Kapitola 7

Extrémy funkcí více proměnných

Celá tato kapitola se zabývá pouze jedinou otázkou: jakým způsobem zjistit bod či body, ve kterých daná funkce nabývá extrém. Z Kapitoly 4 už víme, že každá spojitá funkce na uzavřené omezené množině nabývá svého minima i maxima. Příslušná věta ale nedává ani sebemenší návod, jak body, ve kterých se toto odehrává, nalézt.

Začneme nejdříve s tzv. lokálními extrémy.

1 Lokální extrémy

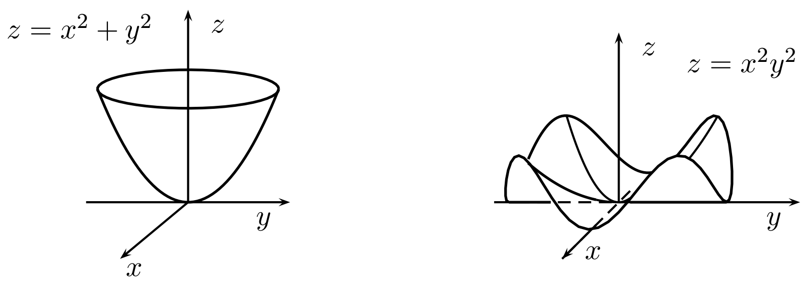
Následující definice přesně vymezuje pojem, který se pokusíme v této části studovat.

Definice 7.1. *Nechť $\mathbb{R}^n$ je euklidovský prostor a $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ funkce. Řekneme, že f má v bodě $\boldsymbol{x} \in \mathbb{R}^n$* **lokální minimum** *(resp.* **maximum***), jestliže existuje okolí U bodu $\boldsymbol{x}$, že*

$$f(\boldsymbol{x}) \le f(\boldsymbol{y}) \quad (resp.\ f(\boldsymbol{x}) \ge f(\boldsymbol{y})\ )$$

pro všechna $\boldsymbol{y} \in U$. Bude-li dokonce $f(\boldsymbol{x}) < f(\boldsymbol{y})$ (resp. $f(\boldsymbol{x}) > f(\boldsymbol{y})$) pro všechny body $y \in U \setminus \{\boldsymbol{x}\}$, mluvíme o ostrém lokálním minimu (resp. maximu).

Nabývá-li f v $\boldsymbol{x}$ lokální minimum nebo maximum, říkáme, že f má v $\boldsymbol{x}$ lokální extrém. Podobně ostrý lokální extrém znamená ostré lokální minimum nebo ostré lokální maximum.



(a) (b)

Obr. 7.1.

Rozdíl mezi ostrým a neostrým extrémem je jistě zřejmý. Na obr. 7.1 jsou příklady ilustrující tuto odlišnost. Obrázek 7.1.(a) představuje ostré lokální minimum funkce $f(x,y) = x^2 + y^2$ v bodě $(0,0)$, kdežto v části 7.1.(b) je $(0,0)$ neostré lokální minimum funkce $f(x,y) = x^2y^2$.

1.1 Stacionární body

Bod, podezřelý z toho, že v něm funkce nabývá lokální extrém, odhalíme z jednoduché geometrické podmínky. Tečná rovina (event. nadrovina v případě tří a více proměnných) musí být v takovém bodě kolmá na osu funkčních hodnot. Diferenciál, který zadává rovinu tohoto typu, je nutně nulový. Pro funkce třídy alespoň C^1 pak z Tvrzení 5.1 plyne, že nulovost diferenciálu je ekvivalentní s požadavkem

$$\frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0.$$

V této chvíli už nás proto nepřekvapí, že platí

Věta 7.2. *Nechť $G \subset \mathbb{R}^n$ je otevřená podmnožina euklidovského prostoru $\mathbb{R}^n$ a nechť funkce $f\colon G \longrightarrow \mathbb{R}$ je třídy C^1 na G. Je-li $\boldsymbol{x} \in G$ bod, ve kterém f nabývá lokální extrém, pak*

$$\partial_{\boldsymbol{h}} f(\boldsymbol{x}) = 0$$

pro každý směr $\boldsymbol{h} \in \mathbb{R}^n$.

Důkaz. Předpokládejme, že $\boldsymbol{x}$ je bod lokálního minima funkce f. (Pro maximum by byl důkaz zcela analogický.) Nechť $\boldsymbol{h} \in \mathbb{R}^n$ je daný vektor. Víme, že pro jisté okolí U bodu $\boldsymbol{x}$ platí

$$f(\boldsymbol{x}) \leq f(\boldsymbol{y}), \tag{7.1}$$

kdykoli $\boldsymbol{y} \in U$. Můžeme najít tak malý interval $\langle -\delta, \delta\rangle \subset \mathbb{R}$, že pro všechna $t \in \langle -\delta, \delta\rangle$ leží bod $\boldsymbol{x} + t\boldsymbol{h}$ stále v okolí U. V tom případě funkce $\varphi\colon \langle -\delta, \delta\rangle \longrightarrow \mathbb{R}$ definovaná předpisem

$$\varphi(t) = f(\boldsymbol{x} + t\boldsymbol{h})$$

má v bodě 0 lokální minimum: podle (7.1) je totiž

$$\varphi(0) = f(\boldsymbol{x}) \leq f(\boldsymbol{x} + t\boldsymbol{h}) = \varphi(t).$$

Protože f je třídy C^1, existuje její derivace ve směru $\boldsymbol{h}$. Pro funkci φ to znamená, že existuje $\varphi'(0)$. Podle známé věty pro funkce jedné proměnné musí být $\varphi'(0) = 0$. Tudíž

$$0 = \varphi'(0) = \lim_{t\to 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t\to 0} \frac{f(\boldsymbol{x} + t\boldsymbol{h}) - f(\boldsymbol{x})}{t} = \partial_{\boldsymbol{h}} f(\boldsymbol{x}).$$

□

Důsledek 7.3. *Má-li C^1 funkce f v bodě $\boldsymbol{x}$ lokální extrém, pak v tomto bodě platí*

$$\frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0. \tag{7.2}$$

Důkaz. Použijeme Větu 7.2 s $\boldsymbol{h} = \boldsymbol{e}_i$ pro $i = 1, \ldots, n$. □

Definice 7.4. *Bod $\boldsymbol{x} \in \mathbb{R}^n$, pro který platí $\dfrac{\partial f}{\partial x_1}(\boldsymbol{x}) = \cdots = \dfrac{\partial f}{\partial x_n}(\boldsymbol{x}) = 0$ se nazývá* **stacionární bod** *funkce f.*

Stručně zapsaná podmínka (7.2) pro stacionární bod je

$$\operatorname{grad} f = 0. \tag{7.3}$$

Někdo by se mohl zeptat, co lze říci o funkci f v bodě lokálního extrému v případě, že f není třídy C^1. Pak samozřejmě nemusí Věta 7.2 platit z jednoduchého důvodu: příslušná derivace ve směru $\boldsymbol{h}$ vůbec neexistuje. Že to může nastat, vidíme např. na funkci

$$f(x, y) = |x| + |y|,$$

viz obr. 7.2. Bod $(0, 0)$ je minimum, ale díky ostrému vrcholu na grafu funkce f neexistuje derivace v žádném směru. Obecně platí následující alternativa: v bodě lokálního extrému buďto v nějakém směru derivace neexistuje nebo jsou všechny nulové.

Příklad 7.5. Nalezněte všechny stacionární body funkce $f(x, y) = \dfrac{x^2}{2} - xy - x + y$ a funkce $g(x, y, z) = \cos(x^2 + y^2 + z^2)$.

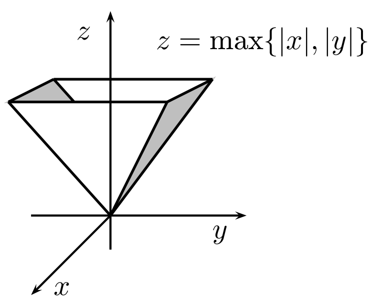
Řešení: V případě funkce f hledáme všechny body $(x, y) \in \mathbb{R}^2$ splňující

$$\frac{\partial f}{\partial x} = x - y - 1 = 0 \quad \text{a} \quad \frac{\partial f}{\partial y} = -x + 1 = 0.$$

Z druhé rovnice ihned plyne, že $x = 1$. Dosazením do první dostaneme $y = 0$. Závěr je takový, že funkce f má jediný stacionární bod $(1, 0)$.

