X, \tag{12}$$

където

$$X = \{x \in \mathbb{R}^n : g_i(x) \le 0\ (I_1),\ l_i(x) \le 0\ (I_2),\ l_i(x) = 0\ (I_3),\ x_j \ge 0\ (J_1)\}.$$

Тук I_1, I_2, I_3, J_1 са множества от индекси, като

$$(I_1 \cap I_2) \cup (I_1 \cap I_3) \cup (I_2 \cap I_3) = \emptyset, \quad I_1 \cup I_2 \cup I_3 = I = \{1, \dots, m\},$$

а $J_1 = \{1, 2, \dots, n_1\} \subset J = \{1, 2, \dots, n\}$. С $g_i(x) \le 0\ (I_1)$ означаваме системата неравенства $g_i(x) \le 0$, $i \in I_1$. За вектора $a \in \mathbb{R}^s$, чиито компоненти с индекси от $K \subset \{1, \dots, s\}$ са неотрицателни, ще пишем $a \ge 0\ (K)$.

Ще предполагаме, че са изпълнени следните условия:

1) Функциите f, g_i, $i \in I_1$, са изпъкнали, а

$$l_i(x) = \sum_J a_{ij} x_j - b_i, \quad i \in I_2 \cup I_3,$$

т. е. l_i са афинни.

2) Изпълнено е условието на Слейтър — съществува такава точка $\overline{x} \in X$, че $g_i(\overline{x}) < 0\ (I_1)$.

Теорема 3 (Кун–Такър)**.** Ако са изпълнени условията 1) и 2) и x^0 е решение на задачата (12), то съществува такъв вектор $\lambda^0 \ge 0\ (I_1 \cup I_2)$, че точката (x^0, λ^0) е седлова точка за функцията на Лагранж

$$L(x, \lambda) = f(x) + \sum_{I_1} \lambda_i g_i(x) + \sum_{I_2} \lambda_i l_i(x) + \sum_{I_3} \lambda_i l_i(x)$$

в областта $x \ge 0\ (J_1)$, $\lambda \ge 0\ (I_1 \cup I_2)$, т. е.

$$L(x^0, \lambda) \le L(x^0, \lambda^0) \le L(x, \lambda^0) \tag{13}$$

за всяко $x \ge 0\ (J_1)$ и $\lambda \ge 0\ (I_1 \cup I_2)$.

Доказателство. Нека x^0 е решение на задачата (12), а

$$\sum_{I_2'} \alpha_i l_i(x) + \sum_{I_3} \beta_i l_i(x) + \sum_{J_1'} \gamma_{m+j} x_j = 0, \tag{14}$$

където $\alpha_i > 0$ (I_2'), $\gamma_{m+j} < 0$ (J_1'), $I_2' \subset I_2$, $J_1' \subset J_1$, е максималната (по брой на елементите на $I_2' \cup J_1'$) линейна комбинация от този вид на афинните функции, тъждествено равна на нула. В $\mathbb{R}^{m+n_1+1}$ дефинираме множеството

$$\begin{aligned} Y = \{y : \text{съществува } x \in \mathbb{R}^n,\ y_0 \ge f(x),\ y_i \ge g_i(x)\ (I_1),& \\ y_i \ge l_i(x)\ (I_2 \setminus I_2'),\ y_i = l_i(x)\ (I_2' \cup I_3),& \\ y_{m+j} = x_j\ (J_1'),\ y_{m+j} \le x_j\ (J_1 \setminus J_1')\}.& \end{aligned} \tag{15}$$

Множеството $Y \neq \emptyset$ и при условие 1) е изпъкнало. Действително, ако $y^1 \in Y$ и $y^2 \in Y$, то съществуват такива точки $x^1 \in \mathbb{R}^n$ и $x^2 \in \mathbb{R}^n$, че

$$\begin{aligned} &y_0^1 \ge f(x^1), && y_0^2 \ge f(x^2), && \\ &y_i^1 \ge g_i(x^1), && y_i^2 \ge g_i(x^2), && i \in I_1, \\ &y_i^1 \ge l_i(x^1), && y_i^2 \ge l_i(x^2), && i \in I_2 \setminus I_2', \\ &y_i^1 \ge l_i(x^1), && y_i^2 = l_i(x^2), && i \in I_2' \cup I_3, \\ &y_{m+j}^1 = x_j^1, && y_{m+j}^2 = x_j^2, && j \in J_1', \\ &y_{m+j}^1 \le x_j^1, && y_{m+j}^2 \le x_j^2, && j \in J_1 \setminus J_1'. \end{aligned}$$

Но тогава за точката

$$\alpha y^1 + \beta y^2, \quad \alpha \ge 0, \quad \beta \ge 0, \quad \alpha + \beta = 1,$$

получаваме

$$\begin{aligned} &\alpha y_0^1 + \beta y_0^2 \ge \alpha f(x^1) + \beta f(x^2) \ge f(\alpha x^1 + \beta x^2), && \\ &\alpha y_i^1 + \beta y_i^2 \ge \alpha g_i(x^1) + \beta g_i(x^2) \ge g_i(\alpha x^1 + \beta x^2), && i \in I_1, \\ &\alpha y_i^1 + \beta y_i^2 \ge \alpha l_i(x^1) + \beta l_i(x^2) = l_i(\alpha x^1 + \beta x^2), && i \in I_2 \setminus I_2', \\ &\alpha y_i^1 + \beta y_i^2 = \alpha l_i(x^1) + \beta l_i(x^2) = l_i(\alpha x^1 + \beta x^2), && i \in I_2' \cup I_3, \\ &\alpha y_{m+j}^1 + \beta y_{m+j}^2 = \alpha x_j^1 + \beta x_j^2 = (\alpha x^1 + \beta x^2)_j, && j \in J_1', \\ &\alpha y_{m+j}^1 + \beta y_{m+j}^2 \le \alpha x_j^1 + \beta x_j^2 = (\alpha x^1 + \beta x^2)_j, && j \in J_1 \setminus J_1', \end{aligned}$$

следователно $\alpha y^1 + \beta y^2 \in Y$. Очевидно точката $z = (f(x^0), 0, \dots, 0) \in Y$ и за всяко $\varepsilon > 0$ точката $z_{+\varepsilon} = (f(x^0) + \varepsilon, 0, \dots, 0) \in Y$, а $z_{-\varepsilon} = (f(x^0) - \varepsilon, 0, \dots, 0) \notin$

Y. Следователно точката z е истинска контурна точка за Y ($z \in \mathrm{r}\partial Y$). Но тогава през z може да се прекара опорна хиперравнина за множеството Y, която не съдържа изцяло Y, т. е. съществуват вектор $v \in \mathbb{R}^{m+n_1+1}$ и $\widetilde{y} \in Y$ такива, че

$$(16) \qquad \langle v, y \rangle \geq \langle v, z \rangle \quad \text{за всяко } y \in Y$$

и

$$\langle v, \widetilde{y} \rangle > \langle v, z \rangle.$$

Тъй като множеството на компонентите на точките от Y с индекси от $\{0\} \cup I_1 \cup \{I_2 \setminus I_2'\}$ е неограничено отгоре, а с индекси от $J_1 \setminus J_1'$ — отдолу, то за да бъде изпълнено (16) за всяко $y \in Y$, трябва

$$v_i \geq 0, \; i \in \{0\} \cup I_1 \cup \{I_2 \setminus I_2'\} \quad \text{и} \quad v_{m+j} \leq 0, \; i \in J_1 \setminus J_1'.$$

Понеже за всяко $x \in \mathbb{R}^n$ точката

$$(f(x), g_j(x) \; (I_1), l_i(x) \; (I_2 \cup I_3), x_{m+j} \; (J_1)) \in Y,$$

то от (16) следва, че за всяко $x \in \mathbb{R}^n$ е изпълнено неравенството

$$(17) \qquad v_0 f(x) + \sum_{I_1} v_i g_i(x) + \sum_{I_2} v_i l_i(x) + \sum_{I_3} v_i l_i(x) + \sum_{J_1} v_{m+j} x_j \geq v_0 f(x).$$

Ще покажем, че $v_0 \neq 0$. Означаваме

$$\varphi(x) = v_0 f(x) + \sum_{I_1} v_i g_i(x) + \sum_{I_2} v_i l_i(x) + \sum_{I_3} v_i l_i(x) + \sum_{J_1} v_{m+j} x_j.$$

Ако допуснем, че $v_0 = 0$, от (17) получаваме $\varphi(x) \geq 0$ за всяко $x \in \mathbb{R}^n$. Оттук и от условие 2) следва, че $v_i = 0$, $i \in I_1$. Тогава

$$\varphi(x) = \sum_{I_2} v_i l_i(x) + \sum_{I_3} v_i l_i(x) + \sum_{J_1} v_{m+j} x_j \geq 0$$

за всяко $x \in \mathbb{R}^n$. Но $\varphi(x)$ е афинна функция, следователно е константа и понеже $\varphi(x) \leq 0$ за всяко $x \in X$, то

$$(18) \qquad \varphi(x) \equiv 0.$$

Като вземем предвид (14), $\varphi(x)$ ще има вида

$$\varphi(x) = \sum_{I_2'} v_i l_i(x) + \sum_{I_3} v_i l_i(x) + \sum_{J_1'} v_{m+j} x_j,$$

т. е. $v_i = 0,\ i \in I_2 \setminus I_2',\ v_{m+j} = 0,\ j \in J_2 \setminus J_2'$.

Ако точката $\widetilde{x} \in \mathbb{R}^n$ е такава, че

$$\widetilde{y}_0 \geq f(\widetilde{x}),\ \widetilde{y}_i \geq g_i(\widetilde{x})\ (I_1),\ \widetilde{y}_i \geq l_i(\widetilde{x})\ (I_2 \setminus I_2'),$$
$$\widetilde{y}_i = l_i(\widetilde{x})\ (I_2' \cup I_3),\ \widetilde{y}_{m+j} = \widetilde{x}_{m+j}\ (J'_1),\ \widetilde{y}_{m+j} \geq \widetilde{x}_{m+j}\ (J_1 \setminus J_1'),$$

то

$$\begin{aligned}\varphi(\widetilde{x}) &= \sum_{I_2'} v_i l_i(\widetilde{x}) + \sum_{I_3} v_i l_i(\widetilde{x}) + \sum_{J'_1} v_{m+j}\widetilde{x}_j \\ &= \sum_{I_2'} v_i \widetilde{y}_i + \sum_{I_3} v_i \widetilde{y}_i + \sum_{J'_1} v_{m+j}\widetilde{y}_j = \langle v, \widetilde{y}\rangle > 0.\end{aligned}$$

Стигаме до противоречие с (18). Следователно $v_0 \neq 0$. Можем да смятаме, че в (17) $v_i \geq 0,\ i \in I_2'$, а $v_{m+j} \leq 0,\ j \in J_1'$, тъй като това винаги можем да си осигурим, като използваме тъждеството (14).

Понеже при $x_j \geq 0\ (J_1)$ имаме $\sum\limits_{J_1} v_{m+j}x_j \leq 0$, от (17) получаваме

(19) $$v_0 f(x) + \sum_{I_1} v_i g_i(x) + \sum_{I_2} v_i l_i(x) + \sum_{I_3} v_i l_i(x) \geq v_0 f(x^0)$$

за всяко $x \geq 0\ (J_1)$. Разделяме (19) на v_0 и като означим

$$\lambda_i^0 \ = \ \frac{v_i}{v_0},\ i \in I \quad (\lambda_i^0 \geq 0,\ i \in I_1 \cup I_2),$$

то

(20) $$f(x) + \sum_{I_1} \lambda_i^0 g_i(x) + \sum_{I_2} \lambda_i^0 l_i(x) + \sum_{I_3} \lambda_i^0 l_i(x) \geq f(x^0)$$

за всяко $x \geq 0\ (J_1)$.

