10 Lokální a globální extrémy

Definice 10.1 (lokálních extrémů)**.** Řekneme, že funkce $f : \mathbb{R}^n \to \mathbb{R}$ nabývá v bodě $X_0 \in \mathbb{R}^n$ *lokálního maxima* (resp. *minima*), jestliže existuje okolí $O(X_0)$ takové, že $O(X_0) \subseteq D(f)$ a pro každé $X \in O(X_0)$ platí $f(X) \leq f(X_0)$ (resp. $f(X) \geq f(X_0)$). Jsou-li tyto nerovnosti pro $X \neq X_0$ ostré, mluvíme o *ostrém* lokálním maximu (resp. minimu). (Ostrá) lokální maxima a minima nazýváme souhrnně *(ostré) lokální extrémy*.

Příklad 10.2. a) Funkce $f(x, y) = \sqrt{x^2 + y^2}$ má v bodě $[0, 0]$ ostré lokální minimum, protože $f(x, y) > 0 = f(0, 0)$ pro $\forall [x, y] \neq [0, 0]$. b) Funkce $f(x, y) = \begin{cases} 1 \text{ pro } [x, y] = [0, 0] \\ 0 \text{ pro } [x, y] \neq [0, 0] \end{cases}$ má v bodě $[0, 0]$ ostré lokální maximum. Příklad a) ukazuje, že funkce nemusí být v bodě lokálního extrému diferencovatelná, příklad b) ukazuje, že nemusí být ani spojitá.

Definice 10.3 (stacionárního bodu)**.** Řekneme, že bod $S \in \mathbb{R}^n$ je *stacionárním bodem funkce* $f : \mathbb{R}^n \to \mathbb{R}$, jestliže platí $\operatorname{grad} f(S) = \vec{o} = (0, 0, \ldots, 0)$.

Věta 10.4. *Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ má v bodě $X^* \in \mathbb{R}^n$ lokální extrém a nechť v tomto bodě existují všechny parciální derivace funkce f. Pak X^* je stacionárním bodem funkce f.*

Důkaz. Kdyby $f'_{x_i}(X^*) > 0$, tj.

$$\lim_{h \to 0} \frac{1}{h} \Big[f(x_1^*, \ldots, x_{i-1}^*, x_i^* + h, x_{i+1}^*, \ldots, x_n^*) - f(x_1^*, \ldots, x_{i-1}^*, x_i^*, x_{i+1}^*, \ldots, x_n^*) \Big] > 0,$$

pak by existovalo $\delta > 0$ takové, že pro každé $h \in (-\delta, \delta)$ je také

$$\frac{1}{h} \Big[f(x_1^*, \ldots, x_{i-1}^*, x_i^* + h, x_{i+1}^*, \ldots, x_n^*) - f(x_1^*, \ldots, x_{i-1}^*, x_i^*, x_{i+1}^*, \ldots, x_n^*) \Big] > 0,$$

viz vlastnosti limity. To znamená, že pro kladná h z $(-\delta, \delta)$ je

$$f(x_1^*, \ldots, x_{i-1}^*, x_i^* + h, x_{i+1}^*, \ldots, x_n^*) > f(x_1^*, \ldots, x_{i-1}^*, x_i^*, x_{i+1}^*, \ldots, x_n^*),$$

a pro záporná h je

$$f(x_1^*, \ldots, x_{i-1}^*, x_i^* + h, x_{i+1}^*, \ldots, x_n^*) < f(x_1^*, \ldots, x_{i-1}^*, x_i^*, x_{i+1}^*, \ldots, x_n^*),$$

což je spor s tím, že bod X^* je bodem lokálního extrému. Analogicky vyloučíme možnost $f'_{x_i}(X^*) < 0$. □

Poznámka. Podobně jako u funkce jedné proměnné, stacionární bod nemusí být bodem lokálního extrému, např. funkce $f(x, y) = x^2 - y^2$ má v bodě $[0, 0]$ stacionární bod, ale není zde extrém (takovým stacionárním bodům říkáme sedlový bod).