U funkce g postupujeme podobně. Hledáme řešení soustavy tří rovnic

$$\frac{\partial g}{\partial x} = 2x\cos(x^2 + y^2 + z^2) = 0,$$
$$\frac{\partial g}{\partial y} = 2y\cos(x^2 + y^2 + z^2) = 0,$$
$$\frac{\partial g}{\partial z} = 2z\cos(x^2 + y^2 + z^2) = 0.$$



Obr. 7.2.

Rozlišíme dva případy. Buď $\cos(x^2 + y^2 + z^2) = 0$. Pak

$$x^2 + y^2 + z^2 = \frac{\pi}{2} + k\pi, \quad k = 0, 1, \dots$$

Nebo $\cos(x^2 + y^2 + z^2) \neq 0$, a tak ke splnění soustavy musí nutně být $x = y = z = 0$. Z těchto dvou případů dostáváme, že množina stacionárních bodů funkce g je

$$(0,0,0) \cup \Big\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = \pi/2 + k\pi,\ k\ = 0, 1, \dots\Big\}.$$

Je to nekonečně mnoho soustředných sfér s poloměry $\sqrt{\pi/2 + k\pi}$, $k = 0, 1 \dots$ spolu s počátkem $(0, 0, 0)$.

Je třeba mít na paměti, že stacionární bod je pouhým kandidátem na nabývání lokálního extrému. Jestli se nabývá lokální minimum či maximum nebo jestli se vůbec žádného extrému nenabývá, rozhoduje ve většině případů chování druhého diferenciálu. Myšlenka v pozadí je celkem jednoduchá. V okolí stacionárního bodu $\boldsymbol{x}$ rozvineme funkci v Taylorovu řadu do řádu 2 podle Věty 6.19. Pro přírůstek $\boldsymbol{h}$ z jistého malého okolí nuly U platí

$$f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x}) \approx \boldsymbol{h} \cdot \operatorname{grad} f(\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{h} \cdot \operatorname{grad})^2 f(\boldsymbol{x}).$$

Protože $\boldsymbol{x}$ je stacionární, je podle (7.3) $\operatorname{grad} f(\boldsymbol{x}) = 0$. Takže

$$\begin{aligned}(7.4) \quad f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x}) &\approx \frac{1}{2}(\boldsymbol{h} \cdot \operatorname{grad})^2 f(\boldsymbol{x}) = \frac{1}{2}\Big(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n}\Big)^2 f(\boldsymbol{x}) \\ &= \frac{1}{2}\Big(\sum_{\substack{i=1 \\ j=1}}^{n} h_i h_j \frac{\partial^2}{\partial x_i \partial x_j}\Big) f(\boldsymbol{x}) = \frac{1}{2}\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}].\end{aligned}$$

(Poslední rovnost plyne z (6.18).) Bude-li funkce $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}]$ kladná pro $\boldsymbol{h} \in U \setminus \{\boldsymbol{0}\}$, tak (7.4) říká, že v $\boldsymbol{x}$ je lokální minimum. Bude-li naopak výraz $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}]$ záporný, pak v $\boldsymbol{x}$ je lokální maximum. Nenastane-li však ani jeden z těchto případů, tzn. druhý diferenciál $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}]$ je pro jistá $\boldsymbol{h} \in U$ kladný a pro jiná záporný, tak v $\boldsymbol{x}$ lokální extrém nenastává.

Vidíme, že role druhého diferenciálu zde velmi podstatně vystupuje do popředí. Proto se musíme na chvíli pozastavit u funkcí, které proměnné $\boldsymbol{h}$ přiřadí hodnotu $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}]$.

1.2 Kvadratické formy

Z Definice 6.14 už víme, co je bilineární forma. Z každé bilineární formy lze snadno vytvořit jiný typ funkce, tzv. kvadratickou formu.

Definice 7.6. *Nechť* $\psi\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ *je bilineární forma na euklidovském prostoru* $\mathbb{R}^n$*. Funkce* $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ *definovaná*

$$q(\boldsymbol{h}) = \psi(\boldsymbol{h}, \boldsymbol{h}),\ \boldsymbol{h} = (h_1, \dots, \boldsymbol{h}_n),$$

se nazývá **kvadratická forma** *na* $\mathbb{R}^n$*.*

Jako je nejběžnější bilineární formou skalární součin, tak nejběžnější kvadratická forma je forma vzniklá ze skalárního součinu

$$q(\boldsymbol{h}) = \boldsymbol{h} \cdot \boldsymbol{h} = h_1^2 + h_2^2 + \cdots + h_n^2 = \|\boldsymbol{h}\|^2.$$

Tvar této funkce už sám napovídá důvod, proč se výrazům tohoto typu říká kvadratické formy. Všechny kvadratické formy jsou homogenní funkce 2. stupně. Tzn.

$$q(t\boldsymbol{h}) = t^2 q(\boldsymbol{h})$$

pro všechna $t \in \mathbb{R}$ a $\boldsymbol{h} \in X$. Z našeho hlediska budeme rozlišovat kvadratické formy do tří druhů.

Definice 7.7. *Kvadratická forma $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ na euklidovském prostoru se nazývá*

(i) **pozitivně definitní**, *jestliže existuje $\alpha > 0$ takové, že*

$$q(\boldsymbol{h}) \geq \alpha \|\boldsymbol{h}\|^2, \text{ pro všechna } \boldsymbol{h} \in \mathbb{R}^n;$$

(ii) **negativně definitní**, *jestliže existuje $\alpha > 0$ takové, že*

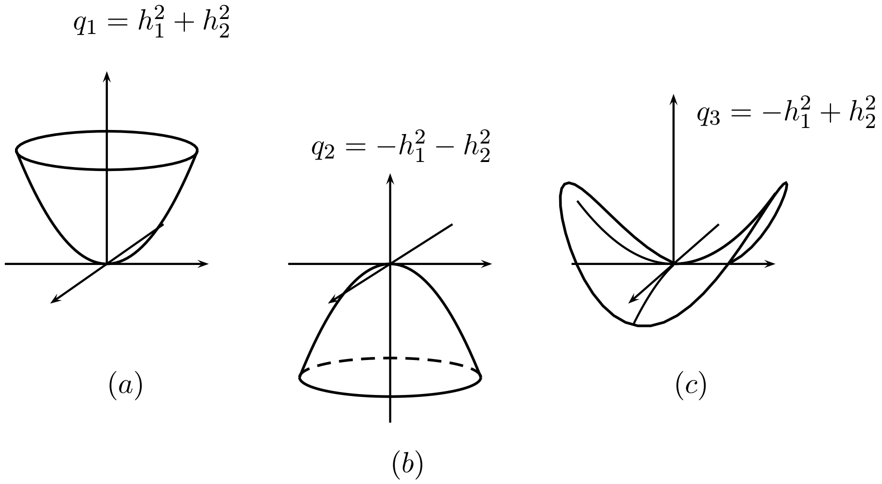
$$q(\boldsymbol{h}) \leq -\alpha \|\boldsymbol{h}\|^2, \text{ pro všechna } \boldsymbol{h} \in \mathbb{R}^n;$$

(iii) **indefinitní**, *jestliže existují $\boldsymbol{h}, \boldsymbol{k} \in \mathbb{R}^n$ taková, že*

$$q(\boldsymbol{h}) > 0 \quad a \quad q(\boldsymbol{k}) < 0.$$

Je ovšem třeba upozornit, že výše vyjmenované tři druhy nepokrývají všechny možnosti, které mohou nastat. Existuje např. kvadratická forma splňující $q(\boldsymbol{h}) \geq 0$ pro všechna $\boldsymbol{h}$ a navíc $q(\boldsymbol{h}_0) = 0$ pro jisté $\boldsymbol{h}_0 \neq 0$. Taková forma nezapadá ani do jedné ze skupin v Definici 7.7.

Na obr.7.3 jsou ukázky grafů kvadratických forem na $\mathbb{R}^2$ odpovídající všem třem typům.



Obr. 7.3.

Forma $q_1(h_1, h_2) = h_1^2 + h_2^2$ je pozitivně definitní ($\alpha = 1$), $q_2(h_1, h_2) = -h_1^2 - h_2^2$ je negativně definitní ($\alpha = -1$) a $q_3(h_1, h_2) = -h_1^2 + h_2^2$ je indefinitní ($q_3(1, 0) < 0$, $q_3(0, 1) > 0$).