Ако в (20) положим $x = x^0$, ще получим

$$\sum_{I_1} \lambda_i^0 g_i(x^0) + \sum_{I_2 \cup I_3} \lambda_i^0 l_i(x^0) \geq 0.$$

Но тези суми са неположителни. Следователно

(21) $$\sum_{I_1} \lambda_i^0 g_i(x^0) + \sum_{I_2 \cup I_3} \lambda_i^0 l_i(x^0) = 0.$$

Като прибавим лявата стана на (21) към дясната на (19), получаваме

(22) $$L(x^0, \lambda^0) \leq L(x, \lambda^0).$$

От неравенството

$$\sum_{I_1} \lambda_i g_i(x^0) + \sum_{I_2} \lambda_i l_i(x^0) + \sum_{I_3} \lambda_i l_i(x^0) \leq 0$$

за всяко $\lambda_i \geq 0$ $(I_1 \cup I_2)$ следва

$$f(x^0) + \sum_{I_1} \lambda_i g_i(x^0) + \sum_{I_2} \lambda_i l_i(x^0) + \sum_{I_3} \lambda_i l_i(x^0)$$
$$\leq f(x^0) + \sum_{I_1} \lambda_i^0 g_i(x^0) + \sum_{I_2} \lambda_i^0 l_i(x^0) + \sum_{I_3} \lambda_i^0 l_i(x^0)$$

или

$$L(x^0, \lambda) \leq L(x^0, \lambda^0). \tag{23}$$

От (22) и (23) получаваме

$$L(x^0, \lambda) \leq L(x^0, \lambda^0) \leq L(x, \lambda^0)$$

за всяко $x \geq 0$ (J_1) и $\lambda \geq 0$ $(I_1 \cup I_2)$. С това теоремата е доказана. □

4. Диференциални условия за седлова точка

Разглеждаме задачата за изпъкналото оптимиране

$$\begin{aligned} & f(x) \to \min, & & \\ & g_i(x) \leq 0, & & i = 1, \dots, m, \\ & \langle c_j, x \rangle = d_j, & & j = 1, \dots, p, \\ & x_j \geq 0, & & j \in J, \end{aligned} \tag{1}$$

където J е подмножество на $\{1, \dots, n\}$ и $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Предполагаме, че е изпълнено условието:

(2) Функциите f и g_i, $i = 1, \dots, m$, са изпъкнали в $\mathbb{R}^n$ и съществува такава допустима точка $\overline{x}$ (т. е. точка, удовлетворяваща всички ограничения), че $g_i(\overline{x}) < 0$ за всички i, които не са афинни.

Въвеждаме функция на Лагранж:

$$L(x, \lambda, \xi) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \xi_j(\langle c_j, x \rangle - d_j). \tag{3}$$

От теоремата на Кун—Такър следва, че съществуват такива вектори $\lambda^0 \in \mathbb{R}^m_+$ и $\xi^0 \in \mathbb{R}^p$, че (x^0, λ^0, ξ^0) е седлова точка на (3) в областта $\widetilde{R} \times \mathbb{R}^m_+ \times \mathbb{R}^p$, където

$$\widetilde{R} = \{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n, x_j \geq 0, j \in J\},$$

т. е.

$$L(x^0, \lambda, \xi) \leq L(x^0, \lambda^0, \xi^0) \leq L(x, \lambda^0, \xi^0) \tag{4}$$

за всяко $x \in \widetilde{R}$, $\lambda^0 \in \mathbb{R}^m_+$, $\xi^0 \in \mathbb{R}^p$. При допълнително предположение за диференцуемост на f и g_i, $i = 1, 2, \dots, m$, по-долу ще дадем условие, еквивалентно на (4)

Теорема 1. Нека f и g_i, $i = 1, 2, \dots, m$, удовлетворяват условията (2) и са диференцуеми в точката x^0. За да бъде x^0 решение на задачата (1), необходимо и достатъчно е да съществуват такива вектори $\lambda^0 \in \mathbb{R}^m_+$ и $\xi^0 \in \mathbb{R}^p$, че да са изпълнени следните съотношения:

$$\frac{\partial L}{\partial x_j} \geq 0, \qquad j \in J, \tag{5а}$$

$$\frac{\partial L}{\partial x_j} = 0, \qquad j \in \{1, \dots, n\} \setminus J, \tag{5б}$$

$$x_j \frac{\partial L}{\partial x_j} = 0, \qquad j \in \{1, \dots, n\}, \tag{5в}$$

$$x_j \geq 0, \qquad j \in J, \tag{5г}$$

$$\frac{\partial L}{\partial \lambda_j} \leq 0, \qquad j \in \{1, \dots, m\}, \tag{6а}$$

$$\lambda_j \frac{\partial L}{\partial \lambda_j} = 0, \qquad j \in \{1, \dots, m\}, \tag{6б}$$

$$\lambda_j \geq 0, \qquad j \in \{1, \dots, m\}, \tag{6в}$$

$$\frac{\partial L}{\partial \xi_j} = 0, \qquad j \in \{1, \dots, p\}. \tag{7}$$

Доказателство. *Необходимост.* Нека x^0 е решение на (1). Тогава съществуват $\lambda^0 \in \mathbb{R}^m_+$ и $\xi^0 \in \mathbb{R}^p$, удовлетворяващи условието (2). При фиксирани $\lambda = \lambda^0$ и $\xi = \xi^0$ функцията $L(x, \lambda^0, \xi^0)$ достига минимума си в областта $\widetilde{R}$ в точката x^0. Следователно са изпълнени условията (5а)–(5г). От (4) се вижда също, че функцията $L(x^0, \lambda, \xi)$ достига максимума си в областта $\mathbb{R}^m_+ \times \mathbb{R}^p$ при $\lambda = \lambda^0$ и $\xi = \xi^0$. Следователно са изпълнени (6а)–(6в) и (7).

Достатъчност. Нека са изпълнени съотношенията (5а)–(7). Понеже функцията $L(x, \lambda^0, \xi^0)$ е изпъкнала като функция на x, то

$$
\begin{aligned}
L(x, \lambda^0, \xi^0) &\geq L(x^0, \lambda^0, \xi^0) + \left\langle \frac{\partial}{\partial x} L(x^0, \lambda^0, \xi^0), x - x^0 \right\rangle \\
&= L(x^0, \lambda^0, \xi^0) + \left\langle \frac{\partial}{\partial x} L(x^0, \lambda^0, \xi^0), x \right\rangle - \left\langle \frac{\partial}{\partial x} L(x^0, \lambda^0, \xi^0), x^0 \right\rangle \\
&\geq L(x^0, \lambda^0, \xi^0)
\end{aligned} \tag{8}
$$

за всяко $x \in \widetilde{R}$, т. е. x^0 е точка на минимум на $L(x, \lambda^0, \xi^0)$ в $\widetilde{R}$.

Функцията $L(x^0, \lambda, \xi)$ е линейна функция на λ и ξ в $\mathbb{R}^m_+ \times \mathbb{R}^p$. Тогава

$$
\begin{aligned}
L(x^0, \lambda, \xi) = L(x^0, \lambda^0, \xi^0) &+ \left\langle \frac{\partial}{\partial \lambda} L(x^0, \lambda^0, \xi^0), \lambda - \lambda^0 \right\rangle \\
&+ \left\langle \frac{\partial}{\partial \xi} L(x^0, \lambda^0, \xi^0), \xi - \xi^0 \right\rangle,
\end{aligned}
$$

и отчитайки (6а)–(6в) и (7), получаваме

$$
L(x^0, \lambda, \xi) \leq L(x^0, \lambda^0, \xi^0) \tag{9}
$$

за всяко $(\lambda, \xi) \in \mathbb{R}^m_+ \times \mathbb{R}^p$. Неравенствата (8) и (9) означават, че (x^0, λ^0, ξ^0) е седлова точка, следователно x^0 е решение на задачата. □

Ако условията за неотрицателност на някои компоненти на x са (евентуално) включени като ограничения от общ вид, т. е. дадена е задачата

$$
\begin{aligned}
&f(x) \to \min, \\
&g_i(x) \leq 0, \qquad i = 1, \dots, m, \\
&\langle c_j, x \rangle = d_j, \quad j = 1, \dots, p,
\end{aligned} \tag{10}
$$

тогава условията (5а)–(7) ще имат вида

$$
\frac{\partial L}{\partial x_i} = 0, \qquad i \in \{1, \dots, n\}, \tag{11}
$$

$$
\frac{\partial L}{\partial \lambda_j} \leq 0, \qquad j \in \{1, \dots, m\}, \tag{12а}
$$

$$
\lambda_j \frac{\partial L}{\partial \lambda_j} = 0, \quad j \in \{1, \dots, m\}, \tag{12б}
$$

$$
\lambda_j \geq 0, \qquad j \in \{1, \dots, m\} \tag{12в}
$$

$$
\frac{\partial L}{\partial \xi_j} = 0, \qquad j \in \{1, \dots, p\}. \tag{13}
$$

Условията (12a) и (13) означават, че x^0 е допустима точка. Останалите условия могат да бъдат записани по следния начин:

$$
\begin{aligned}
&f'(x^0) + \sum_{i=1}^{m} \lambda_i^0 g_i'(x^0) + \sum_{i=1}^{p} \xi_i^0 c_i = 0, \\
&\lambda_j^0 \geq 0, \quad j = 1, \ldots, m, \\
&\langle \lambda^0, g(x^0) \rangle = 0.
\end{aligned}
$$

И така доказахме още веднъж теоремата на Джон, като използвахме условието на Слейтър. Нещо повече, поради изпъкналостта полученият резултат дава не само необходимо, но и достатъчно условие за (глобален) минимум.

Пример 1.

$$
\begin{aligned}
\frac{1}{2}(x_1 - 3)^2 + \frac{1}{2}(x_2 - 2)^2 &\to \min, \\
x_1^2 + x_2^2 &\leq 5, \\
x_1 + 2x_2 &= 4, \\
x_1 \geq 0,\ x_2 &\geq 0.
\end{aligned}
$$

Решение. Функциите, участващи в задачата, са изпъкнали и диференцуеми, следователно можем да използваме условията (5a)–(7). Записваме функцията на Лагранж

$$
L(x, \lambda, \xi) = \frac{1}{2}(x_1 - 3)^2 + \frac{1}{2}(x_2 - 2)^2 + \lambda(x_1^2 + x_2^2 - 5) + \xi(x_1 + 2x_2 - 4).
$$

Условията (5a)–(7) ще имат вида

$$
\begin{aligned}
&\frac{\partial L}{\partial x_1} = x_1 - 3 + 2\lambda x_1 + \xi \geq 0, \\
&x_1 \frac{\partial L}{\partial x_1} = x_1(x_1 - 3 + 2\lambda x_1 + \xi) = 0, \\
&x_1 \geq 0, \\
&\frac{\partial L}{\partial x_2} = x_2 - 2 + 2\lambda x_2 + 2\xi \geq 0, \\
&x_2 \frac{\partial L}{\partial x_2} = x_2(x_2 - 2 + 2\lambda x_2 + 2\xi) = 0, \\
&x_2 \geq 0, \\
&\frac{\partial L}{\partial \lambda} = x_1^2 + x_2^2 - 5 \leq 0,
\end{aligned}
$$

$$\lambda\frac{\partial L}{\partial\lambda} = \lambda(x_1^2 + x_2^2 - 5) = 0,$$
$$\lambda \geq 0,$$
$$\frac{\partial L}{\partial\xi} = x_1 + 2x_2 - 4 = 0$$

Ако предположим, че $x_1^0 = 0$, то $x_2^0 = 2$ и $\lambda^0 = 0$, откъдето $\xi^0 = 0$. Но от първото неравенство $\xi^0 \geq 3$, а това е противоречие. До противоречие ще достигнем и ако предположим, че $x_2^0 = 0$. Ако $x_1^0 > 0$ и $x_2^0 > 0$, то първото и четвъртото неравенство са изпълнени като равенства, откъдето получаваме $(x_1^0, x_2^0) = (2, 1)$, $\lambda^0 = \frac{1}{6}$, $\xi^0 = \frac{1}{3}$.