Definice 10.5 (kvadratické formy)**.** Funkce $K(X) = \sum_{i,j=1}^n b_{ij} x_i x_j$, kde $b_{ij} \in \mathbb{R}$ a $X \in \mathbb{R}^n$, se nazývá *kvadratická forma na* $\mathbb{R}^n$ (jedná se tedy o polynom druhého stupně n proměnných, ve kterém se nevyskytuje lineární a absolutní člen).

Pokud bod $X = [x_1, x_2, \ldots, x_n]$ identifikujeme s (vázaným) vektorem $\vec{x} = X - O = (x_1, x_2, \ldots, x_n)$, pak každé kvadratické formě lze jednoznačně přiřadit symetrickou matici $A = (a_{ij})_{i,j=1}^n$ tak, že platí $K(X) = \vec{x} A \vec{x}^T$. Prvky matice jsou dány $a_{ii} = b_{ii}$, $a_{ij} = a_{ji} = \frac{1}{2}(b_{ij} + b_{ji})$. Říkáme potom, že *matice A reprezentuje kvadratickou formu K*.

Definice 10.6 (definitnosti kvadratické formy)**.** Nechť A je symetrická matice. Řekneme, že kvadratická forma $\vec{x} A \vec{x}^T$ je

pozitivně definitní	, jestliže pro každý vektor $\vec{x} \neq \vec{o}$ platí	$\vec{x} A \vec{x}^T > 0$
negativně definitní		$\vec{x} A \vec{x}^T < 0$
pozitivně semidefinitní		$\vec{x} A \vec{x}^T \geq 0$,
negativně semidefinitní		$\vec{x} A \vec{x}^T \leq 0$

indefinitní, existují-li vektory $\vec{x}_1, \vec{x}_2$ tak, že $\vec{x}_1 A \vec{x}_1^T < 0$ a $\vec{x}_2 A \vec{x}_2^T > 0$ (tj. není-li definitní ani semidefinitní).

Poznámka. Protože každá kvadratická forma je symetrickou maticí A určena jednoznačně, lze hovořit přímo o definitnosti matice (namísto formy).

Definice 10.7 (hlavních minorů)**.** Nechť $A = (a_{ij})_{i,j=1}^n$ je čtvercová matice (ne nutně symetrická). Pak *základními (rohovými) hlavními minory* rozumíme determinanty

$$D_1 = \det(a_{11}), \quad D_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad D_2 = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \ldots, \quad D_n = \det A.$$

Základní hlavní minory tvoří podmnožinu všech hlavních minorů M_k^ℓ. *Hlavní minory řádu k* jsou determinanty čtvercových matic $k \times k$ ($k = 1, \ldots, n$), které vzniknou z matice A vynecháním $n - k$ řádků a $n - k$ sloupců, přičemž vynecháváme vždy odpovídající si sloupce a řádky (tj. i-tý řádek s i-tým sloupcem). Počet hlavních minorů řádu k je $\binom{n}{k}$, tj. $\ell = 1, 2, \ldots, \binom{n}{k}$. Celkový počet hlavních minorů je pak $\binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n - 1$.

Věta 10.8 (Sylvestrovo kritérium, James Joseph Sylvester 1814–1897, Angličan). *Nechť A je reálná symetrická $n \times n$-matice, D_k ($k = 1, \dots, n$) jsou její základní hlavní minory a M_k^ℓ ($k = 1, \dots, n$, $\ell = 1, \dots, \binom{n}{k}$) jsou její hlavní minory. Pak kvadratická forma $\vec{x} A \vec{x}^T$ je:*

(i) *pozitivně definitní $\Longleftrightarrow$ všechny základní hlavní minory D_k ($k = 1, \dots, n$) jsou kladné,*

(ii) *negativně definitní $\Longleftrightarrow$ pro všechny základní hlavní minory D_k platí $\operatorname{sgn} D_k = (-1)^k$, tj. $D_1 < 0$, $D_2 > 0$, $D_3 < 0, \dots$,*

(iii) *pozitivně semidefinitní $\Longleftrightarrow$ všechny hlavní minory M_k^ℓ jsou nezáporné,*

(iv) *negativně semidefinitní $\Longleftrightarrow$ pro všechny hlavní minory M_k^ℓ platí $\operatorname{sgn} M_k^\ell = (-1)^k$ nebo $M_k^\ell = 0$.*

Nenastane-li ani jeden z předchozích případů, pak $\vec{x} A \vec{x}^T$ je indefinitní.