Protože v našich příkladech budeme většinou používat kvadratické formy na $\mathbb{R}^2$, podíváme se blíže na to, jak na nich rozpoznat, do které ze skupin (i), (ii) a (iii) patří. Obecná kvadratická forma na $\mathbb{R}^2$ má tvar

$$q(h_1, h_2) = ah_1^2 + 2bh_1h_2 + ch_2^2,$$

pro nějaké konstanty a, b, c. Číslo 2 u prostředního členu je pouze pro naše pohodlí. Můžeme tak snáze poslední rovnici přepsat na tvar

$$q(h_1, h_2) = (h_1, h_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \tag{7.5}$$

Symetrická matice uprostřed jednoznačně zadává formu q. Ukážeme si, jak z ní určíme, o jaký typ kvadratické formy se jedná. Každá kvadratická forma $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ je určena symetrickou maticí řádu n

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

Pro ni platí následující kritéria.

Věta 7.8. *(Sylvestrovo kritérium) Kvadratická forma $q\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ je*

(i) *pozitivně definitní, jestliže všechny hlavní subdeterminanty matice $\mathbf{A}$ jsou kladné, tj.*

$$a_{11} > 0,\ \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0,\ \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} > 0, \dots,\ \det \mathbf{A} > 0;$$

(ii) *negativně definitní, jestliže hlavní subdeterminanty střídají znaménka počínaje záporným,*

$$a_{11} < 0,\ \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0,\ \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} < 0, \dots,\ (-1)^n \det \mathbf{A} > 0;$$

(iii) *indefinitní, jestliže* $\det \mathbf{A} \neq 0$ *a přitom neplatí ani pravidlo (i) ani pravidlo (ii).*

Předchozí věta nezahrnuje případ, kdy $\det \mathbf{A} = 0$. To odpovídá tomu, že ani Definice 7.7 nepokrývala všechny možnosti.

1.3 Kritérium pro extrémy

Nyní jsme už vyzbrojeni dostatečnými poznatky z předešlých sekcí, abychom mohli dokázat větu o nabývání extrému pro funkce více proměnných.

Věta 7.9. *Nechť f je funkce třídy C^2 na otevřené množině G euklidovského prostoru $\mathbb{R}^n$ a nechť $\boldsymbol{x} \in G$ je stacionární bod. Je-li kvadratická forma $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}]$*

(i) *pozitivně definitní, pak je v $\boldsymbol{x}$ ostré lokální minimum,*

(ii) *negativně definitní, pak je v $\boldsymbol{x}$ ostré lokální maximum,*

(iii) *indefinitní, pak v $\boldsymbol{x}$ není lokální extrém ($\boldsymbol{x}$ je tzv. sedlový bod).*

Důkaz. (i) Předpokládejme, že $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}]$ je pozitivně definitní,

$$\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}] \geq \alpha \|\boldsymbol{h}\|^2 \tag{7.6}$$

pro jisté $\alpha > 0$. Druhý diferenciál je určen svou Hessovou maticí, jejíž složky tvoří všechny druhé parciální derivace. Protože f je třídy C^2, jsou tyto parciální derivace spojité. Znamená to, že pro $\boldsymbol{y}$ blízká bodu $\boldsymbol{x}$ se Hessova matice v bodě $\boldsymbol{y}$ příliš neliší od Hessovy matice v bodě $\boldsymbol{x}$. Proto i pro $\mathrm{d}^2 f(\boldsymbol{y})$ bude platit (7.6) s eventuelně trochu pozměněným α. Přesně řečeno: existuje okolí U bodu $\boldsymbol{x}$, že pro každé $\boldsymbol{y} \in U$ platí

$$\mathrm{d}^2 f(\boldsymbol{y})[\boldsymbol{h}, \boldsymbol{h}] \geq \alpha_0 \|\boldsymbol{h}\|^2 \tag{7.7}$$

pro jisté $\alpha_0 > 0$ a všechna $\boldsymbol{h} \in \mathbb{R}^n$. Zvolme teď takové dostatečně malé $\boldsymbol{h}$, aby stále bylo $\boldsymbol{x}+\boldsymbol{h} \in U$. Nyní užijeme Taylorův rozvoj pro funkci $f(\boldsymbol{x})$ do druhého řádu. Podle Věty 6.19 máme

$$f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x}) = \boldsymbol{h} \cdot \operatorname{grad} f(\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{h} \cdot \operatorname{grad})^2 f(\boldsymbol{x} + \vartheta), \tag{7.8}$$

pro jisté $\theta \in (0, 1)$. Dále víme, že $\boldsymbol{x}$ je stacionární bod, a tedy $\boldsymbol{h} \cdot \operatorname{grad} f(\boldsymbol{x}) = 0$. Zbývá vyjádření $(\boldsymbol{h} \cdot \operatorname{grad})^2$. Připomeňme si dva vztahy, které jsme odvodily. Z rovnice (6.24) vidíme, že

$$(\boldsymbol{h} \cdot \operatorname{grad})^2 f = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

A rovnice (6.18) dále dává

$$\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{h}] = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Tím se nám (7.8) zredukuje na tvar

$$f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x}) = \mathrm{d}^2 f(\boldsymbol{x} + \vartheta \boldsymbol{h})[\boldsymbol{h}, \boldsymbol{h}]. \tag{7.9}$$

Protože $\boldsymbol{h}$ bylo zvoleno, aby $\boldsymbol{x}+\vartheta\boldsymbol{h} \in U$ pro všechna $\vartheta \in \langle 0, 1\rangle$, je možné použít odhad (7.7) pro druhý diferenciál. Dostaneme tak

$$f(\boldsymbol{x} + t\boldsymbol{h}) - f(\boldsymbol{x}) \geq \frac{\alpha_0}{2} \|\boldsymbol{h}\|^2.$$

Odtud vidíme, že $f(\boldsymbol{x}+\boldsymbol{h}) > f(\boldsymbol{x})$ pro všechna $\boldsymbol{h}$ nenulová a taková, že $\boldsymbol{x}+\boldsymbol{h} \in U$.

(ii) Pro negativně definitní $\mathrm{d}^2 f(\boldsymbol{x})$ je důkaz úplně stejný, až na to, že v (7.9) je na pravé straně záporný člen. Proto je v $\boldsymbol{x}$ lokální maximum.

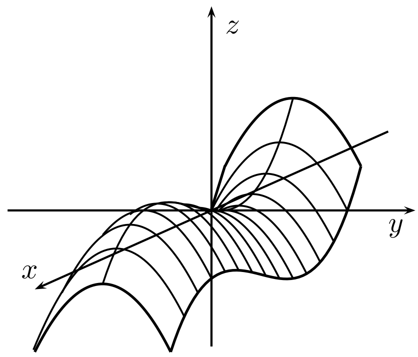
(iii) Zbývá případ, kdy kvadratická forma $\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h},\boldsymbol{h}]$ je indefinitní. Existují tedy vektory $\boldsymbol{h}$ a $\boldsymbol{k}$ takové, že

$$\mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{h},\boldsymbol{h}] > 0 \quad \text{a} \quad \mathrm{d}^2 f(\boldsymbol{x})[\boldsymbol{k},\boldsymbol{k}] < 0.$$

Z (7.9) opět dostáváme, že při posunu z bodu $\boldsymbol{x}$ do bodu $\boldsymbol{x}+\boldsymbol{h}$ hodnota funkce vzroste. Při posunu do bodu $\boldsymbol{x}+\boldsymbol{k}$ hodnota funkce klesne. Odtud plyne, že v $\boldsymbol{x}$ není lokální extrém.
□

Jako ukázku použití Věty 7.9 spočteme následující

Příklad 7.10. Vyšetřete lokální extrémy funkce $z = x(3-x^2) - y^2$.



Obr. 7.4.