Задача 1.

$$\begin{aligned} &x_1^2 + x_2^2 \to \min,\\ &x_1^2 + x_2^2 \leq 5,\\ &x_1 + 2x_2 = 4,\\ &x_1 \geq 0,\ x_2 \geq 0. \end{aligned}$$

Задача 2.

$$\begin{aligned} -x_1 &\to \min,\\ x_2 - (1 - x_1)^3 &\leq 0,\\ x_2 &\geq 0. \end{aligned}$$

Задача 3.

$$\begin{aligned} \left(x_1 - \frac{9}{4}\right)^2 + (x_2 - 2)^2 &\to \min,\\ -x_1^2 + x_2 &\geq 0,\\ x_1 + x_2 &\geq 6,\\ x_1 \geq 0,\ x_2 &\geq 0. \end{aligned}$$

Задача 4. Нека A е положително дефинитна симетрична матрица с размери $n\times n$, B е матрица с размери $m\times n$, $b \in \mathbb{R}^m$ и $c \in \mathbb{R}^n$. Напишете условията за седлова точка за задачата

$$\frac{1}{2}x^T Ax + c^T x \to \min,$$
$$Bx \leq b.$$

Задача 5. Нека $f : \mathbb{R}^n \to \mathbb{R}^1$ и $g_j : \mathbb{R}^n \to \mathbb{R}^1$, $j = 1, 2, \dots, m$, са изпъкнали функции. Нека x^0 е решение на задачата

$$f(x) \to \min, \quad g_j(x) \le 0, \quad j = 1, 2, \dots, m.$$

Нека M е собствено подмножество на $\{1, 2, \dots, m\}$ и $\overline{x}$ е решение на задачата

$$f(x) \to \min, \quad g_j(x) \le 0, \quad j \in M.$$

Да допуснем, че множеството $V = \{i \in \{1, \dots, m\} : g_i(\overline{x}) > 0\}$ не е празно. Докажете, че може да се намери такова $i \in V$, че $g_i(x^0) = 0$.

Задача 6. Нека a_i, b и c_i са положителни числа, $i = 1, 2, \dots, n$. Решете задачата

$$\sum_{i=1}^{n} \frac{c_i}{x_i} \to \min,$$
$$\sum_{i=1}^{n} a_i x_i = b, \quad x_i \ge 0, \quad i = 1, 2, \dots, n.$$

Глава 6

Специални (основни) класове оптимизационни задачи

В тази глава се разглеждат три специални вида оптимизационни задачи — линейни, хиперболични и квадратични. Освен че имат голямо самостоятелно значение за теорията и практиката на оптималните процеси, те са и ефективен спомагателен апарат при решаването на оптимизационни задачи с по-сложна структура.

1. Линейно оптимиране

Задачите за намиране на екстремум (максимум или минимум) на линейна функция при линейни ограничения се наричат задачи на линейно оптимиране. До такива задачи се стига при математическото моделиране на широк кръг икономически и технико-икономически процеси. Самото възникване и развитие на линейното оптимиране е непосредствено свързано с необходимостта от решаване на задачи, възникващи в икономиката.

Първите резултати по линейно оптимиране са получени независимо от Л. В. Канторович и Т. Купманс през 30-те години. За своите научни изследвания в областта на линейното оптимиране и икономическите модели Канторович и Купманс стават през 1975 г. лауреати на Нобелова награда.

През 1949 г. Дж. Данциг предлага т. нар. симплекс-метод за решаване на линейни задачи, който и до днес е един от най-широко използваните методи. Идеите, заложени в симплекс-метода, продължават да се развиват и усъвършенстват.

1.1. Свойства на линейните функции в изпъкнали множества. Ще посочим някои полезни за следващото изложение свойства на линейните функции, дефинирани в $\mathbb{R}^n$, които, както е известно, имат вида

$$l(x) = \sum_{i=1}^{n} c_i x_i = \langle c, x\rangle.$$

Лема 1. Нека $l(x) = \langle c, x\rangle$, K е конус (с връх в 0) и S е подпространство на $\mathbb{R}^n$. Тогава:

а) ако $\sup\limits_K l(x) < +\infty$, то $l(x) \le 0$, $x \in K$;

б) ако $\sup\limits_S l(x) < +\infty$, то $l(x) = 0$, $x \in S$.

Доказателство. а) Допускаме, че съществува $\bar{z} \in K$ такова, че $l(\bar{z}) > 0$. Но $\lambda\bar{z} \in K$ за всяко $\lambda > 0$ и $l(\lambda\bar{z}) = \lambda l(\bar{z}) \to +\infty$, $\lambda \to +\infty$. Стигаме до противоречие със $\sup\limits_K l(x) < +\infty$. Аналогично се доказва б). □

Следствие 1. Ако $X = X_1 + K$, където K е конус, и линейната функция $l(x)$ е ограничена отгоре в X, то

$$\sup_X l(x) = \sup_{X_1} l(x).$$

Действително $X_1 \subset X$ (тъй като K е конус, съдържащ 0) и тогава от линейността на $l(x)$ следва

$$\sup_X l(x) = \sup_{X_1} l(x) + \sup_K l(x) = \sup_{X_1} l(x),$$

тъй като $\sup\limits_K l(x) = 0$.

Забележка. Ако X е компактно, $l(x)$ достига екстремумите си в X.

Лема 2. За произволно множество $X \subset \mathbb{R}^n$ и линейната функция $l(x)$

$$\sup_X l(x) = \sup_{\operatorname{co} X} l(x).$$

Доказателство. Да вземем произволна точка $\bar{x} \in \operatorname{co} X$. Тогава

$$\bar{x} = \sum_{i=1}^{n+1} \alpha_i x_i, \quad \alpha_i \ge 0, \quad x_i \in X, \quad i = 1, \ldots, n+1, \quad \sum_{i=1}^{n+1} \alpha_i = 1.$$

Но

$$l(\bar{x}) = \langle c, \bar{x}\rangle = \left\langle c, \sum_{i=1}^{n+1} \alpha_i x_i \right\rangle = \sum_{i=1}^{n+1} \alpha_i \langle c, x_i\rangle$$
$$\le \sum_{i=1}^{n+1} \alpha_i \sup_X \langle c, x\rangle = \sup_X \langle c, x\rangle = \sup_X l(x),$$

т. е.

$$\sup_{\operatorname{co} X} l(x) \leq \sup_{X} l(x).$$

Но от $X \subset \operatorname{co} X$ имаме

$$\sup_{X} l(x) \leq \sup_{\operatorname{co} X} l(x).$$

От предишните две неравенства следва

$$\sup_{X} l(x) = \sup_{\operatorname{co} X} l(x). \qquad \square$$

Преди да формулираме следващото твърдение, ще припомним, че всяко изпъкнало затворено множество X се представя във вида

$$X = \operatorname{co} \widehat{X}_1 + K(X),$$

където $X_1 = X \cap T$, $K(X)$ е рецесивният конус на X, а T — допълнителното подпространство на рецесивното подпространство $S(X)$ на X. Припомняме, че $\widehat{X}$ означава множество от крайните точки на X.

Теорема 1. Ако X е изпъкнало затворено множество и $l(x) = \langle c, x\rangle$ е ограничена отгоре в X, то задачите

$$\sup\{l(x) : x \in X\}, \tag{1}$$

$$\sup\{l(x) : x \in \widehat{X}_1\} \tag{2}$$

са еквивалентни в следния смисъл:

1) $X \neq \emptyset \iff \widehat{X}_1 \neq \emptyset$;

2) $\sup\limits_{X} l(x) = \sup\limits_{\widehat{X}_1} l(x)$;

3) ако X^0 и $\widehat{X}_1^0$ са множества на решенията съответно на задачи (1) и (2), а $K_0 = \{z = K(X),\ l(z) = 0\}$, то

$$X^0 = \operatorname{co} \widehat{X}_1^0 + K_0.$$

Доказателство. 1) е очевидно изпълнено, а 2) следва непосредствено от следствие 1 и от леми 1 и 2.

3) Нека $x^0 \in X^0$,

$$x^0 = \sum_{i=1}^{s} \alpha_i \widehat{x}_i + z,\ \alpha_i > 0,\ \widehat{x}_i \in \widehat{X}_1,\ i = 1, \dots, s,\ \sum_{i=1}^{s} \alpha_i = 1,\ z \in K(X).$$

От
$$l(x^0) = l\left(\sum_{i=1}^{s} \alpha_i \widehat{x_i} + z\right) = \sum_{i=1}^{s} \alpha_i l(\widehat{x_i}) + l(z)$$
и $l(\widehat{x_i}) \leq l(x^0)$, $l(z) \leq 0$, следва
$$l(\widehat{x_i}) = l(x^0), \quad l(z) = 0,$$
т. е. $\widehat{x_i} \in \widehat{X_1^0}$ и $z \in K_0$. Следователно
$$X^0 \subset \operatorname{co} \widehat{X_1^0} + K_0.$$
Обратното включване е очевидно. Оттук
$$X^0 = \operatorname{co} \widehat{X_1^0} + K_0. \qquad \square$$

1.2. Основна задача на линейното оптимиране. Една от общите постановки на линейна оптимизационна задача е следната: Дадени са линейна функция
$$l(x) = \sum_{i=1}^{n} c_i x_i = \langle c, x \rangle$$
и множество $X \subset \mathbb{R}^n$, определено с линейни условия (ограничения)
$$X = \{x \subset \mathbb{R}^n : Ax \leq b\},$$
където $A = (a_{ij})$ е матрица с размери $m \times n$ и $b = (b_1, \ldots, b_m)$ е m-мерен вектор. В дефинирането на X могат да участвуват и равенства, но това както вече отбелязахме, не променя нещата (в теоретично отношение), тъй като всяко равенство може да се замени с две неравенства.

Основната задача на линейното оптимиране е да се намери такава точка $x^0 \in X$, че $l(x^0) = \max\{l(x) : x \in X\}$.

Разбира се, може да се търси и минимумът на функцията, но това не променя по същество нещата, тъй като, както знаем,
$$\min\{l(x) : x \in X\} = -\max\{-l(x) : x \in X\}.$$

Както и за всяка оптимизационна задача функцията $l(x)$, която максимизираме, ще наричаме *целева функция*. Всяка точка $x \in X$ ще наричаме *план* на задачата, а множеството X — *множество на плановете*. Точката (плана) x^0, за която $l(x^0) = \max\{l(x) : x \in X\}$, ще наричаме *оптимален план* или *решение*.

Крайните точки, т. е. върховете (ако има такива) на X, ще наричаме *опорни планове*.

Всяко неравенство (или равенство), участващо в дефинирането на X, наричаме *условие* или *ограничение*. Числата b_i, $i = 1, 2, \dots, m$, наричаме *десни страни на условията*, а матрицата A — *матрица на условията* (*ограниченията*).

Ще казваме, че линейната задача е *разрешима*, ако съществува точка $x^0 \in X$, за която $l(x^0) = \sup\{l(x) : x \in X\}$. В противен случай ще наричаме задачата *неразрешима*.

По-нататък ще пропускаме прилагането „оптимизационна“, тъй като става въпрос само за такива задачи. Ще разгледаме основните свойства на линейните задачи, като изкажем редица твърдения, без да ги доказваме. По същество тези твърдения са доказани в предишните глави.

Теорема 1. Опорните планове (ако има такива) на линейната задача са краен брой.

Този резултат следва от по-рано доказаното твърдение (вж. следствие 1 от § 5, гл. 3), че едно многостенно множество има краен брой крайни точки (опорни планове).

Теорема 2. Линейната задача има опорни планове тогава и само тогава, когато $X \neq \emptyset$ и $\operatorname{rank} A = n$ (вж. теорема 1 от § 5, гл. 3).

Теорема 3. Планът $\widehat{x}$ е опорен тогава и само тогава, когато $\operatorname{rank} A_{\widehat{x}} = n$ (вж. теорема 2 от § 5, гл. 3).