Důkaz. Omezíme se na důkaz tvrzení (i) a (ii). Využijeme přitom faktu, že každá reálná symetrická matice A má reálná vlastní čísla, přičemž vlastní vektory jsou vzájemně ortogonální (i v případě vícenásobných vlastních čísel, tj. algebraická násobnost vlastního čísla je rovna geometrické násobnosti). Navíc, kvadratická forma $\vec{x} A \vec{x}^T$ je pozitivně definitní, právě když všechna vlastní čísla jsou kladná (důkaz těchto vlastností je založen na větě o hlavních osách, která říká, že každou reálnou symetrickou matici lze diagonalizovat pomocí matice ortogonální transformace). Důsledkem je, že pro uvažovanou matici A platí $a_{ii} > 0$ ($i = 1, 2, \dots, n$) a $D_n = \det(A) > 0$. Skutečně, položíme-li $\vec{x} = (0, \dots, 0, x_i, 0, \dots, 0)$, $x_i \neq 0$, pak z pozitivní definitnosti formy $\vec{x} A \vec{x}^T$ ihned máme $\vec{x} A \vec{x}^T = a_{ii} x_i^2 > 0$, tedy $a_{ii} > 0$. Protože všechna vlastní čísla λ_i matice A jsou kladná, platí $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n > 0$.

Označme symbolem A_k submatice matice A vystupující v definici 10.7, jejichž determinanty jsou čísla D_k. Pro libovolný nenulový vektor $\vec{y}_k = (y_1, y_2, \dots, y_k) \in \mathbb{R}^k$, $k \leq n$, položme $\vec{x}_k = (x_1, x_2, \dots, x_k, 0, \dots, 0) \in \mathbb{R}^n$.

„$\Longrightarrow$“ Nechť $\vec{x} A \vec{x}^T$ je pozitivně definitní a k je libovolný index z množiny $\{1, 2, \dots, n\}$. Pak $\vec{y}_k A_k \vec{y}_k^T = \vec{x}_k A \vec{x}_k^T > 0$, tj. kvadratická forma $\vec{y}_k A_k \vec{y}_k^T$ je pozitivně definitní a podle prvního odstavce $\det(A_k) = D_k > 0$.