Stacionární body musí vyhovovat podmínkám

$$0 = \frac{\partial z}{\partial x} = 3 - 3x^2 \quad \text{a} \quad 0 = \frac{\partial z}{\partial y} = -2y.$$

Ty nám dávají $x = \pm 1$ a $y = 0$. Funkce z má dva stacionární body $\boldsymbol{x}_1 = (1,0)$ a $\boldsymbol{x}_2 = (-1,0)$. Zjistíme nyní, jakého typu je kvadratická forma dz v příslušných bodech. K tomu je třeba spočítat determinanty Hessovy matice. Její složky jsou

$$\frac{\partial^2 z}{\partial x^2} = -6x, \ \frac{\partial^2 z}{\partial y^2} = -2, \ \text{a} \ \frac{\partial^2 z}{\partial x \partial y} = 0.$$

V bodě $\boldsymbol{x}_1$:

$$\det \begin{pmatrix} -6 & 0 \\ 0 & -2 \end{pmatrix} = 12 > 0, \quad \frac{\partial^2 z}{\partial x^2}(\boldsymbol{x}_1) = -6 < 0,$$

a tedy $\mathrm{d}^2 z$ je v $\boldsymbol{x}_1$ negativně definitní. Podle Věty 7.9 je v $\boldsymbol{x}_1$ lokální maximum funkce z.
V bodě $\boldsymbol{x}_2$:

$$\det \begin{pmatrix} 6 & 0 \\ 0 & -2 \end{pmatrix} = -12 < 0,$$

a tedy $\mathrm{d}^2 z$ je indefinitní. Funkce z nemá v $\boldsymbol{x}_2$ lokální extrém, $\boldsymbol{x}_2$ je sedlový bod. Graf funkce z je na obr. 7.4.

2 Vázané extrémy

V této části popíšeme, jak řešit jiný typ úloh na extrémy než jsme doposud viděli. Jedná se o extrémy *vzhledem k množině*. Znamená to, že zkoumáme funkci pouze v bodech dané množiny a mezi nimi hledáme největší nebo nejmenší funkční hodnotu. Metoda, kterou budeme užívat, se obzvláště hodí v případech, kdy příslušná množina je popsána jednou či více rovnicemi. Zatím se ale podíváme na první jednoduchý příklad hledání extrému vzhledem k množině.

Příklad 7.11. Nalezněte minimum a maximum funkce $f(x,y) = \sqrt{3}x - y + 2$ na množině M zadané rovnicí $x^2 + 2x + y^2 = 0$.

Řešení: V tomto jednoduchém případě není nutné vynalézat nějakou speciální metodu řešení. Množina M je vlastně kružnice, neboť původní rovnice se nechá přepsat do tvaru

$$(x+1)^2 + y^2 = 1.$$

Odtud vidíme, že střed je $(-1,0)$ a poloměr 1. Protože M je uzavřená a omezená a zadaná funkce f spojitá, nabývá f svého minima i maxima vzhledem k M podle Věty 4.1. Stojí za povšimnutí, že bez omezení se na množinu M by funkce f neměla žádný extrém, neboť nemá žádné stacionární body.

K cíli teď vedou dvě cesty. Ta první, elegantnější, spočívá v postřehu, že funkce f je lineární. Směr jejího největšího růstu udává $\operatorname{grad} f = (\sqrt{3}, -1)$. Vyjdeme-li od středu kružnice M ve směru gradientu, dorazíme k bodu na M, ve kterém je maximum. Vydáme-li se od středu v opačném směru, musíme dojít do bodu minima. Protože poloměr kružnice je 1, přičtením jednotkového vektoru ve směru gradientu

$$\frac{\operatorname{grad} f}{\|\operatorname{grad} f\|} = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$$

ke středu $(-1,0)$ dostaneme bod maxima

$$\boldsymbol{x}_{\max} = (-1,0) + \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \left(\frac{\sqrt{3}-2}{2}, -\frac{1}{2}\right).$$

Podobně odečtení téhož vektoru od středu získáme bod minima

$$\boldsymbol{x}_{\min} = (-1,0) - \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \left(-\frac{\sqrt{3}+2}{2}, \frac{1}{2}\right).$$

Druhá cesta je početní. Množina M je parametricky popsána

$$\begin{aligned} x &= -1 + \cos t \\ y &= \sin t \end{aligned} \tag{7.10}$$

pro $t \in \langle 0, 2\pi \rangle$. Dosadíme toto vyjádření do $f(x,y)$ a dostaneme funkci proměnné t

$$\sqrt{3}(-1+\cos t) - \sin t + 2 = \sqrt{3}\cos t - \sin t + 2 - \sqrt{3}.$$

Úloha se zredukovala na nalezení minima a maxima posledního výrazu pro $t \in \langle 0, 2\pi \rangle$. Položíme tedy první derivaci rovnou nule,

$$-\sqrt{3}\sin t - \cos t = 0, \quad \text{tj.} \quad \operatorname{tg} t = -\frac{1}{\sqrt{3}}.$$

Tomu odpovídají dvě řešení v $\langle 0, 2\pi\rangle$: $t_1 = \frac{5\pi}{6}$ a $t_2 = \frac{11\pi}{6}$. Dosazením do (7.10) zjistíme, že pro hodnotu parametru t_1 dostáváme bod $\boldsymbol{x}_{\min}$ a pro hodnotu t_2 bod $\boldsymbol{x}_{\max}$.

K právě uvedenému příkladu se vrátíme až budeme mít dokázánu hlavní větu této části a spočítáme jej také pomocí nové metody.

Obecná situace, kterou budeme studovat je tato: nechť $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ je třídy C^1 na euklidovském prostoru. Mějme dále p funkcí $g_1, g_2, \dots, g_p$ opět třídy C^1, které nám zadávají množinu M, a to tak, že v bodech M platí současně následující rovnice

$$\begin{aligned} g_1(\boldsymbol{x}) &= 0, \\ g_2(\boldsymbol{x}) &= 0, \\ &\vdots \\ g_p(\boldsymbol{x}) &= 0. \end{aligned}$$

Matematický zápis množiny M je

$$M = \left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; g_1(\boldsymbol{x}) = 0,\ g_2(\boldsymbol{x}) = 0, \dots,\ g_p(\boldsymbol{x}) = 0\right\} = \bigcap_{i=1}^{p} \left\{\boldsymbol{x} \in \mathbb{R}^n \mid g_i(\boldsymbol{x}) = 0\right\}.$$

Pro představu: $\mathbb{R}^n$ euklidovský prostor dimenze n. Množina bodů

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; g_1(\boldsymbol{x}) = 0\right\}$$

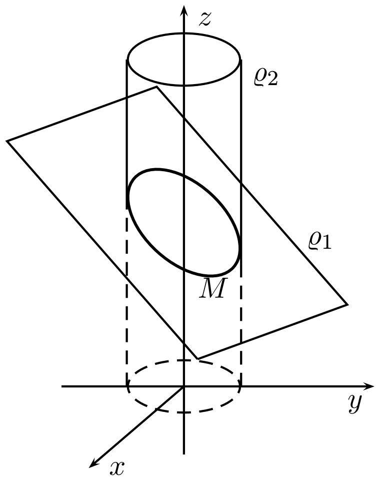
představuje v typických případech nadplochu, jejíž dimenze je $n - 1$. Průnik

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; g_1(\boldsymbol{x}) = 0\right\} \cap \left\{\boldsymbol{x} \in \mathbb{R}^n \mid g_2(\boldsymbol{x}) = 0\right\}$$

je útvar v $\mathbb{R}^n$ mající dimenzi již $n - 2$. Přidáváme-li do průniku další množiny, dimenze výsledného útvaru se sníží vždy o jednu. Množina M je tak $(n - p)$-dimenzionální plocha v n-rozměrném prostoru $\mathbb{R}^n$. Např. při $n = 3$ položme

$$g_1(x, y, z) = x + y + z - 2 \quad \text{a} \quad g_2(x, y, z) = x^2 + y^2 - 1.$$

První udává rovinu ϱ_1 procházející body $(2, 0, 0)$, $(0, 2, 0)$ a $(0, 0, 2)$. Druhá funkce zadává plášť ϱ_2 nekonečného válce, jehož osu tvoří osa z a jehož poloměr podstavy je 1. Každá z podmínek $g_1(x, y, z) = 0$ a $g_2(x, y, z) = 0$ sama o sobě definuje dvourozměrné plochy. Společně pak určují křivku danou jejich průnikem, v našem případě elipsu na obr.7.5.



Obr. 7.5

Slíbená metoda spočívá v následující větě. Její důkaz zde uvádět nebudeme, ale je ho možné nalézt např. v [[3]].