Теорема 4 (основна теорема)**.** Ако $X \neq \emptyset$ и $l(x) < \infty$ в X, то линейната задача е разрешима и

1. $\max\{l(x) : x \in X\} = \max\{l(x) : x \in \widehat{X}_1\}$;
2. Ако X^0 и $\widehat{X}_1^0$ са множества от оптимални планове съответно в X и $\widehat{X}_1$, то $X^0 = \operatorname{co} \widehat{X}_1^0 + K_0$, където $K_0 = \{z \in K(X) : l(z) = 0\}$.

Нека $\widehat{x}$ е опорен план на оптимизационна задача и l_i, $i = 1, 2, \dots, s$, са всички ръбове на X, излизащи от $\widehat{x}$, а z_i, $i = 1, 2, \dots, s$, са допустимите посоки за съответните ръбове. Отново ще означим $Z(\widehat{x}) = \{z_1, \dots, z_s\}$.

Теорема 5 (критерий за оптималност)**.** Опорният план $\widehat{x}$ е оптимален тогава и само тогава, когато $l(z) = \langle c, z \rangle \leq 0$ за всяко $z \in Z(\widehat{x})$.

Доказателство. Тъй като конусът на възможните посоки е $K(\widehat{x}, X) = \operatorname{con} Z(\widehat{x})$, то $l(u) \leq 0$ за $u \in K(\widehat{x}, X)$. Следователно $\widehat{x}$ е оптимален план (линейната функция е изпъкнала и следователно локалните екстремуми са и глобални). □

Теорема 6. Ако $l(x) = \langle c, x\rangle = 0$ за $x \in S(X)$ двете задачи

$$\max\{l(x) : x \in X\}, \tag{1}$$
$$\max\{l(x) : x \in X_1 = X \cap T\}. \tag{2}$$

са едновременно разрешими или неразрешими.

Доказателството следва непосредствено от представянето на X във вида

$$X = X_1 + K(X_1) + S(X).$$

По-нататък можем да смятаме, че ако $X \neq \emptyset$, то задачата има опорни планове. При $\widehat{X} = \emptyset$ ще изследваме $l(x)$ в $S(X)$. Ако $l(x) = \langle c, x\rangle$ не е тъждествено равна на нула в $S(X)$, задачата е неразрешима ($l(x)$ расте неограничено в X). Ако $l(x) \equiv 0$ в $S(X)$ — заменяме задачата (1) със задачата (2).

Основните свойства на линейната задача (теореми 1, 3 и 5) подсказват следния метод за решаване на линейни задачи: Намираме всички опорни планове на задачата (те са краен брой). От тях избираме точка x^0, за която целевата функция $l(x)$ има най-голяма стойност. Намираме всички ръбове, излизащи от тази точка (те са краен брой). Тогава:

1. Ако по всички ръбове $l(x)$ не нараства, то x^0 е оптимален план.

2. Ако по някой от ръбовете $l(x)$ нараства, то задачата е неразрешима ($l(x)$ расте неограничено в X).

От теоретична гледна точка един такъв метод е напълно приемлив, но на практика е неприложим. Прости сметки показват, че за да се реши една сравнително неголяма задача (например с размери $m = 100$, $n = 200$) на съвременен компютър, е необходимо време от порядъка на години. Това обаче не омаловажава основната идея на метода — търсене на оптимален план само измежду опорните планове. Само че това търсене трябва да се прави последователно и целенасочено. Например така:

1. Намираме един опорен план $\widehat{x}$.

2. Проверяваме дали е изпълнен критерият за оптималност.

3. Ако е изпълнен, то $\widehat{x}$ е оптимален план и задачата е решена. Ако не е изпълнен, намираме ръб, по който целевата функция нараства.

4. Ако ръбът е неограничен, задачата е неразрешима (целевата функция расте неограничено в X). Ако е ограничен, минаваме (по този ръб) в друг (съседен) опорен план и се връщаме към т. 2.

Основните свойства на линейната задача гарантират, че с този метод за краен брой стъпки или ще намерим опорен план, или ще установим, че задачата е неразрешима. Такъв е механизмът и на предложения от Данциг симплекс-метод.

1.3. Двойственост в линейното оптимиране Сега от диференциалните условия на Кун—Такър ще получим важни резултати за теорията и приложенията на линейните оптимизационни задачи.

Ще разглеждаме следната линейна задача:

$$\text{(P)} \qquad \max\left\{p(x) = \sum_{i=1}^{n} c_i x_i : x \in P\right\},$$

където множеството P е зададено със следната система уравнения и неравенства:

$$\text{(1)} \qquad \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i \in I_1 = \{1, \dots, m_1\},$$

$$\text{(2)} \qquad \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i \in I_2 = \{m_1 + 1, \dots, m\},$$

$$\text{(3)} \qquad x_j \ge 0, \qquad j \in J_1 = \{1, 2, \dots, n_1\}.$$

Наред с (P) разглеждаме задачата

$$\text{(Q)} \qquad \min\left\{q(y) = \sum_{i=1}^{m} b_i y_i : y \in Q\right\},$$

където Q се определя от системата

$$\text{(4)} \qquad \sum_{i=1}^{m} a_{ij} y_i \ge c_j, \quad j \in J_1,$$

$$\text{(5)} \qquad \sum_{i=1}^{m} a_{ij} y_i = c_j, \quad j \in J_2 = \{n_1 + 1, \dots, n\},$$

$$\text{(6)} \qquad y_i \ge 0, \qquad y \in I_1.$$

Определение 1. Двойката задачи (P), (Q) наричаме *спрегнати* (*двойнствени*, *дуални*) задачи. Всяка от тях наричаме спрегната (двойнствена, дуална) на другата.

Да означим още J=$\{1,2,\dots,n\}$ и I=$\{1,2,\dots,m\}$.

Теорема 1. За всяко $x \in P$ и всяко $y \in Q$ е изпълнено $p(x) \le q(y)$.

Доказателство. Да вземем произволни $x \in P$ и $y \in Q$. Имаме

$$p(x) = \sum_{j=1}^{n} c_j x_j \le \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} y_i \right) x_j = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) y_i \le \sum_{i=1}^{m} b_i y_i = q(y).$$

Първото неравенство във веригата следва от (4), (5) и (3), а второто — от (1), (2) и (6). □

Следствие 1. Ако за $x^0 \in P$ и $y^0 \in Q$ е изпълнено $p(x^0) = q(y^0)$, то x^0, y^0 са оптимални планове на съответните задачи.

Доказателство. Твърдението непосредствено следва от неравенствата

$$p(x) \le q(y^0) = p(x^0), \quad x \in P,$$
$$q(y) \ge p(x^0) = q(y^0), \quad y \in Q.$$

Както ще видим по-нататък, условието $p(x^0) = q(y^0)$ е не само достатъчно, но и необходимо за оптималност на плановете x^0 и y^0. □

Следствие 2. Ако $P \ne \emptyset$ и $Q \ne \emptyset$, двойката спрегнати задачи са разрешими.

Твърдението следва непосредствено от ограничеността на функциите $p(x)$ (отгоре) и $q(y)$ (отдолу), тъй като $p(x) \le q(y)$ за $x \in P$ и $y \in Q$.

Следствие 3. Ако $p(x)$ е неограничена (отгоре) в P, то $Q = \emptyset$ и ако $q(y)$ е неограничена (отдолу) в Q, то $P = \emptyset$.

За да запишем условието на Кун—Такър за задачата (P), ще я заменим с еквивалентната ѝ:

$$\text{(P}'\text{)} \qquad \min \left\{ -p(x) = -\sum_{j=1}^{n} c_j x_j : x \in P \right\}.$$

Понеже множествата от решения на (P) и (P′) съвпадат, то всяко условие за оптималност за едната задача е условие за оптималност за другата задача.

Функцията на Лагранж за задачата (P′) има вида

$$L_1(x, y) = -\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} y_i \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right),$$

където $y = (y_1, \dots, y_m)^{\mathrm{T}}$ е вектор от множителите на Лагранж. Сега можем да напишем системата от диференциални условия за седлова точка:

$$\text{(7)} \qquad \sum_{i=1}^{m} a_{ij} y_i - c_j \ge 0, \qquad j \in J_1,$$

$$\sum_{i=1}^{m} a_{ij}y_i - c_j = 0, \qquad j \in J_2, \tag{8}$$

$$x_j\left(\sum_{i=1}^{m} a_{ij}y_i - c_j\right) = 0, \quad j \in J, \tag{9}$$

$$x_j \geq 0, \qquad j \in J_1, \tag{10}$$

$$\sum_{j=1}^{n} a_{ij}x_j - b_i \leq 0, \qquad i \in I_1, \tag{11}$$

$$\sum_{j=1}^{n} a_{ij}x_j - b_i = 0, \qquad i \in I_2, \tag{12}$$

$$y_i\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right) = 0, \quad i \in I, \tag{13}$$

$$y_i \geq 0, \qquad i \in I_1. \tag{14}$$

Както вече знаем, тази система (за задачата (P′)) е свързана с оптимизационната задача (P) по следния начин:

1. Ако x^0 е решение на задачата (P), то съществува y^0 такова, че (x^0, y^0) е решение на системата (7)–(14).

2. Ако (x^0, y^0) е решение на системата (7)–(14), то x^0 е решение на задачата (P).

За задачата (Q) функцията на Лагранж е

$$L_2(y, x) = \sum_{i=1}^{m} b_i y_i + \sum_{j=1}^{n} x_j\left(c_j - \sum_{i=1}^{m} a_{ij}y_i\right) = -L_1(x, y)$$

с вектор на Лагранж $x = (x_1, x_2, \ldots, x_n)$. Оттук се получава, че системата на Кун—Такър за задачата (Q) е също системата (7)–(14).

Следователно, ако x^0 и y^0 са решения на задачите (P) и (Q), то (x^0, y^0) е решение на системата (7)–(14) и ако (x^0, y^0) е решение на системата (7)–(14), то x^0 е решение на задачата (P), а y^0 — на задачата (Q), при това от (9) и (13) имаме

$$p(x^0) = \sum_{j=1}^{n} c_j x_j^0 = \sum_{j=1}^{n} x_j^0 \sum_{i=1}^{m} a_{ij}y_i^0 = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}x_j^0 y_i^0 = \sum_{i=1}^{m} b_i y_i^0 = q(y^0).$$

Тези резултати се обобщават в следната

Теорема 2 (I основна теорема за двойнственост)**.** Двойка спрегнати задачи са едновременно разрешими и $x^0 \in P$, $y^0 \in Q$ са съответно решения на задачите (P) и (Q) тогава и само тогава, когато

$$p(x^0) = q(y^0).$$

Двойките условия

$$\text{(а)} \qquad \sum_{j=1}^{n} a_{ij}x_j \le b_i \quad \text{и} \quad y_i \ge 0, \quad i = 1, 2, \ldots, m_1,$$

$$\text{(б)} \qquad x_j \ge 0 \quad \text{и} \quad \sum_{i=1}^{m} a_{ij}y_i \ge c_j, \quad j = 1, 2, \ldots, n_1$$

ще наричаме *двойки спрегнати условия* за всяко фиксирано $i \in I_1$ и $j \in J_1$.

Определение 2. Едно условие ще наричаме *свободно*, ако за поне един оптимален план то е строго неравенството, и *закрепено*, ако за всеки оптимален план е равенство (т. е. то е активно за всеки оптимален план).

Теорема 3 (II основна теорема за двойнственост)**.** От двойка спрегнати условия едното е свободно, а другото закрепено.

Доказателство. Ще докажем най-напред, че двойка спрегнати условия не могат да бъдат едновременно свободни.