„$\Longleftarrow$“ Tuto implikaci dokážeme matematickou indukcí. Nechť $D_k > 0$ pro každé $k = 1, 2, \dots, n$. V případě $k = 1$ máme $D_1 = \det(a_{11}) = a_{11} > 0$, což znamená, že $\vec{y}_1 A_1 \vec{y}_1 = a_{11} y_1^2 > 0$, tj. forma $\vec{y}_1 A_1 \vec{y}_1$ je pozitivně definitní. Pro indukční krok předpokládejme, že pro libovolné přirozené číslo $k = m-1$ je kvadratická forma $\vec{y}_{m-1} A_{m-1} \vec{y}_{m-1}^T$ pozitivně definitní, tj. $\vec{y}_{m-1} A_{m-1} \vec{y}_{m-1}^T > 0$ pro každý nenulový vektor $\vec{y}_{m-1} \in \mathbb{R}^{m-1}$. Ukažme, že potom je pozitivně definitní i forma $\vec{y}_m A_m \vec{y}_m^T$. Sporem: předpokládejme, že není pozitivně definitní. To znamená, že matice A_m musí mít alespoň 2 záporná vlastní čísla λ_1, λ_2. Kdyby měla pouze jedno záporné vlastní číslo, pak by muselo být $D_m = \lambda_1 \lambda_2 \cdots \lambda_n \leq 0$, což je spor s předpokladem kladnosti všech základních hlavních minorů. Ze stejného důvodu nemůže mít A_m nulové vlastní číslo. Ke dvojici vlastních čísel λ_1, λ_2 existují dva ortogonální vektory $\vec{u}, \vec{v} \in \mathbb{R}^m$ (tj. splňující $\vec{u} \cdot \vec{v} = 0$). Zvolme lineární kombinaci $\vec{w} = \alpha \vec{u} + \beta \vec{v}$ takovou, že vektor $\vec{w}$ je nenulový a má poslední složku nulovou. Protože $A_m \vec{u}^T = \lambda_1 \vec{u}^T$ a $A_m \vec{v}^T = \lambda_2 \vec{v}^T$, vynásobením příslušnými vektory zleva máme $\vec{u} A_m \vec{u}^T = \lambda_1 \vec{u} \vec{u}^T < 0$ a $\vec{v} A_m \vec{v}^T = \lambda_2 \vec{v} \vec{v}^T < 0$, a tedy $\vec{w} A_m \vec{w}^T = \alpha^2 (\vec{u} A_m \vec{u}^T) + \beta^2 (\vec{v} A_m \vec{v}^T) < 0$. Protože ale poslední složka vektoru $\vec{w}$ je nulová, musí také být $\vec{w}_{m-1} A_{m-1} \vec{w}_{m-1}^T < 0$, kde jsme (podobně jako výše) označili $\vec{w}_{m-1} = (w_1, w_2, \dots, w_{m-1})$. To je ale spor s indukčním předpokladem, tj. s pozitivní definitností formy s maticí A_{m-1}. Matice A_m tedy musí mít všechna vlastní čísla kladná, což podle prvního odstavce znamená, že kvadratická forma s maticí A_m je pozitivně definitní. Princip matematické indukce pak říká, že platí uvažovaná implikace pro libovolné $n \in \mathbb{N}$.

Větev (ii) potom plyne z (i) uvážíme-li, že když je forma $\vec{x} A \vec{x}^T$ pozitivně definitní, pak forma $-\vec{x} A \vec{x}^T = \vec{x}(-A)\vec{x}^T$ je negativně definitní (a naopak) a platí $\det(-A_k) = (-1)^k \det(A_k)$. $\square$

Důsledek 10.9. a) *Je-li alespoň jeden základní hlavní minor D_k záporný pro k sudé, pak $\vec{x} A \vec{x}^T$ je indefinitní.*

b) *Jsou-li všechny základní hlavní minory D_k nenulové a nenastává ani jeden z případů (i), (ii) předchozího tvrzení, pak $\vec{x} A \vec{x}^T$ je indefinitní.*

Příklad 10.10. Rozhodněte o typu kvadratických forem:

a) $K(x, y, z) = 3x^2 + y^2 + 4z^2 - 2xy + 4xz - 6yz$;

b) $K(x, y, z) = 16yz - 2z^2$.

Řešení. ad a) Kvadratická forma je reprezentována symetrickou maticí

$$A = \begin{pmatrix} 3 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 4 \end{pmatrix}. \quad \text{Máme } D_1 = 3,\ D_2 = 2,\ D_3 = -11.$$

Podle Sylvestrova kritéria je kvadratická forma pozitivně definitní.

ad b) V tomto případě kvadratické formě odpovídá matice

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 8 \\ 0 & 8 & -2 \end{pmatrix}, \quad \text{přičemž } M_1^1 = 0,\ M_1^2 = 0,\ M_1^3 = -2,\ M_2^1 = 0,\ M_2^2 = 0,\ M_2^3 = -16,\ M_3^1 = 0.$$

Podle Sylvestrova kritéria je forma indefinitní. V tomto případě šlo rozhodnout i snadněji, například $K(1, 1, 1) = 14 > 0$, ale $K(1, 1, -1) = -18 < 0$, tj. forma mění znaménko, což znamená indefinitnost.