Věta 7.12. *(o Lagrangeových multiplikátorech) Nechť f, $g_1, \dots, g_p$ jsou funkce třídy C^1 na otevřené množině G v euklidovském prostoru $\mathbb{R}^n$, $n > p$. Mějme množinu M zadánu*

$$M = \bigcap_{i=1}^{p} \Big\{ \boldsymbol{x} \in \mathbb{R}^n \mid g_i(\boldsymbol{x}) = 0 \Big\} \subset G.$$

Dále předpokládejme, že vektory $\operatorname{grad} g_1(\boldsymbol{x}), \operatorname{grad} g_2(\boldsymbol{x}), \dots, \operatorname{grad} g_p(\boldsymbol{x})$ *jsou lineárně nezávislé ve všech bodech množiny M. Je-li bod $\boldsymbol{x}_0 \in M$ bodem lokálního extrému funkce f vzhledem k množině M, pak existují taková čísla $\lambda_1, \dots, \lambda_p \in \mathbb{R}$, že bod $\boldsymbol{x}_0$ je stacionární bod tzv. Lagrangeovy funkce*

$$L = f(\boldsymbol{x}) - \sum_{i=1}^{p} \lambda_i g_i(\boldsymbol{x}). \tag{7.11}$$

Poznámka 7.13. Rovnicím $g_1(\boldsymbol{x}) = 0,\ g_2(\boldsymbol{x}) = 0, \dots,\ g_p(\boldsymbol{x}) = 0$ se někdy říká vazebné podmínky a extrému vzhledem k množině „vázaný extrém“. Čísla $\lambda_1, \dots, \lambda_p$ jsou označována názvem *Lagrangeovy multiplikátory* a Věta 7.12 bývá v literatuře uváděna kromě názvu věta o Lagrangeových multiplikátorech taká pod názvem věta o vázaném extrému. Podmínka pro stacionární bod Lagrangeovy funkce rozepsaná do složek představuje n rovnic

$$\begin{aligned} \frac{\partial f}{\partial x_1} - \lambda_1 \frac{\partial g_1}{\partial x_1} - \dots - \lambda_p \frac{\partial g_p}{\partial x_1} &= 0, \\ \vdots \qquad\qquad & \vdots \\ \frac{\partial f}{\partial x_n} - \lambda_1 \frac{\partial g_1}{\partial x_n} - \dots - \lambda_p \frac{\partial g_p}{\partial x_n} &= 0. \end{aligned} \tag{7.12}$$

Neznámých je však $n+p$: $x_1, \dots, x_n, \lambda_1, \dots, \lambda_p$. Proto k výše uvedené soustavě přidáme p vazebných podmínek $g_1(\boldsymbol{x}) = 0$, $g_2(\boldsymbol{x}) = 0, \dots,$ $g_p(\boldsymbol{x}) = 0$ a vyrovnáme tak počet rovnic i neznámých. Tuto soustavu pak v konkrétních případech řešíme, abychom našli body podezřelé z extrému.

Podmínka, aby gradienty $\operatorname{grad} g_1(\boldsymbol{x}), \operatorname{grad} g_2(\boldsymbol{x}), \dots, \operatorname{grad} g_p(\boldsymbol{x})$ byly lineárně nezávislé vektory je ekvivalentně vyjádřena požadavkem, aby matice typu $p \times n$

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial g_p}{\partial x_1} & \cdots & \dfrac{\partial g_p}{\partial x_n} \end{pmatrix}$$

měla hodnost p ve všech bodech množiny M. V tomto tvaru je formulován předpoklad ve Věty 7.4 v mnoha učebnicích.

Vraťme se ještě k podmínce stacionarity bodu $\boldsymbol{x}_0$ pro Lagrangeovu funkci $L(\boldsymbol{x})$. Kromě podrobného rozpisu do složek, který je v (7.12), můžeme podmínku stacionarity vyjádřit následovně:

$$\operatorname{grad} L(\boldsymbol{x}_0) = 0, \text{ tj. } \operatorname{grad} f(\boldsymbol{x}_0) = \lambda_1 \operatorname{grad} g_1(\boldsymbol{x}_0) + \cdots + \lambda_p \operatorname{grad} g_p(\boldsymbol{x}_0).$$

Vidíme, že v bodě lokálního extrému funkce f vzhledem k množině M je gradient funkce f lineární kombinací gradientů funkcí z vazebných podmínek. Takto vyjádřený závěr Věty 7.12 o Lagrangeových multiplikátorech má názornou geometrickou interpretaci. Podívejme se na ni blíže v jednoduchém případě roviny $\mathbb{R}^2$. Představme si, že množina M je graf funkce $y = g(x)$. Přepíšeme si tuto rovnici do tvaru vazební podmínky

$$g_1(x, y) = 0, \quad \text{kde } g_1(x, y) = g(x) - y. \tag{7.13}$$

V části o geometrickém významu gradientu jsme zjistili, že $\operatorname{grad} g_1$ je normálový vektor ke grafu funkce $y = g(x)$:

$$\boldsymbol{n} = \operatorname{grad} g_1 = \Big(\frac{\partial g_1}{\partial x}, \frac{\partial g_1}{\partial y}\Big) = (g'(x), -1). \tag{7.14}$$

Předpokládejme, že v bodě $(x, y) \in M$ nabývá funkce $f(x, y)$ svého extrému vzhledem k M. Vyšetřujeme-li naší funkci pouze v bodech množiny M, sledujeme chování složené funkce $f(x, g(x))$, která závisí už jen na jedné proměnné. Má-li tato funkce v bodě x_0 extrém, je její derivace nulová:

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x, g(x)) \Big|_{x=x_0} = 0.$$

Podle vzorce o derivaci složené funkce (Věta 6.1(i) nebo Důsledek 6.5) dostaneme

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} g'(x_0) = 0.$$

To ovšem není nic jiného než skalární součin dvou vektorů

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \bigl(1, g'(x_0)\bigr) = 0. \tag{7.15}$$

První vektor je $\operatorname{grad} f$. Ve druhém vektoru není těžké rozpoznat tečný vektor ke grafu funkce g v bodě x_0: Tento vektor je totiž zjevně kolmý na vektor $\boldsymbol{n}$ v (7.14). Protože tam se jednalo o normálový vektor, je vektor na něj kolmý nutně tečný vektor. Poslední rovnice (7.15) pak říká, že $\operatorname{grad} f$ musí mít směr normály:

$$\operatorname{grad} f = \lambda \boldsymbol{n} = \lambda \operatorname{grad} g_1.$$

To je ovšem přesně tvrzení Věty 7.12 v tomto speciálním případě.

Uvedeme si dva ilustrační příklady na použití Věty o vázaných extrémech. V prvním se vrátíme k Příkladu 7.11 a vyřešíme jej metodou Lagrangeových multiplikátorů.

Příklad 7.14. Zjistěte extrém funkce $\sqrt{3}x - y + 2$ za podmínky $x^2 + 2x + y^2 = 0$.

Máme funkci $f(x, y) = \sqrt{3}x - y + 2$ a vazebnou podmínku $g(x, y) = x^2 + 2x + y^2$. Sestavíme Lagrangeovu funkci (7.11):

$$L(x, y) = f(x, y) - \lambda g(x, y) = \sqrt{3}x - y + 2 - \lambda(x^2 + 2x + y^2).$$

Nyní si vypíšeme podmínky pro stacionární bod funkce $L(x, y)$ a přidáme k nim vazebnou podmínku:

$$\begin{aligned} \frac{\partial L}{\partial x} - \lambda \frac{\partial g}{\partial x} &= 0 & & & \sqrt{3} - 2\lambda(x+1) &= 0, \\ \frac{\partial L}{\partial y} - \lambda \frac{\partial g}{\partial y} &= 0 & &\text{tj.} & -1 - 2\lambda y &= 0, \\ g(x, y) &= 0 & & & x^2 + 2x + y^2 &= 0. \end{aligned}$$

Vyloučíme-li λ z prvních dvou rovnic, dostaneme vztah mezi x a y: $y = -(x+1)/\sqrt{3}$. Ten dosadíme do třetí rovnice. Vzniklou kvadratickou rovnici

$$x^2 + 2x + \frac{(x+1)^2}{3} = 0$$

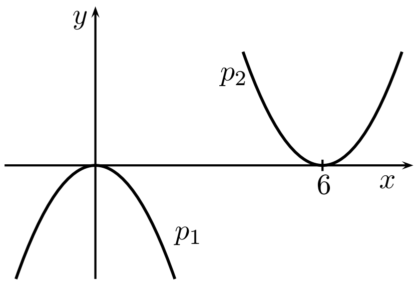
vyřešíme. Dostaneme $x_1 = (\sqrt{3} - 2)/2$ a $x_2 = -(\sqrt{3} + 2)/2$. Dopočtením příslušných y-ových souřadnic získáme dva body

$$\left(\frac{\sqrt{3} - 2}{2}, -\frac{1}{2}\right), \quad \left(-\frac{\sqrt{3} + 2}{2}, \frac{1}{2}\right).$$

Jak se snadno přesvědčíme porovnáním funkčních hodnot, první bod je bod maxima a druhý je bod minima. Máme tak potřetí ověřeno řešení z příkladu 7.11.