Да допуснем, че условията

$$\sum_{j=1}^{n} a_{kj}x_j \le b_k \quad \text{и} \quad y_k \ge 0, \quad k \in I_1,$$

са свободни, т. е. съществуват оптимални планове x^0 и y^0, за които

$$\sum_{j=1}^{n} a_{kj}x_j^0 < b_k \quad \text{и} \quad y_k^0 > 0, \quad k \in I_1.$$

Но те трябва да изпълняват и условието на Кун—Такър (13). Стигаме до противоречие.

Сега ще покажем, че двойка спрегнати условия не са едновременно закрепени. Без да губим от общността, ще разгледаме двойката условия

$$x_1 \ge 0 \quad \text{и} \quad \sum_{j=1}^{n} a_{j1}y_i \ge c_1.$$

Нека условието $x_1 \geq 0$ в задачата (P) е закрепено, т. е. $x_1 = 0$ за всеки оптимален план. Да означим с l^* оптималната стойност на $p(x)$ в областта (P). При достатъчно малки $\mu > 0$ задачата

$$(P_\mu) \qquad \max\{p_\mu(x) = p(x) + \mu x_1 : x \in P\}$$

е разрешима и $\max\{p_\mu(x), x \in P\} = l^*$.

Действително, ако $P' = \{x : x \in X, x_1 \leq 1\}$, задачата $\max\{p_\mu(x), x \in P'\}$ е разрешима. Избираме

$$\mu_0 = \min_{x \in \widehat{P}_1} \frac{l^* - p(x)}{x_1} > 0.$$

Ако $\mu \in (0, \mu_0)$, то $p_\mu(x) \leq l^*$ за всяко $x \in \widehat{P}_1$, а оттук — и за всяко $x \in P'$ и $x \in P$ (условието $x_1 \leq 1$ е свободно).

Двойнствената задача на (P_μ) е задачата

$$(Q_\mu) \qquad \min\{q(y) : y \in Q_\mu\},$$

където областта Q_μ е зададена със системата

$$\begin{aligned}
&\sum_{i=1}^m a_{i1} y_i \geq c_1 + \mu, \\
&\sum_{i=1}^m a_{ij} y_i \geq c_j, \qquad && j \in J_1 \setminus \{1\}, \\
&\sum_{i=1}^m a_{ij} y_i = c_j, \qquad && j \in J_2, \\
&y_i \geq 0, && i \in I_1.
\end{aligned}$$

От теорема 2 следва, че задачата (Q_μ) е разрешима и $\min\{q(y), y \in Q_\mu\} = l^*$. Ако y^0 е решение на задачата (Q_μ), то

$$q(y^0) = l^*$$

и

$$\begin{aligned}
&\sum_{i=1}^m a_{i1} y_i^0 \geq c_1 + \mu, \\
&\sum_{i=1}^m a_{ij} y_i^0 \geq c_j, \qquad && j \in J_1 \setminus \{1\},
\end{aligned}$$

$$\sum_{i=1}^{m} a_{ij} y_i^0 = c_j, \qquad j \in J_2,$$
$$y_i^0 \geq 0, \qquad i \in I_1.$$

т. е. y^0 е решение на задачата (Q) и $\sum\limits_{i=1}^{m} a_{i1} y_i^0 > c_1$. Следователно условието $\sum\limits_{i=1}^{m} a_{i1} y_i^0 \geq c_1$ в задачата (Q) е свободно.

Обратно, ако условието $\sum\limits_{i=1}^{m} a_{i1} y_i^0 \geq c_1$ е закрепено, условието $x_1 \geq 0$ е свободно. Ако допуснем, че $x_1 \geq 0$ е закрепено, от доказаното ще следва, че спрегнатото му условие е свободно. Стигаме до противоречие. □

Задача 1. Напишете двойнствената задача на задачата (т. нар. канонична форма на линейната оптимизационна задача):

$$l(x) = \langle c, x \rangle \to \max,$$
$$Ax = b,$$
$$x \geq 0.$$

Задача 2. Напишете двойнствената задача на задачата (т. нар. стандартна форма на линейната оптимизационна задача):

$$l(x) = \langle c, x \rangle \to \max,$$
$$Ax \leq b,$$
$$x \geq 0.$$

Задача 3 (теорема на Фаркаш)**.** Докажете, че ако A е матрица с размери $m \times n$, $c \in \mathbb{R}^n$, то от двете системи

1. $Ax \leq 0$, $\langle c, x \rangle > 0$, $x \in \mathbb{R}^n$,
2. $A^{\mathrm{T}} y = c$, $y \geq 0$, $y \in \mathbb{R}^m$,

само едната има решение.

Задача 4. Докажете, че ако A е матрица в размери $m \times n$, $c \in \mathbb{R}^n$, то от двете системи

1. $Ax \leq 0$, $\langle c, x \rangle > 0$, $x \geq 0$, $x \in \mathbb{R}^n$,
2. $A^{\mathrm{T}} y \geq c$, $y \geq 0$, $y \in \mathbb{R}^m$,

само едната е разрешима.

Задача 5. Докажете, че ако A е матрица с размери $m \times n$, B — матрица с размери $l \times m$ и $c \in \mathbb{R}^n$, то от двете системи

1. $Ax \leq 0$, $Bx = 0$, $\langle c, x \rangle > 0$, $x \in \mathbb{R}^n$,
2. $A^{\mathrm{T}} y + B^{\mathrm{T}} z = c$, $y \geq 0$, $y \in \mathbb{R}^m$, $z \in \mathbb{R}^l$,

само едната е разрешима.

Задача 6. Докажете, че ако неравенството $\langle c, x \rangle \leq 0$ е следствие от системата

$$\begin{aligned} Ax &\leq 0, \\ Bx &= 0, \end{aligned}$$

то съществуват $y \geq 0$ и z такива, че

$$c = A^{\mathrm{T}} + B^{\mathrm{T}} z.$$

2. Хиперболично оптимиране

В този параграф ще разгледаме следния клас оптимизационни задачи: Да се намери максимумът на функцията

$$H(x) = \frac{p_0 + p_1 x_1 + \cdots + p_n x_n}{q_0 + q_1 x_1 + \cdots + q_n x_n} = \frac{P(x)}{Q(x)} \tag{1}$$

върху областта X, зададена с условията

$$\begin{aligned} & a_{i0} + \sum_{j=1}^{n} a_{ij} x_j = 0, \qquad && i = 1, 2, \ldots, m, \\ & x_j \geq 0, && j = 1, 2, \ldots, n. \end{aligned} \tag{2}$$

Тъй като целевата функция е дробно-линейна (хиперболична), тази задача ще наричаме задача на *хиперболичното оптимиране*. Тя е записана в каноничен вид — ограниченията са от тип равенство и върху всички променливи е наложено условието за неотрицателност. Към този вид се свежда всяка задача на хиперболичното оптимиране (при условия от тип равенства и неравенства и наличие на свободни променливи).

Тук нямаме за цел да изградим цялостно теорията на хиперболичната задача, защото ще излезем от рамките на основната тема. Обаче поради особената ѝ важност за практиката ще покажем как при известни условия тя може да се сведе до *линейна задача*, така че да може да бъде решавана с изложените вече методи.

Ще предполагаме, че $Q(x) \neq 0$ при $x \in X$. При това предложение, тъй като $Q(x)$ е непрекъсната функция, тя ще запазва постоянен знак върху областта X. За определеност да приемем, че $Q(x) > 0$ при $x \in X$.

Ще казваме, че задачата (1), (2) е *разрешима* тогава и само тогава, когато съществува точка $x^0 \in X$, за която

$$H(x^0) = \sup\{H(x) : x \in X\}.$$

Тук сме употребили sup вместо max, тъй като за разлика от линейната функция хиперболичната функция може да не достига супремума си върху областта X даже когато е ограничена.

Теорема 1. Ако хиперболичната задача (1), (2) е решима, тя има опорен оптимален план.

Доказателство. Да означим

$$H^0 = \sup\{H(x) : x \in X\}$$

и да разгледаме линейната задача

$$\max[P(x) - H^0 Q(x)]. \tag{3}$$

Тази задача е решима, защото $X \neq \emptyset$ и линейната функция $P(x) - H^0 Q(x)$ е ограничена върху X:

$$P(x) - H^0 Q(x) = Q(x)\left[\frac{P(x)}{Q(x)} - H^0\right] = Q(x)(H(x) - H^0) \leq 0.$$

При това очевидно

$$\max_{x \in X}(P(x) - H^0 Q(x)) = 0.$$

Следователно съгласно теорията на линейното оптимиране задачата (3) има опорен оптимален план — например $x' = (x'_1, \ldots, x'_n)$. Ще покажем, че x' е оптимален план и за задачата (1), (2). Наистина

$$P(x') - H^0 Q(x') = 0,$$

следователно

$$H(x') = \frac{P(x')}{Q(x')} = H^0. \qquad \square$$

Да разгледаме сега линейната задача: Да се намери максимумът на функцията

$$\beta(y) = \sum_{j=0}^{n} p_j y_j \tag{4}$$

в областта Λ, зададена с ограниченията

$$\begin{aligned} &\sum_{j=0}^{n} a_{ij} y_j = 0, \quad i = 1, 2, \ldots .m, \\ &\sum_{j=0}^{n} a_j y_j = 1, \end{aligned} \tag{5}$$

$$y_j \geq 0, \quad j = 0, 1, \ldots, m. \tag{6}$$

Връзката между задачите (1), (2) и (4)–(6) ни дават следните теореми:

Теорема 2. Ако $x' = (x'_1, \ldots, x'_n)$ е план на хиперболичната задача (1), (2), то $y' = \dfrac{1}{Q(x')}(1, x'_1, \ldots, x'_n)$ (предполагаме $Q(x') > 0$) е план на линейната задача (4)–(6). Ако $\overline{y} = (\overline{y}_0, \overline{y}_1, \ldots, \overline{y}_n)$ е план на задачата (4)–(6) и $\overline{y}_0 > 0$, то $\overline{x} = \dfrac{1}{\overline{y}_0}(\overline{y}_1, \ldots, \overline{y}_n)$ е план на задачата (1), (2). При това за всеки два съответни плана x' и y' имаме

$$y'_0 = \frac{1}{Q(x')} \quad \text{и} \quad H(x') = \beta(y').$$

Доказателство. Нека $x' = (x'_1, \ldots, x'_n)$ е план на задачата (1), (2). Тогава

$$\begin{aligned} &\sum_{j=0}^{n} a_{ij} y'_j = \frac{1}{Q(x')} \left[a_{i0} + \sum_{j=1}^{n} a_{ij} x'_j \right] = 0, \quad i = 1, 2, \ldots, m, \\ &\sum_{j=0}^{n} q_j y'_j = \frac{1}{Q(x')} \left[q_0 + \sum_{j=1}^{n} q_j x'_j \right] = 1, \\ &y'_j \geq 0, \quad j = 0, 1, \ldots, n. \end{aligned}$$

следователно y' е плана на задачата (4)–(6). Обратно, ако $\overline{y} = (\overline{y}_0, \overline{y}_1, \ldots, \overline{y}_n)$ е план на задачата (4)–(6), то

$$a_{i0} + \sum_{j=1}^{n} a_{ij} \overline{x}_j = \frac{1}{\overline{y}_0} \sum_{j=0}^{n} a_{ij} \overline{y}_j = 0, \qquad i = 1, 2, \ldots, m,$$

$$\overline{x}_j \geq 0, \qquad j = 1, 2, \ldots, n.$$

т. е. $\overline{x}$ е план на задачата (1), (2) и

$$Q(\overline{x}) = q_0 + \frac{1}{\overline{y}_0}\sum_{j=1}^{n} q_j \overline{y}_j = \frac{1}{\overline{y}_0}\sum_{j=0}^{n} q_j \overline{y}_j = \frac{1}{\overline{y}_0},$$

$$H(x') = \frac{P(x')}{Q(x')} = \frac{P_0 + \sum_{j=1}^{n} p_j x'_j}{Q(x')} = p_0 \frac{1}{Q(x')} + \sum_{j=1}^{n} p_j \frac{x'_j}{Q(x')}$$

$$= \sum_{j=0}^{n} p_j y'_j = \beta(y'). \qquad \square$$

Теорема 3. Ако $Q(x) > 0$ за всяко $x \in X$, то компонентата y_0 е ограничена отгоре при $y \in \Lambda$. Тъй като $Q(x)$ е линейна функция, ограничена отдолу в X, то съществува точка $x' \in X$, за която

$$Q(x') = \min_{x \in X} Q(x) = \mu, \tag{7}$$

и понеже $Q(x) > 0$ за всяко $x \in X$, то $\mu > 0$.