Druhý totální diferenciál $d^2 f(X)$ je kvadratická forma reprezentovaná Hessovou[4] maticí druhých parciálních derivací

$$H_f(X) = \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots & f''_{x_1x_n} \\ f''_{x_2x_1} & f''_{x_2x_2} & \cdots & f''_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f''_{x_nx_1} & f''_{x_nx_2} & \cdots & f''_{x_nx_n} \end{pmatrix}(X),$$

tj. platí $d^2 f(X) = (h_1, h_2, \dots, h_n) H_f(X)(h_1, h_2, \dots, h_n)^T$.

Věta 10.11. *Nechť $f : \mathbb{R}^n \to \mathbb{R}$ je funkce, S její stacionární bod a nechť Hessova matice H_f je spojitá na nějakém okolí bodu S. Je-li $d^2 f(S)$*

(i) *pozitivně definitní, pak f má v bodě S ostré lokální minimum,*
(ii) *negativně definitní, pak f má v bodě S ostré lokální maximum,*
(iii) *indefinitní, pak f nemá v bodě S lokální extrém.*

Důkaz. (i) Z předpokladu spojitosti Hessovy matice na nějakém okolí bodu S plyne podle vět 8.2 a 8.4 spojitost parciálních derivací prvního řádu a funkce samotné na tomto okolí. Podle Taylorovy věty pak platí

$$f(X) = f(S) + d f(S; X - S) + \frac{1}{2} d^2 f(S + \vartheta(X - S); X - S), \quad \vartheta \in (0, 1).$$

Protože S je stacionární bod, platí $d f(S) = 0$, a tedy

$$f(X) - f(S) = \frac{1}{2} d^2 f(S + \vartheta(X - S); X - S).$$

Je-li $d^2 f(S)$ pozitivně definitní (tj. $d^2 f(S; X - S) > 0$ pro libovolný bod $X \neq S$), pak (z důvodu spojitosti) existuje okolí $O(S)$, na kterém je pozitivně definitní i $d^2 f(\hat{X})$ pro každý bod $\hat{X} \in O(S)$. Uvažujeme-li eukleidovskou metriku, tak do tohoto okolí náleží i bod $S + \vartheta(X - S)$, protože $\vartheta \in (0, 1)$. To znamená, že $d^2 f(S + \vartheta(X - S); X - S) > 0$, a tedy $f(X) - f(S) > 0$, tj. $f(X) > f(S)$, tj. v S je ostré lokální minimum.

Důkaz (ii) by se provedl stejně, důkaz (iii) je mírně technicky náročnější. □

Poznámka. Je-li $d^2 f(S)$ pozitivně nebo negativně semidefinitní, nelze o extrému rozhodnout.

Příklad 10.12. Vyšetřete lokální extrémy funkce $f(x, y, z) = 2xy^2 - 4xy + x^2 + z^2 - 2z$.

Řešení. Položíme-li gradient funkce roven nulovému vektoru, dostaneme soustavu tří rovnic

$$2y^2 - 4y + 2x = 0, \quad 4xy - 4x = 0, \quad 2z - 2 = 0,$$

ze které obdržíme tři stacionární body $S_1 = [1, 1, 1]$, $S_2 = [0, 0, 1]$ a $S_3 = [0, 2, 1]$. Hessova matice zadané funkce má tvar

$$H_f(x, y, z) = \begin{pmatrix} 2 & 4y - 4 & 0 \\ 4y - 4 & 4x & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

a po dosazení stacionárních bodů máme

$$H_f(S_1) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad H_f(S_2) = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad H_f(S_3) = \begin{pmatrix} 2 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Pomocí Sylvestrova kritéria snadno určíme, že $d^2 f(S_1)$ je pozitivně definitní kvadratická forma ($D_1 = 2 > 0$, $D_2 = 8 > 0$, $D_3 = 16 > 0$), $d^2 f(S_2)$ je indefinitní kvadratická forma ($D_1 = 2 > 0$, $D_2 = -16 < 0$, $D_3 = -32 < 0$) a $d^2 f(S_3)$ je také indefinitní ($D_1 = 2 > 0$, $D_2 = -16 < 0$, $D_3 = -32 < 0$). Lokální extrém tedy nastává pouze ve stacionárním bodě S_1 (a to ostré lokální minimum).