Pokud je vazebných podmínek více a nejsou lineární, obvykle řešení vede k rovnicím vyššího stupně než dva nebo dokonce k rovnicím transcendentním.

Příklad 7.15. Jaká je vzdálenost parabol $p_1 \colon y = -x^2$ a $p_2 \colon y = (x - 6)^2$? Viz obr.7.6.



Obr. 7.6

Funkce, kterou budeme minimalizovat je funkce f vzdálenosti dvou bodů (x_1, y_1) a (x_2, y_2) v rovině, z nichž jeden leží na p_1 a druhý na p_2. Funkce f závisí na čtyřech proměnných x_1, y_1, x_2, y_2. Máme tak

$$\begin{aligned} f(x_1, y_1, x_2, y_2) &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \\ g_1(x_1, y_1) &= y_1 + x_1^2, \\ g_2(x_2, y_2) &= y_2 - (x_2 - 6)^2. \end{aligned}$$

Protože minimum funkce f se nabývá v tom samém bodě jako minimum f^2, můžeme pro zjednodušení výpočtů uvažovat místo vzdálenosti její druhou mocninu, kterou označíme opět symbolem f

$$f(x_1, y_1, x_2, y_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

Vazebné podmínky jsou $g_1(x_1, y_1) = 0$ a $g_2(x_2, y_2) = 0$ a Lagrangeova funkce

$$L = f(x_1, y_1, x_2, y_2) - \lambda_1 g_1(x_1, y_1) - \lambda_2 g_2(x_2, y_2).$$

Zderivujeme tuto funkci podle x_1, x_2, y_1 a y_2 a sestavíme následující soustavu šesti rovnic

$$\begin{aligned} \frac{\partial f}{\partial x_1} - \lambda_1 \frac{\partial g_1}{\partial x_1} - \lambda_2 \frac{\partial g_2}{\partial x_1} &= 0 & \qquad\qquad & & x_1 - x_2 &= \lambda_1 x_1, \\ \frac{\partial f}{\partial x_2} - \lambda_1 \frac{\partial g_1}{\partial x_2} - \lambda_2 \frac{\partial g_2}{\partial x_2} &= 0 & & & x_2 - x_1 &= -\lambda_2 (x_2 - 6), \\ \frac{\partial f}{\partial y_1} - \lambda_1 \frac{\partial g_1}{\partial y_1} - \lambda_2 \frac{\partial g_2}{\partial y_1} &= 0 & & \text{tj.} & y_1 - y_2 &= \frac{\lambda_1}{2}, \\ \frac{\partial f}{\partial y_2} - \lambda_1 \frac{\partial g_1}{\partial y_1} - \lambda_2 \frac{\partial g_2}{\partial y_2} &= 0 & & & y_2 - y_1 &= \frac{\lambda_2}{2}, \\ g_1(x_1, y_1) &= 0 & & & y_1 &= -x_1^2, \\ g_2(x_2, y_2) &= 0 & & & y_2 &= (x_2 - 6)^2. \end{aligned}$$

Ze třetí a čtvrté rovnice ihned plyne, že $\lambda_1 = -\lambda_2$. Když toho využijeme v první a v druhé rovnici, dostaneme $x_1 + x_2 = 6$. Z tohoto vztahu plyne, že $x_1 = x_2 - 6$, a proto nám pátá a šestá rovnice dávají $y_1 = -y_2$. Dosadíme-li za y_2 do třetí (nebo čtvrté) rovnice, dostaneme $y_1 = \lambda_1/4$. Zatím jsme tedy vyvodili následující

$$y_1 = -y_2 = \frac{\lambda_1}{4} = -\frac{\lambda_2}{4} \quad \text{a} \quad x_1 + x_2 = 6. \tag{7.16}$$

Budeme teď vylučovat z původních rovnic všechny proměnné kromě x_1 a y_1, tzn. za ostatní budeme dosazovat z (7.16). Tím se nám šest rovnic zredukuje na tyto dvě

$$x_1 - 3 = 2x_1y_1, \qquad y_1 = x_1^2.$$

Vyloučením y_1 dostaneme rovnici třetího stupně

$$2x_1^3 + x_1 - 3 = 0.$$

Tato rovnice má pouze jediné reálné řešení $x_1 = 1$. Odtud pomocí (7.16) dopočítáme všechny ostatní

$$x_1 = 1, \quad y_1 = -1, \quad x_2 = 5, \quad y_2 = 1.$$

Body, ve kterých se realizuje minimální vzdálenost jsou $(1,-1) \in p_1$ a $(5,1) \in p_2$. Závěr: vzdálenost mezi parabolami p_1 a p_2 je vzdálenost bodů $(1,-1)$ a $(5,1)$:

$$\|(1,-1) - (5,1)\| = \sqrt{20}.$$

Stejně jako Věta 7.9 udávala postačující podmínky pro to, aby stacionární bod byl bodem lokálního minima či maxima, máme i v případě vázaných extrémů podobné kritérium. Z toho, jak se chová jistá kvadratická forma můžeme zjistit, zda je v příslušném bodě lokální extrém.

Věta 7.16. *Nechť $f, g_1, \dots, g_p$ jsou funkce třídy C^2 na otevřené množině G v euklidovském prostoru $\mathbb{R}^n$, $n > p$. Mějme dále množinu M zadánu*

$$M = \bigcap_{i=1}^{p} \Big\{ \boldsymbol{x} \in \mathbb{R}^n \mid g_i(\boldsymbol{x}) = 0 \Big\} \subset G.$$

Dále předpokládejme, že vektory $\operatorname{grad} g_1, \dots, \operatorname{grad} g_p$ *jsou lineárně nezávislé ve všech bodech množiny M. Nechť $\boldsymbol{x}_0 \in M$ je bod s následujícími vlastnostmi*

(i) *existují čísla $\lambda_1, \dots, \lambda_p \in \mathbb{R}$, že Lagrangeova funkce*

$$L(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda_1 g_1(\boldsymbol{x}) - \dots - \lambda_p g_p(\boldsymbol{x})$$

má v bodě $\boldsymbol{x}_0$ stacionární bod, $\operatorname{grad} L(\boldsymbol{x}_0) = 0$,

(ii) *kvadratická forma $\mathrm{d}^2 L(\boldsymbol{x}_0)[\boldsymbol{h}, \boldsymbol{h}]$ uvažovaná pouze pro vektory $\boldsymbol{h}$ z množiny*

$$T = \bigcap_{i=1}^{p} \Big\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \perp \operatorname{grad} g_i(\boldsymbol{x}_0) \Big\}$$

je pozitivně definitní (resp. negativně definitní, resp. indefinitní).

Pak funkce f má v bodě $\boldsymbol{x}_0$ ostré lokální minimum (resp. maximum, resp. extrém nenastává) vzhledem k množině M.

Množina T z bodu (ii), na které uvažujeme kvadratickou formu $\mathrm{d}^2 f(\boldsymbol{x}_0)$, je tečná „nadrovina“ k množině M v bodě $\boldsymbol{x}_0$. Uvozovky u slova „nadrovina“ znamenají, že její dimenze nemusí být jen o jednu menší než je dimenze prostoru. Dimenze T závisí na poču vazebných podmínek a obecně T má dimenzi $n-p$. Zmenšení oboru pro testování definitnosti formy $\mathrm{d}^2 f(\boldsymbol{x}_0)$ je podstatné. Může se totiž stát, že kvadratická forma je indefinitní na $\mathbb{R}^n$, ale její zúžení na T je např. pozitivně definitní. Uvidíme to ostatně v následujícím příkladu.

Příklad 7.17. Uvažujme funkci $f(x,y) = x^2 - y^2$ na množině

$$M = \{(x,y) \in \mathbb{R}^2 \mid y + e^{-x^2} - 1 = 0\}.$$

Má zde funkce f nějaký lokální extrém?