Доказателство. Да допуснем, че y_0 е неограничена отгоре в Λ. Тогава съществува редица от точки в Λ

$$y^{(1)}, y^{(2)}, \ldots, y^{(k)}, \ldots$$

такива, че $y_0^{(k)} > k$ за всяко k. Но на всяко $y^{(k)} \in \Lambda$ съответства $x^{(k)} \in X$ (теорема 2) и

$$Q(x^{(k)}) = \frac{1}{y_0^{(k)}}.$$

Ако изберем $k > \frac{1}{\mu}$, то

$$Q(x^{(k)}) = \frac{1}{y_0^{(k)}} < \mu,$$

което противоречи на (7). Следователно y_0 е ограничена отгоре в Λ. $\square$

Следствие 1. Ако y_0 расте неограничено в Λ, то съществува точка $x \in X$, за която $Q(x) = 0$.

Теорема 4. Хиперболичната задача (1), (2) ($Q(x) > 0$ при $x \in X$) е разрешима тогава и само тогава, когато е разрешима линейната задача (4)–(6) и за поне един оптимален план $y^0 = (y_0^0, y_1^0, \ldots, y_n^0)$ имаме $y_0^0 > 0$. При това $x^0 = \frac{1}{y_0^0}(y_1^0, \ldots, y_n^0)$ е оптимален план за задачата (1), (2) и $H(x^0) = \beta(y^0)$.

Доказателство. *Необходимост.* Нека задачата (1), (2) е разрешима и $x^0 = (x_1^0, \dots, x_n^0) \in X$ е оптимален план, т. е.

$$H(x^0) = \max\{H(x) : x \in X\}.$$

Ще покажем, че $y^0 = (y_0^0, y_1^0, \dots, y_n^0)$, където

$$y_0^0 = \frac{1}{Q(x^0)}, \quad y_j^0 = \frac{x_j}{Q(x^0)}, \quad j = 1, \dots, n$$

($Q(x^0) > 0$ по предположение), е решение на задачата (4)–(6). Наистина $y^0 \in \Lambda$ (Теорема 2). Да допуснем, че y^0 не е оптимален план, т. е. съществува точка $y' = (y'_0, y'_1, \dots, y'_n) \in \Lambda$, за която

$$\beta(y') > \beta(y^0).$$

Имаме следните две възможности:

а) Ако $y'_0 > 0$, то точката $x' = \frac{1}{y'_0}(y'_1, \dots, y'_n) \in X$ и

$$H(x') = \frac{P(x')}{Q(x')} = \frac{p_0 + \frac{1}{y'_0} \sum_{j=1}^{n} p_j y'_j}{q_0 + \frac{1}{y'_0} \sum_{j=1}^{n} q_j y'_j} = \beta(y') > \beta(y^0) = \frac{P(x^0)}{Q(x^0)} = H(x^0),$$

което противоречи на оптималността на x^0.

б)Ако $y'_0 = 0$, то

$$\sum_{j=1}^{n} a_{ij} y_j^0 = 0, \quad i = 1, 2, \dots, m,$$

и ако означим $z = (y'_1, y'_2, \dots, y'_n)$, то $x^0 + \lambda z \in X$ за всяко $\lambda \geq 0$ и

$$H(x^0 + \lambda z) = \frac{P(x^0 + \lambda z)}{Q(x^0 + \lambda z)} = \frac{P(x^0) + \lambda(P(z) - p_0)}{Q(x^0) + \lambda(Q(z) - q_0)}.$$

Но

$$P(z) - p_0 = \beta(y') > \beta(y^0) = \frac{P(x^0)}{Q(x^0)},$$

$$Q(z) - q_0 = \sum_{j=1}^{n} q_j y'_j = 1$$

и при $\lambda \geq 0$ имаме

$$H(x^0 + \lambda z) > \frac{P(x^0) + \lambda \dfrac{P(x^0)}{Q(x^0)}}{Q(x^0) + \lambda} = \frac{P(x^0)}{Q(x^0)} = H(x^0),$$

което противоречи на оптималността на x^0.

Следователно y^0 е оптимален план на задачата (4)–(6) и $y_0^0 > 0$.

Достатъчност. Нека задачата (4)–(6) е разрешима и $y^0 = (y_0^0, y_1^0, \dots, y_n^0)$ е оптимален план, при който $y_0^0 > 0$. Ще покажем, че $x^0 = \dfrac{1}{y_0^0}(y_1^0, \dots, y_n^0)$ е решение на задачата (1), (2). Наистина $x^0 \in X$ (Теорема 2). Да допуснем, че x^0 не е оптимален план, т. е. съществува план $x' = (x_1', x_2', \dots, x_n')$, за който $H(x') > H(x^0)$. Но тогава точката $y' = \dfrac{1}{Q(x')}(1, x_1', \dots, x_n') \in \Lambda$ и

$$\beta(y') = \frac{P(x')}{Q(x')} = H(x') > H(x^0) = \beta(y^0),$$

което противоречи на оптималността на y^0. С това теоремата е доказана. □

Теорема 5. Ако $x^0 = (x_1^0, x_2^0, \dots, x_n^0)$ е оптимален опорен план на задачата (1), (2), то $y^0 = \dfrac{1}{Q(x^0)}(1, x_1^0, \dots, x_n^0)$ е опорен план на задачата (4)–(6).

Доказателство. Нека различните от нула компоненти на плана x^0 са първите му s компоненти $x_1^0, \dots, x_s^0$, $s \leq m$. Тогава матрицата

$$B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1s} \\ a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{ms} \end{pmatrix}$$

има ранг s. Разглеждаме матрицата от векторите на условията, съответстващи на различните от нула компоненти на плана y^0:

$$\overline{B} = \begin{pmatrix} a_{10} & a_{11} & a_{12} & \dots & a_{1s} \\ a_{20} & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m0} & a_{m1} & a_{m2} & \dots & a_{ms} \\ q_0 & q_1 & q_2 & \dots & q_s \end{pmatrix}.$$

Ако допуснем, че $\operatorname{rank} \overline{B} < s + 1$, тъй като $\operatorname{rank} B = s$, то съществуват s числа $u_1, u_2, \ldots, u_s$ такива, че

$$a_{i0} + \sum_{j=1}^{s} a_{ij} u_j = 0, \quad i = 1, 2, \ldots, m, \tag{8}$$

$$q_0 + \sum_{j=1}^{s} q_j u_j = 0. \tag{9}$$

Но

$$a_{i0} + \sum_{j=1}^{s} a_{ij} x_j^0 = 0, \quad i = 1, 2, \ldots, m, \tag{10}$$

$$q_0 + \sum_{j=1}^{s} q_j x_j^0 = Q(x^0) > 0. \tag{11}$$

От (8) и (10) получаваме

$$\sum_{j=1}^{s} a_{ij}(u_j - x_j^0) = 0, \quad i = 1, 2, \ldots, m,$$

и понеже $\operatorname{rank} B = s$, то

$$u_j = x_j^0, \quad j = 1, 2, \ldots, s,$$

т. е. условията (9) и (11) са противоречиви.

Следователно $\operatorname{rank} \overline{B} = s+1$, т. е. y^0 е опорен план. Неговата оптималност следва непосредствено от доказателството на теорема 2. □

Следствие 2. Ако задачата (1), (2) е разрешима, то задачата (4), (5) има опорен оптимален план $y^0 = (y_0^0, y_1^0, \ldots, y_n^0)$, за който $y_0^0 > 0$.

Следствие 3. Ако за всички оптимални опорни планове на задачата (4), (5) имаме $y_0 = 0$, то задачата (1), (2) е неразрешима (целевата функция не достига супремума си върху областта от планове на задачата).

3. Квадратично оптимиране

Квадратичните оптимизационни задачи са следващият добре изучен клас оптимизационни задачи след линейните. Създадени са много методи за решаване на квадратични задачи. Програми, реализиращи някои от тези методи, са част от базовото осигуряване на съвременните компютри.

3.1. Постановка на задачата. Една обща постановка на квадратичната задача е следната: Да се намери точката x^0 от множеството

$$X = \{x \in \mathbb{R}^n : Ax \le b\},$$

в която квадратичната функция $Q(x) = \langle Cx, x\rangle + \langle p, x\rangle$ има най-малка стойност, т. е.

$$Q(x^0) \le Q(x), \quad x \in X.$$

В дефинирането на X могат да участват и равенства, но това засега не е съществено, тъй като всяко равенство е еквивалентно на две неравенства.

Разликата между квадратичната и линейната задача е в целевата функция. Тук целевата функция е квадратична. Това води до съществена разлика в свойствата на двата класа задачи. Например линейната функция достига екстремума си във връх на многостена, а квадратичната функция може да достигне екстремума си във всяка точка на многостена (фиг. 28).

Преди да пристъпим към разглеждане на свойствата на квадратичните задачи, ще се спрем на някои свойства на квадратичните функции.

3.2. Квадратични форми и квадратични функции. Както знаем, квадратична форма наричаме хомогенна функция от втора степен, т. е. функция от вида

$$\begin{aligned} Q(x) = Q(x_1, x_2, \dots, x_n) = c_{11}x_1^2 + 2c_{12}x_1x_2 + 2c_{13}x_1x_3 + \dots \\ + 2c_{1n}x_1x_n + c_{22}x_2^2 + 2c_{23}x_2x_3 + \dots + 2c_{2n}x_2x_n + \dots + c_{nn}x_n^2. \end{aligned}$$

Ако за $i > j$ приемем $c_{ij} = c_{ji}$, квадратичната форма можем да запишем в следният симетричен вид:

$$\begin{aligned} Q(x) = {} & c_{11}x_1^2 + c_{12}x_1x_2 + \dots + c_{1n}x_1x_n \\ + {} & c_{21}x_2x_1 + c_{22}x_2^2 + \dots + c_{2n}x_2x_n \\ & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ + {} & c_{n1}x_nx_1 + c_{n2}x_nx_2 + \dots + c_{nn}^2x_n. \end{aligned}$$

Като означим с C квадратната симетрична матрица:

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

$Q(x)$ се записва като скаларно произведение на векторите Cx и x, т. е.

$$Q(x) = \langle Cx, x\rangle.$$

Ако приемем, че x е вектор-стълб, $x = (x_1, \dots, x_n)^{\mathrm{T}}$, и го разгледаме като матрица $n \times 1$, то $Q(x)$ може да се запише и в матрична форма:

$$Q(x) = x^{\mathrm{T}} Cx.$$

Както се вижда, всяка квадратична форма е свързана с квадратна симетрична матрица — матрицата от коефициентите ѝ.

Да припомним, че квадратичната форма $\langle Cx, x\rangle$ наричаме неотрицателно дефинирана (положително полудефинитна), ако $\langle Cx, x\rangle \geq 0$ за всяко $x \in \mathbb{R}^n$, и положително дефинитна, ако $\langle C, x\rangle > 0$ за всяко $x \neq 0$. Ако са изпълнени обратните неравенства, квадратичната форма наричаме съответно неположително дефинитна или отрицателно дефинитна. Казваме, че квадратичната форма е дефинитна, ако притежава някое от изброените видове дефинитност. Неотрицателно (неположително) дефинитните форми се наричат още положително (отрицателно) полудефинитни.

Ето няколко примера за дефинитни форми:

1) $x_1^2 - 2x_1x_2 + x_2^2$ — положително полудефинитна;

2) $x_1^2 + x_2^2$ — положително дефинитна;

3) $x_1^2 - x_2^2$ — не е дефинитна.