Příklad 10.13. Vyšetřete lokální extrémy funkce $f(x, y) = x^2(1 + y^2)$.

Řešení. Podmínka grad $f = \vec{o}$ vede na soustavu

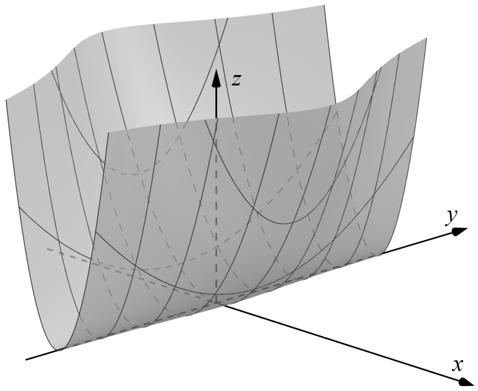
$$2x(1 + y^2) = 0, \quad 2x^2 y = 0,$$

která má nekonečně mnoho řešení $S = [0, y]$, $y \in \mathbb{R}$ (stacionárním bodem je tedy každý bod ležící na ose y). Hessova matice má potom tvar

$$H_f(x, y) = \begin{pmatrix} 2(1 + y^2) & 4xy \\ 4xy & 2x^2 \end{pmatrix}, \quad \text{vyčísleno v bodech } S\text{:} \quad H_f(x, y) = \begin{pmatrix} 2(1 + y^2) & 0 \\ 0 & 0 \end{pmatrix}.$$

[4] Ludwig Otto Hesse 1811–1874, Němec

Protože $M_1^1 = 2(1 + y^2) > 0$, $M_1^2 = 0$, $M_2^1 = 0$, $d^2 f(S)$ je pozitivně semidefinitní kvadratickou formou a podle poznámky výše nelze na základě věty 10.11 o extrému rozhodnout. Uvážíme-li však, že $f(S) = 0$ a $f(x, y) > 0\ \forall [x, y] \in \mathbb{R}^2$ mimo osu y, je zřejmé, že f má v každém bodě S lokální minimum, které však není ostré (graf funkce je na obrázku).



Obrázek 6: Graf funkce $f(x, y) = x^2(1 + y^2)$. Každý bod osy y je bodem (neostrého) lokálního minima

Definice 10.14. Buď $f : \mathbb{R}^n \to \mathbb{R}$ funkce a $M \subseteq D(f)$. Řekneme, že f nabývá na množině M v bodě $X_0 \in M$ *globálního minima* (resp. *maxima*), jestliže $f(X_0) \leq f(X)$ (resp. $f(X_0) \geq f(X)$) pro $\forall X \in M$. Jsou-li nerovnosti ostré pro každé $X \neq X_0$, mluvíme o *ostrých globálních minimech* (resp. *maximech*) na M. (Ostrá) globální maxima a minima nazýváme souhrnně *(ostré) globální extrémy funkce* f na množině M.

Poznámka. Je-li bod globálního extrému X_0 na množině M vnitřním bodem této množiny, pak je X_0 bodem lokálního extrému. Funkce f tedy může mít globální extrém na množině M v bodě lokálního extrému nebo v hraničním bodě množiny M, pokud tento bod do množiny M patří. Např. funkce $f(x, y) = x^2 + y^2$ má na uzavřeném jednotkovém kruhu $x^2 + y^2 \leq 1$ (ostré) globální minimum 0 v bodě $[0, 0]$ a (neostré) globální maximum 1 v každém bodě hranice kruhu (tj. kružnice $x^2 + y^2 = 1$). Kdybychom uvažovali otevřený kruh, tak funkce má stále (ostré) globální minimum 0 v bodě $[0, 0]$, ale globální maximum nemá (ostré ani neostré).

Praktický postup při výpočtu globálních extrémů na kompaktní množině (tj. ohraničené a uzavřené) bude ukázán později v kapitole o vázaných extrémech.