Vazbová podmínka je zde jedna

$$g(x,y) = 0, \text{ kde } g(x,y) = y + e^{-x^2} - 1.$$

Zjistíme nejprve, jaké body vyhovují podmínce (i) Věty 7.16. Lagrangeova funkce je

$$L(x,y) = f(x,y) - \lambda g(x,y).$$

Příslušné rovnice pak mají tvar

$$\begin{aligned} \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} &= 0 & & & 2x &= -2\lambda x e^{-x^2}, \\ \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} &= 0 & &\text{tj.} & -2y &= \lambda, \\ g(x,y) &= 0 & & & 0 &= y + e^{-x^2} - 1. \end{aligned}$$

Řešení rozdělíme na dva případy: $\lambda = 0$ a $\lambda \neq 0$. V prvním případě dostaneme řešení $x = y = \lambda = 0$. Pro $\lambda \neq 0$ máme z prvních dvou rovnic

$$e^{-x^2} = -\frac{1}{\lambda}, \quad y = -\frac{\lambda}{2}.$$

Dosazením do třetí rovnice a po úpravě získáme kvadratickou rovnici

$$\lambda^2 + 2\lambda + 2 = 0,$$

která nemá reálné řešení. Našli jsme tak pouze jediný bod vyhovující podmínce (i), a to $\boldsymbol{x}_0 = (0,0)$.

Postoupíme teď k bodu (ii). Druhý diferenciál $\mathrm{d}^2 L(\boldsymbol{x}_0)$ s $\lambda = 0$ je dán Hessovou maticí

$$\mathrm{d}^2 L(\boldsymbol{x}_0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}. \tag{7.17}$$

Zjistíme, jaká je kvadratická forma daná maticí (7.17). Vyšetřovat ji ovšem budeme pouze pro taková $\boldsymbol{h}$ splňující požadavky z bodu (ii):

$$\boldsymbol{h} \perp \operatorname{grad} g(\boldsymbol{x}_0), \text{ tj. } \boldsymbol{h} \perp (0,1).$$

To jsou pouze vektory tvaru $\boldsymbol{h} = (h, 0)$, $h \in \mathbb{R}$. Takže

$$\mathrm{d}^2 L(\boldsymbol{x}_0)[\boldsymbol{h}, \boldsymbol{h}] = (h, 0) \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} h \\ 0 \end{pmatrix} = 2h^2 = 2\|\boldsymbol{h}\|^2.$$

Vidíme, že tato forma je pozitivně definitní. Z Věty 7.16 plyne, že bod $\boldsymbol{x}_0 = (0, 0)$ je ostré lokální minimum funkce f na množině M.

Pozastavme se u tohoto příkladu ještě chvilku. Kvadratická forma daná maticí (7.17) uvažovaná na celém $\mathbb{R}^2$ je indefinitní, neboť její první subdeterminant je kladný a druhý záporný. Pokud ji ovšem testujeme pouze pro vektory s druhou souřadnicí nulovou, stává se pozitivně definitní.

3 Nejmenší a největší hodnota funkce

Ne každá množina se nechá popsat rovnicí nebo soustavou rovnic. Příklad takové množiny je čtverec $\langle -1, 1\rangle \times \langle -1, 1\rangle$ v $\mathbb{R}^2$ nebo část rotačního paraboloidu

$$P = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq z \leq 1\},$$

viz obr. 2.5(a) a pod. Zjištění největší a nejmenší hodnoty spojité funkce na takové množině se skládá ze dvou kroků. Nejprve vyšetříme funkci na vnitřku množiny a posléze na její hranici. Pro vnitřek užijeme metodu lokálních extrémů a pro hranici metodu extrémů vázaných.

Příklad 7.18. Zjistíme maximální a minimální hodnotu funkce

$$f(x, y) = x^2 + y^2 - 6x - 4y + 11$$

na množině

$$M = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 - 4x \leq 5\}.$$

Množina M je kruh se středem $(2, 0)$ a poloměrem 3, neboť nerovnost pro M je možné přepsat jako $(x-2)^2 + y^2 \leq 9$. Podíváme se nejprve na stacionární body zadané funkce.

$$0 = \frac{\partial f}{\partial x} = 2x - 6, \quad 0 = \frac{\partial f}{\partial y} = 2y - 4.$$

Tato soustava má jediné řešení $\boldsymbol{x}_0 = (3, 2)$. Bod $\boldsymbol{x}_0$ leží v M, je tedy jedním z kandidátů pro extrém funkce f.

Podíváme se nyní na hranici ∂M. Ta se skládá z kružnice o rovnici

$$0 = g(x, y) = x^2 - 4x + y^2 - 5.$$

Vyšetříme tedy vázaný extrém funkce f. Podle Věty 7.9 musí v bodě extrému platit, že Lagrangeova funkce $L(x, y) = f(x, y) - \lambda g(x, y)$ má stacionární bod. Tedy

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0, \quad \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0, \quad g(x, y) = 0$$

pro jisté $\lambda \in \mathbb{R}$. Výpočtem dostáváme soustavu

$$\begin{aligned} 2x - 6 &= \lambda(2x-4) & & & x &= \frac{3-2\lambda}{1-\lambda} = 2 + \frac{1}{1-\lambda} \\ 2y - 4 &= \lambda\, 2y & &\text{tj.} & y &= \frac{2}{1-\lambda} \\ x^2 - 4x + y^2 &= 5 & & & (x-2)^2 + y^2 &= 9. \end{aligned}$$

Z prvních dvou rovnic plyne, že $y = 2(x-2)$. Dosadíme-li za x do třetí, máme

$$(x-2)^2 + 4(x-2)^2 = 9.$$

Tato rovnice má dvě řešení $= 2 \pm \dfrac{3}{\sqrt{5}}$. Volbě znaménka minus odpovídá výsledný bod

$$\boldsymbol{x}_1 = \left(2 - \frac{3}{\sqrt{5}}, -\frac{6}{\sqrt{5}}\right).$$

Hodnota funkce v tomto bodě je $f(\boldsymbol{x}_1) = 12 + 30/\sqrt{5}$. Pro druhou volbu znaménka máme bod

$$\boldsymbol{x}_2 = \left(2 + \frac{3}{\sqrt{5}}, \frac{6}{\sqrt{5}}\right).$$

Hodnota f v tomto bodě je $f(\boldsymbol{x}_2) = 12 - 30/\sqrt{5}$. Hodnota v $\boldsymbol{x}_0$ je $f(\boldsymbol{x}_0) = -2$. Porovnáním těchto hodnot jsme vedeni k následujícímu závěru:

$$\max_M f = f(\boldsymbol{x}_1) = 12 + \frac{30}{\sqrt{5}}, \quad \min_M f = f(\boldsymbol{x}_0) = -2.$$

Pokud množina, na které vyšetřujeme spojitou funkci f, je uzavřená a omezená, vždy největší a nejmenší hodnota existuje (viz Větu 4.1). Nesplňuje-li množina M tyto požadavky, může se stát, že funkce f nenabývá žádný extrém na M. Uvažujme např. funkci

$$f(x, y) = x + y$$

na otevřeném čtverci $Q = (0, 1) \times (0, 1)$. To je sice omezená množina, nikoli však uzavřená. Je jasné, že kdyby se jednalo o uzavřený čtverec, tak minimum funkce f je v bodě $(0, 0)$ a maximum v bodě $(1, 1)$. Tyto body ale do Q nepatří. A v žádném jiném funkce f extrém nenabývá, neboť žádný bod z Q není stacionárním bodem.

Jiný příklad je s uzavřenou, ale neomezenou množinou

$$M = \{(x, y) \in \mathbb{R}^2 \mid x = y\}.$$

Je to vlastně přímka procházející počátkem. Funkce $f(x, y) = x + y$ je na M neomezená shora i zdola, a tak nenabývá ani největší ani nejmenší hodnoty.