Понятието дефинитност автоматично се пренася и върху квадратните симетрични матрици. Една квадратна симетрична матрица C наричаме положително полудефинитна, ако съответната ѝ квадратична форма $\langle Cx, x\rangle$ е такава. Аналогично се определят и останалите видове дефинитност. Често с $C \geq 0$ означаваме, че C е неотрицателно дефинитна, а с $C > 0$ — положително дефинитна.

Сега ще се спрем на някои свойства на квадратичните форми и квадратните симетрични матрици, свързани с понятието дефинитност.

Най-напред ще отбележим някои очевидни неща.

Ако в една квадратична форма положим някои от променливите равни на нула, ще получим пак квадратична форма (с по-малък брой променливи). Ако изходната квадратична форма е дефинитна, то получената квадратична форма запазва или подобрява дефинитността си (от полудефинитна може да стане дефинитна), така че във всяка симетрична подматрица (получена чрез задраскване на съответни редове и стълбове) се запазва или подобрява дефинитността на изходната матрица. В частност диагоналните елементи на положително полудефинитна (дефинитна) матрица са неотрицателни (положителни).

Ако диагоналният елемент на полудефинитна матрица е нула, всички елементи в реда и стълба, съдържащи този елемент, са нули.

Ще докажем и някои по-малко очевидни свойства, свързани с дефинитността.

Свойство 1. Ако C е неотрицателно дефинитна, то $Cx = 0$ за всяко x, за което $\langle Cx, x\rangle = 0$.

Доказателство. Нека $\langle Cx, x\rangle = 0$. За произволни $y \in \mathbb{R}^n$ и $\lambda \in \mathbb{R}^1$ имаме

$$
\begin{aligned}
0 &\le \langle C(y + \lambda x), y + \lambda x\rangle = \langle Cy, y\rangle + 2\lambda\langle Cx, y\rangle + \lambda^2\langle Cx, x\rangle \\
&= \langle Cy, y\rangle + 2\lambda\langle Cx, y\rangle.
\end{aligned}
$$

Но това е възможно, само ако

$$\langle Cx, y\rangle = 0$$

и понеже y е произволно — само ако

$$Cx = 0.$$

Така доказахме необходимостта на условието $Cx = 0$ за равенството $\langle Cx, x\rangle = 0$. Разбира се, това условие е и достатъчно, но това е очевидно. □

Свойство 2. а) Ако C е положително дефинитна матрица, тя е неособена.

б) Ако C е неотрицателно дефинитна и неособена, тя е положително дефинитна.

Доказателство. а) Нека $C > 0$. Тогава $\langle x, Cx\rangle > 0$ за всяко $x \neq 0$ и $Cx \neq 0$ за всяко $x \neq 0$. Следователно C е неособена.

б) Ако C е неособена неотрицателно дефинитна матрица, то $\langle Cx, x\rangle \ge 0$ за всяко x. Ако допуснем че, за някое $\overline{x} \neq 0$ имаме $\langle C\overline{x}, \overline{x}\rangle = 0$, то от свойство 1 следва, че $C\overline{x} = 0$. Стигаме до противоречие. Следователно $\langle Cx, x\rangle > 0$ за всяко $x \neq 0$, т. е. C е положително дефинитна. □

Свойство 3. Нека C е квадратна симетрична матрица с размери $n \times n$ и A е произволна матрица с размери $n \times m$. Тогава:

а) ако C е неотрицателно дефинитна, матрицата $A^{\mathrm{T}}CA$ е неотрицателно дефинитна.

б) ако C е положително дефинитна и $\operatorname{rank} A = m$, матрицата $A^{\mathrm{T}}CA$ е положително дефинитна.

Доказателство. Ако положим $Ax = y$, то

$$\langle A^{\mathrm T}CAx, x\rangle = \langle CAx, Ax\rangle = \langle Cy, y\rangle \geq 0$$

за всяко $y \in \mathbb{R}^m$ и следователно за всяко $x \in \mathbb{R}^n$, т. е. $A^{\mathrm T}CA \geq 0$. Ако $\operatorname{rank} A = m$, то от $x \neq 0$ следва $y \neq 0$. Тогава, ако матрицата $C > 0$, то

$$\langle A^{\mathrm T}CAx, x\rangle = \langle CAx, Ax\rangle = \langle Cy, y\rangle > 0$$

за всяко $y \neq 0$ и следователно за всяко $x \neq 0$, т. е. $A^{\mathrm T}CA > 0$. □

Свойство 4. При произволна матрица A с размери $n \times m$ матрицата $A^{\mathrm T}A$ е неотрицателно дефинитна и положително дефинитна, ако $\operatorname{rank} A = m$.

Следва от свойство 3, ако положим $C = I$ (I е единичната матрица).

Свойство 5. Ако C е положително дефинитна, то и C^{-1} е положително дефинитна.

Следва от свойство 3 при $A = A^{\mathrm T} = C^{-1}$.

Свойство 6. Ако C е положително (полу)дефинитна матрица с размери $n \times n$ и $C = \left(\begin{smallmatrix} P & Q \\ Q^{\mathrm T} & S \end{smallmatrix}\right)$, където P е положително дефинитна матрица с размери $n_1 \times n_1$, S — неотрицателно дефинитна матрица с размери $n_2 \times n_2$, $n_2 = n - n_1$, и Q — матрица с размери $n_1 \times n_2$, то и матрицата $S - Q^{\mathrm T}P^{-1}Q$ е положително (полу)дефинитна.

Доказателство. Тук ще си послужим с матричен запис. Ще разделим вектора $x = (x_1, \dots, x_n)^{\mathrm T}$ на две части $u = (x_1, \dots, x_{n_1})^{\mathrm T}$ и $v = (x_{n_1+1}, \dots, x_n)^{\mathrm T}$ така, че $x = \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$. И така

$$\langle Cx, x\rangle = x^{\mathrm T}Cx = (u^{\mathrm T}v^{\mathrm T})\left(\begin{smallmatrix} P & Q \\ Q^{\mathrm T} & S \end{smallmatrix}\right)\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) = u^{\mathrm T}Pu + v^{\mathrm T}Q^{\mathrm T}u + u^{\mathrm T}Qv + v^{\mathrm T}Sv \geq 0$$

за всяко $u \in \mathbb{R}^{n_1}$ и $v \in \mathbb{R}^{n_2}$. Като положим $u = -P^{-1}Qv$, ще получим

$$-v^{\mathrm T}Q^{\mathrm T}P^{-1}Qv + v^{\mathrm T}Sv = v^{\mathrm T}(S - Q^{\mathrm T}P^{-1}Q)v \geq 0$$

за всяко $v \in \mathbb{R}^{n_2}$. Следователно $S - Q^{\mathrm T}P^{-1}Q$ е неотрицателно дефинитна. Очевидно, ако C е положително дефинитна, то $S - Q^{\mathrm T}P^{-1}Q$ е положително дефинитна. □

Очевидно свойства 1–6 остават в сила, ако заменим понятието положително (неотрицателно) дефинитна с отрицателно (неположително) дефинитна.

Да припомним че, ако $Q(x) = \langle Cx, x\rangle$, то

$$\begin{aligned} Q'(x) &= \left(\frac{\partial Q(x)}{\partial x_1}, \frac{\partial Q(x)}{\partial x_2}, \dots, \frac{\partial Q(x)}{\partial x_n}\right)^{\mathrm T} \\ &= 2\left(\sum_{j=1}^{n} c_{1j}x_j, \sum_{j=1}^{n} c_{2j}x_j, \dots, \sum_{j=1}^{n} c_{nj}x_j\right) = 2Cx \end{aligned}$$

и

$$Q''(x) = 2C.$$

Тогава от теорема 2, § 2, гл. 4, получаваме, че квадратичната форма $\langle x, Cx\rangle$ е изпъкнала (вдлъбната) тогава и само тогава, когато C е неотрицателно (неположително) дефинитна (пример 1, § 2, гл. 4).

Сега ще се спрем на някои свойства на квадратичните функции, т. е. функции от вида

$$Q(x) = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle + q,$$

където C е квадратна симетрична матрица с размери $n \times n$, $p \in \mathbb{R}^n$ и $q \in \mathbb{R}^1$.

Тъй като за редица свойства (например екстремалните) на квадратичната функция свободният член няма значение, често ще смятаме, че $q = 0$.

И тук да припомним че

$$Q'(x) = Cx + p, \quad Q''(x) = C.$$

Така че $Q(x)$ е изпъкнала (вдлъбната) тогава и само тогава, когато C е неотрицателно (неположително) дефинитна и строго изпъкнала (строго вдлъбната) тогава и само тогава, когато C е положително (отрицателно) дефинитна.

Теорема 1 (съществуване на екстремум)**.** Ако квадратичната функция $Q(x) = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle$ е ограничена отдолу (отгоре) в множеството $X = Y + K$, където Y е компактно множество, а K — крайно породен конус, то $Q(x)$ достига инфимума (супремума) си в X.

Доказателство. Нека за определеност функцията $Q(x)$ е ограничена отдолу в X. Тъй като за всички $y \in Y$, $z \in K$ и $\lambda \geq 0$ имаме $x = y + \lambda z \in X$ и

$$Q(x) = Q(y + \lambda z) = \frac{1}{2}\langle Cy, y\rangle + \langle p, y\rangle + \lambda\langle Cy + p, z\rangle + \frac{1}{2}\lambda^2\langle Cz, z\rangle,$$

то $\langle Cz, z\rangle \geq 0$ за всяко $z \in K$ и ако за някое $\bar{z} \in K$ имаме $\langle C\bar{z}, \bar{z}\rangle = 0$, то $\langle Cy + p, \bar{z}\rangle \geq 0$ за всяко $y \in Y$. Да означим $L = \{z \in K : \|z\| = 1\}$. Разглеждаме двата възможни случая:

Случай 1. Функцията $\langle Cz, z\rangle > 0$ за всяко $z \in L$. От непрекъснатостта на функцията $\langle Cz, z\rangle$, компактността на множеството L и $\langle Cz, z\rangle > 0$ в L следва, че съществува константа $\delta > 0$ такава, че $\langle Cz, z\rangle \geq \delta$ за всяко $z \in L$. Тогава, ако $x = y + \lambda z$, $y \in Y$, $z \in L$, $\lambda \geq 0$,то

$$\begin{aligned} Q(x) = Q(y + \lambda z) &= \frac{1}{2}\langle Cy, y\rangle + \langle p, y\rangle + \lambda\langle Cy + p, z\rangle + \frac{1}{2}\lambda^2\langle Cz, z\rangle = \varphi(y, z, \lambda) \\ &\geq \min_{\lambda\geq 0}\varphi(y, z, \lambda) = \varphi(y, z, \max) = \psi(y, z), \end{aligned}$$

така че

$$\inf_X Q(x) = \inf_{(y,z)\in Y\times L} \psi(y,z).$$

Но $\psi(y,z)$ е непрекъсната функция в компактното множество $Y \times L$ и следователно съществува точка $(y^0, z^0) \in Y \times L$, за която

$$\psi(y^0, z^0) = \inf\{\psi(y,z) : (y,z) \in Y \times L\}.$$

Но

$$x^0 = y^0 + \max\left\{0, \frac{\langle Cy^0 + p, z^0\rangle}{\langle z^0, Cz^0\rangle}\right\}, \quad z^0 \in X,$$

и

$$Q(x^0) = \psi(y^0, z^0).$$

Следователно

$$Q(x^0) = \inf\{Q(x) : x \in X\}.$$

Случай 2. Нека съществува такова $\bar{z} \in K\setminus\{0\}$, че $\langle C\bar{z}, \bar{z}\rangle = 0$. В този случай ще докажем твърдението с индукция по броя на елементите, пораждащи конуса K.