4 Cvičení

1. U následujících funkcí zjistěte body lokálních extrémů.

   a) $z = x^3y^2(6 - x - y)$,
   b) $z = x^3 + y^3 + 9xy + 27$,
   c) $z = xe^{y+x\sin x}$,
   d) $z = x^3 + y^3 - 3axy$,
   e) $z = \sqrt{(a-x)(a-y)(x+y-a)}$,
   f) $z = e^{-x^2-y^2}(ax^2 + by^2)$,

2. Určete největší a nejmenší hodnoty daných funkcí na předepsaných množinách.

   a) $z = e^{xy}$ na množině $x + y = 1$,
   b) $z = \frac{1}{x} + \frac{1}{y}$ na množině $x + y = 2a$, $a > 0$, $x \neq 0$, $y \neq 0$,
   c) $z = xy$ na množině $x^2 + y^2 = 2a^2$,
   d) $u = xyz$ na množině $x + y + z = 5$, $xy + yz + zx = 8$,
   e) $z = x - 2y - 3$ na množině $0 \leq x, y \leq 1$, $0 \leq x + y \leq 1$,
   f) $z = x^2 - 3y^2 - x + 18y - 4$ na množině $0 \leq x \leq y \leq 4$,
   g) $z = (x - y^2)(x-1)^{\frac{2}{3}}$ na množině $y^2 \leq x \leq 2$,
   h) $z = x^2 + 2xy - 3y^2 + y$ na množině $0 \leq x, y \leq 1$, $0 \leq x + y \leq 1$,
   ch) $z = x^2 - xy + y^2$ na množině $|x| + |y| \leq 1$,
   i) $z = x^2 - y^2$ na množině $x^2 + y^2 \leq 4$,
   j) $z = x^2 + 2xy + 4x + 8y$ na množině $\langle 0, 1\rangle \times \langle 0, 2\rangle$,
   k) $z = \sin x + \sin y + \sin(x + y)$ na množině $\langle 0, \pi/2\rangle \times \langle 0, \pi/2\rangle$,
   l) $u = x + y + z$ na množině $x^2 + y^2 \leq z \leq 1$.

3. V rovině $2x - z = 0$ nalezněte bod, pro nějž je součet čtverců vzdáleností od bodů $(1, 1, 1)$ a $(2, 3, 4)$ co nejmenší.

4. Mějme n bodů v prostoru, $A_1 = (x_1, y_1, z_1), \dots, A_n = (x_n, y_n, z_n)$. Určete takový bod $P = (x, y, z)$, pro nějž je součet druhých mocnin vzdáleností k jednotlivým $A_1, \dots, A_n$ co nejmenší.

5. Nechť jsou dány $A = (4, 0, 4)$, $B = (4, 4, 4)$ a $C = (4, 4, 0)$. Na povrchu koule se středem v počátku a poloměrem 2 nalezněte bod D tak, aby objem čtyřstěnu $ABCD$ byl a) co největší, b) co nejmenší.

6. Jaké rozměry má mít kvádr daného objemu, aby měl minimální povrch? Co v případě maximálního povrchu?

7. Jakou největší hodnotu může mít součin tří nezáporných čísel, je-li jejich součet a? Zobecněte pro součin n činitelů.

8. Pomocí předchozího cvičení dokažte nerovnost mezi geometrickým a aritmetickým průměrem
$$(x_1 \cdots x_n)^{\frac{1}{n}} \le \frac{x_1 + \cdots + x_n}{n}.$$

9. Rozložte kladné číslo x na součet kladných sčítanců $x = x_1 + x_2 + x_3$ tak, aby hodnota výrazu $x_1^n x_2^m x_3^p$ byla maximální ($n, m, p \in \mathbb{N}$).

10. Dokažte nerovnost
$$\left(\frac{x+y}{2}\right)^n \le \frac{x^n + y^n}{2},$$
kde $n \in \mathbb{N}$, $x \ge 0$, $y \ge 0$. (Návod: určete minimum funkce $\frac{1}{2}(x^n + y^n)$ za podmínky $x + y = a$).

11. Rozložte dané číslo $x > 0$ na součin n kladných činitelů $x = x_1 \cdots x_n$ tak, aby součet jejich převrácených hodnot byl minimální.

12. V množině všech elips mající součet poloos roven $2L$ nalezněte elipsu s největším obsahem.

13. Určete trojúhelník daného obvodu $2p$, který při rotaci kolem jedné ze svých stran vytvoří těleso maximálního objemu.

14. Nechť K je rotační kužel výšky h a s poloměrem podstavy R. Jaké rozměry musí mít kvádr vepsaný do K, aby měl maximální objem?

15. Nalezněte vzdálenost paraboly $y = x^2$ od přímky $y = x - 2$.

16. Na elipse $\dfrac{x^2}{4} + \dfrac{y^2}{9} = 1$ nalezněte body, které mají největší a nejmenší vzdálenost od přímky $3x + y - 9 = 0$.

17. Jaká je největší vzdálenost plochy $2x^2 + 3y^2 + 2z^2 + 2xy = 6$ od roviny xy?

18. Na elipse jsou dva body A a B. Najděte třetí bod C na téže elipse tak, aby trojúhelník ABC měl největší obsah.

19. $*$ Planeta A obíhá po elipse $\dfrac{(x-3)^2}{4} + (y+1)^2 = 1$. Její poloha v čase t je popsána
$$x = 3 + 2\cos t$$
$$y = -1 + \sin t.$$
V čase $t = 0$ je vypuštěna sonda z bodu $(0, 0)$ a pohybuje se rovnoměrně po přímce $2\sqrt{3}y - x = 0$ směrem do 1. kvadrantu. Určete, jaká musí být rychlost v sondy, aby se míjela s planetou A v co nejmenší vzdálenosti.

20. $*$ Cestovatel se ocitne bez zdroje vody. Podle mapy zjistil, že je v bodě $C = (0, 2\sqrt{5})$. Má dvě množnosti: řeka v nížině, jejíž tok je zakreslen křivkou $x^2 - y^2 = 1$, $x > 0$, a horské jezero, které zabírá na mapě oblast $(x+2)^2 + y \le \dfrac{1}{2} + 2\sqrt{5}$. Cestou dolů k řece se může pohybovat třikrát rychleji než cestou nahoru k jezeru. Kam se cestovatel vydá, aby došel k vodě co nejdříve?

Výsledky

1.a) $(3,2)$ – max; b) $(-3,-3)$ – max; c) nemá extrém d) (a,a) – max pro $a<0$ a min pro $a>0$; e) $(0,0)$ – max; f) $(0,0)$ – min, pro $a>b$ je $(\pm1,0)$ – max, pro $a<b$ je $(0,\pm1)$ – max; 2.a) $(\frac{1}{2},\frac{1}{2})$ – max; b) (a,a) – min; c) $(\pm a,\pm a)$ – max, $(\pm a,\mp a)$ – min; d) max: $(2,2,1),(2,1,2),(1,2,2)$, min: $(\frac{4}{3},\frac{4}{3},\frac{7}{3}),(\frac{4}{3},\frac{7}{3},\frac{4}{3}),(\frac{7}{3},\frac{4}{3},\frac{4}{3})$; e) $(0,1)-$min, $(1,0)-$max; f) $(4,4)-$max, $(0,0)-$min; g) $(2,0)-$max, min: celá hranice; h) $(\frac{7}{8},\frac{1}{8})-$max, $(0,1)-$min; ch) $(0,0)$ – min, $(\pm1,0),(0,\pm1)$ – max; i) $(0,\pm2)$ – min, $(\pm2,0)$ – max; j) $(0,0)$ – min, $(1,2)$ – max; k) $(\frac{\pi}{3},\frac{\pi}{3})$ – max, $(0,0)$ – min; l) $(-\frac{1}{2},-\frac{1}{2},\frac{1}{2})$ – min, $(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},1)$ – max; 3. $(\frac{9}{10},2,\frac{9}{5})$; 4. $P=\frac{1}{n}(\sum x_i,\sum y_i,\sum z_i)$; 5.a) $(-2,0,0)$, b) $(2,0,0)$; 6. čtverec, maximum neexistuje; 7. $(a/3)^3$, $(a/n)^n$; 9. $x_1=\frac{mx}{m+n+p}$, $x_2=\frac{nx}{m+n+p}$, $x_3=\frac{px}{m+n+p}$; 11. $x_i=x^{\frac{1}{n}}$; 12. kruh s poloměrem L; 13. $a=b=\frac{3}{4}p$, $c=\frac{1}{2}p$; 14. výška kvádru je $h/3$; 15. $\frac{7}{4\sqrt{2}}$; 16. $(\frac{4}{\sqrt{5}},\frac{3}{\sqrt{5}}),(-\frac{4}{\sqrt{5}},-\frac{3}{\sqrt{5}})$; 17. $\sqrt{3}$; 18. C leží na ose úsečky AB; 19. $v=\frac{27\sqrt{3}-6}{4\pi\sqrt{13}}$; 20. vzdálenost k jezeru je $\frac{1}{2}\sqrt{5}$ a k řece $\sqrt{11}$ — vydá se k řece.