Ако $K = \mathrm{con}(z_1, z_2, \ldots, z_s)$, ще означаваме

$$K_q = \mathrm{con}(z_1, z_2, \ldots, z_{q-1}, z_{q+1}, \ldots, z_s)$$

и $X_q = Y + K_q$, $q = 1, 2, \ldots, s$. Нека $s = 1$. Тогава $K = \{\lambda z_1, \lambda \geq 0\}$ и според условията $\langle Cz_1, z_1\rangle = 0$, а оттук $\langle Cy + p, z_1\rangle \geq 0$ за всяко $y \in Y$. Да вземем произволно $x \in X$. Тъй като $x = y + \lambda z_1$, $y \in Y$, $\lambda \geq 0$, и

$$Q(x) = Q(y + \lambda z_1) = Q(y) + \lambda\langle Cy + p, z_1\rangle + \frac{1}{2}\lambda^2\langle Cz_1, z_1\rangle \geq Q(y),$$

имаме

$$\inf\{Q(x) : x \in X\} = \inf\{Q(y) : y \in Y\}.$$

Множеството Y е компактно и функцията $Q(x)$ достига инфимума си в него. Следователно $Q(x)$ достига инфимума си и в X.

Нека твърдението е вярно за всяко $s < k$. Ще покажем, че е вярно и при $s = k$. От индуктивното предположение или от първия случай следва, че $Q(x)$ достига инфимума си върху всяко множество X_q, $q \in 1, 2, \ldots, k$. Нека точката $\bar{z} \in K \setminus \{0\}$ удовлетворява равенството $\langle C\bar{z}, \bar{z}\rangle = 0$, т. е. $\langle Cx + p, \bar{z}\rangle \geq 0$ за всяко $x \in X$. Избираме произволно $x \in X$. Ако

$$x = y + z, \quad y \in Y, \quad z \in K,$$

$$z = \sum_{i=1}^{k} \alpha_i z_i, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, k,$$

$$\overline{z} = \sum_{i=1}^{k} \overline{\alpha}_i z_i, \quad \overline{\alpha}_i \geq 0, \quad i = 1, 2, \dots, k,$$

и ако $\lambda_q = \min\limits_{\overline{\alpha}_i > 0} \alpha_i / \overline{\alpha}_i = \alpha_q / \overline{\alpha}_q$, $i = 1, 2, \dots, k$, то

$$\widehat{z} = z - \lambda_q \overline{z} \in K_q \quad \text{и} \quad \widehat{x} = y + \widehat{z} \in X_q.$$

Но

$$\begin{aligned}
Q(\widehat{x}) &= Q(y + \widehat{z}) = Q(y + z - \lambda_q \overline{z}) \\
&= \frac{1}{2}\langle Cy, y\rangle + \langle p, y\rangle + \langle Cy + p, z\rangle + \frac{1}{2}\langle Cz, z\rangle - \lambda_q \langle C(y + z) + p, \overline{z}\rangle \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - \frac{1}{2}\lambda_q^2 \langle C\overline{z}, \overline{z}\rangle \\
&= Q(y + z) - \lambda_q \langle C(y + z) + p, \overline{z}\rangle - \frac{1}{2}\lambda_q^2 \langle C\overline{z}, \overline{z}\rangle \\
&\leq Q(y + z) = Q(x),
\end{aligned}$$

т. е. за всяко $x \in X$ има такава точка $\widehat{x} \in \bigcup\limits_{q=1}^{k} X_q$, че $Q(\widehat{x}) \leq Q(x)$. Следователно

$$\inf_X Q(x) = \inf_{\bigcup\limits_{q=1}^{k} X_q} Q(x) = \min_{1 \leq q \leq k} \{ \inf_{x \in X_q} Q(x) \}.$$

Но $Q(x)$ достига инфимума си върху всяко X_q, така че $Q(x)$ достига инфимума си в X. $\square$

Задача 1. Докажете теорема 1 при условие, че функцията

$$Q(x) = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle$$

е: а) строго изпъкнала; б) изпъкнала.

3.3. Условия за екстремум на квадратични функции Квадратичните оптимизационни задачи могат да се запишат по един от следните начини:

$$\text{(А)} \qquad \min\left\{ Q(x) = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle,\ Ax \leq b, x \geq 0 \right\},$$

$$\text{(Б)} \qquad \min\left\{ Q(x) = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle,\ Ax = b, x \geq 0 \right\},$$

(В) $$\min\left\{Q(x) = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle,\ Ax \le b\right\}.$$

Ако задачи (Б) или (В) искаме да приведем в задача (А), можем да заменим всяко равенство с две неравенства, всяка променлива без ограничения за знак — с разлика на две неотрицателни променливи.

Функцията на Лагранж за задачите (А), (Б) и (В) е

$$L(x, \lambda) = Q(x) + \langle \lambda, Ax - b\rangle = \frac{1}{2}\langle Cx, x\rangle + \langle p, x\rangle + \langle \lambda, Ax - b\rangle,$$

където $\lambda = (\lambda_1, \dots, \lambda_m)$ е векторът на Лагранж, а λ_i, $i = 1, 2, \dots, m$, са множителите на Лагранж.

Ако означим $y = -L'_\lambda(x, \lambda) = -Ax + b$, $v = L'_x(x, \lambda) = Cx + p + A^T\lambda$, то условията за седлова точка за трите основни форми са съответно:

- за форма (А):

(1) $$\begin{aligned} &\text{а) } Ax + y = b, && \text{в) } x \ge 0,\ v \ge 0,\ y \ge 0,\ \lambda \ge 0,\\ &\text{б) } Cx - v + A^T\lambda = -p, && \text{г) } \langle x, v\rangle + \langle \lambda, y\rangle = 0; \end{aligned}$$

- за форма (Б):

(2) $$\begin{aligned} &\text{а) } Ax = b, && \text{в) } x \ge 0,\ v \ge 0,\\ &\text{б) } Cx - v + A^T\lambda = -p, && \text{г) } \langle x, v\rangle = 0; \end{aligned}$$

- за форма (В):

(3) $$\begin{aligned} &\text{а) } Ax + y = b, && \text{в) } y \ge 0,\ \lambda \ge 0,\\ &\text{б) } Cx - v + A^T\lambda = -p, && \text{г) } \langle \lambda, y\rangle = 0; \end{aligned}$$

Да се спрем на квадратичната задача (А) и съответната ѝ система (1). Както знаем, съществуването на решение на (1) е необходимо условие, а в случай на изпъкналост на функцията $Q(x)$ — и достатъчно за съществуване на решение на оптимизационната задача (А). И така, ако матрицата C е неотрицателно дефинитна, то n-мерен вектор x^0 е решение на задача (А) тогава и само тогава, когато съществуват n-мерен вектор v^0 и m-мерни вектори λ^0 и y^0 такива, че $(x^0, v^0, \lambda^0, y^0)$ е решение на системата (1).

По-нататък ще предполагаме, че C е неотрицателно дефинитна, т. е. функцията $Q(x)$ е изпъкнала.

Достатъчността на условието на Кун—Такър дава възможност да се замени решаването на оптимизационната задача с решаване на система от равенства и неравенства. Така например, ако намерим едно решение $(x^0, v^0, \lambda^0, y^0)$ на системата (1), то x^0 ще бъде решение на задачата (А). Голяма част от числените методи за решаване на квадратични оптимизационни задачи използват тази идея и се различават само по начина, по който решават съответната система неравенства. Понеже първите три групи условия — а), б) и в) — в (1) са линейни, трудността при решаването на тази система идва от последното нелинейно условие г). Затова често се оказва полезна следната:

Теорема 2. Ако $Q(x)$ е изпъкнала, то условията а), б) и в) в (1) са необходими и достатъчни за съществуване на решение на оптимизационната задача.

Доказателство. *Необходимост.* Очевидна е, тъй като това са част от условията на Кун—Такър.

Достатъчност. Нека $\overline{x}$, $\overline{v}$, $\overline{\lambda}$, $\overline{y}$ удовлетворяват условията а), б) и в) в (1). Точката $\overline{x}$ е допустима за оптимизационната задача (А) и за всяка друга точка x имаме $Q(x) \geq Q(\overline{x}) + \langle Q'(\overline{x}), x - \overline{x}\rangle$, т. е.

$$\begin{aligned}
Q(x) - Q(\overline{x}) &\geq \langle C\overline{x} + p, x - \overline{x}\rangle = \langle \overline{v} - A^T\overline{\lambda}, x - \overline{x}\rangle \\
&= \langle x, \overline{v}\rangle - \langle \overline{x}, \overline{v}\rangle - \langle A^T\overline{\lambda}, x\rangle + \langle A^T\overline{\lambda}, \overline{x}\rangle \\
&= \langle x, v\rangle - \langle x, \overline{v}\rangle - \langle A\overline{x}, \overline{\lambda}\rangle + \langle Ax, \overline{\lambda}\rangle \\
&= \langle x, \overline{v}\rangle - \langle \overline{x}, \overline{v}\rangle - \langle b - y, \overline{\lambda}\rangle + \langle b - \overline{y}, \overline{\lambda}\rangle \\
&= \langle x, v\rangle - \langle x, \overline{v}\rangle + \langle \overline{y}, \overline{\lambda}\rangle - \langle y, \overline{\lambda}\rangle \\
&\geq -\langle \overline{x}, \overline{v}\rangle - \langle \overline{y}, \overline{\lambda}\rangle.
\end{aligned}$$

Следователно функцията $Q(x)$ е ограничена отдолу и в множеството от допустими точки и достига инфимума си в това множество. Ако x^0 е решение, от горните неравенства се вижда, че

$$Q(\overline{x}) - \langle \overline{x}, \overline{v}\rangle - \langle \overline{y}, \overline{\lambda}\rangle \leq Q(x^0) \leq Q(\overline{x}).$$

Тези неравенства се използуват често при приближено решаване на задачата за оценка на дадено приближено решение. □

Библиография

[1] Алексеев, В.М., Э.М. Галеев, В.М. Тихомиров. Сборник задач по оптимизации. М., Наука, 1984.

[2] Астафьев, Н.Н. Линейные неравенства и выпуклость. М., Наука, 1982.

[3] Базара, М., К. Шетти. Нелинейное программирование. М., Мир, 1982.

[4] Васильев, Ф. П. Численные методы решения экстремальных задач. М., Наука, 1980.

[5] Илин, В.А., В.А. Садовничи, Бл. Х. Сендов. Математически анализ, ч. 1. С., Наука и изкуство, 1979.

[6] Карманов, В.Т. Математическое программирование. М., Наука, 1975.

[7] Кендеров, П.С. Изпъкнали множества. С., Народна просвета, 1985.

[8] Иванов, Г.Х., Р.П. Иванов. Теореми существования экстремума для многочленов на неограниченных множествах. Год. Соф. унив., Фак. мат. и мех., **70**, 1975/1976.

[9] Пшеничный, Б.П. Выпуклый анализ и экстремальные задачи. М., Наука, 1980.

[10] Рокафеллар, Р.Т. Выпуклый анализ. М., Мир, 1973.

[11] Христов, Г., Р. Калтинска. Математическо оптимиране, ч. 1. Линейно оптимиране. С., Наука и изкуство, 1972.

[12] Христов, Г. Математическо оптимиране. С., Софийски университет, 1973.

Петър Кендеров, Георги Христов,
Асен Дончев

МАТЕМАТИЧЕСКО ОПТИМИРАНЕ

Българска
Първо издание

Рецензенти Йордан Митев, Тодор Гичев
Редактор Елисавета Михайлова
Художник на корицата Пенчо Мутафчиев
Худ. редактор Красимира Михайлова
Техн. редактор Божидар Стоянов

Изд. индекс 205
Код 26 / 9534622231 / 4805–331–89
Подписана за печат май 1989 г.
Формат 59 × 84 / 16 Тираж 2200+85 ЛГ 111–4
Печатни коли 10 Изд. коли 9,30 УИК 8,78
Цена 0,53 лв.

Университетско издателство „Климент Охридски“
Печат и подвързия — Печатна база БЕК.