Jiří Bečvář; Miloslav Nekvinda

Poznámka o extrémech funkcí dvou a více proměnných

Časopis pro pěstování matematiky, Vol. 81 (1956), No. 3, 267--271

Persistent URL: `http://dml.cz/dmlcz/117194`

ČASOPIS PRO PĚSTOVÁNÍ MATEMATIKY
Vydává Matematický ústav ČSAV
SVAZEK 81 * PRAHA, 12.VII. 1956 * ČÍSLO 3

POZNÁMKA O EXTRÉMECH FUNKCÍ DVOU A VÍCE PROMĚNNÝCH

JIŘÍ BEČVÁŘ a MILOSLAV NEKVINDA, Liberec.

(Došlo dne 14. února 1955.)
DT: 517.514
517.27

Článek se zabývá případem, kdy u funkce dvou proměnných je determinant z druhých parciálních derivací ve vyšetřovaném bodě roven nule, v jeho okolí je však od nuly různý. Tento případ je zároveň zobecněn na konvexní (konkávní) funkce libovolného počtu proměnných.

Nechť funkce F dvou proměnných je definována v nějakém okolí bodu $A(a, b)$ a má v tomto okolí spojité druhé parciální derivace. Jestliže $F'_x(A) = F'_y(A) = 0$ a pro funkci $D = F''_{x^2}F''_{y^2} - (F''_{xy})^2$ platí $D(A) = 0$, pak podle běžné theorie nelze bez dalšího rozhodnout, zda funkce F má v bodě A extrém či ne. Přesto lze udat v jistých případech jednoduché postačující podmínky pro existenci resp. neexistenci extrému v takovém bodě. Tomuto případu a jistému jeho zobecnění je věnován tento článek.

Všude v dalším značí slova „derivace", „limita" konečnou derivaci resp. limitu. Vzdálenost dvou bodů X, Y v euklidovském prostoru $\mathbf{E}^n$ ($n \geq 1$) značíme $|X - Y|$. Derivace (resp. parciální derivace) značíme čárkou u označení funkce (resp. vyznačením proměnné, podle které se derivuje). Okolím vždy rozumíme, není-li řečeno jinak, otevřené souvislé okolí uvažovaného bodu.

Formulujme a dokažme nejprve větu, která si všímá speciálně funkcí dvou proměnných.

Věta 1. *Nechť funkce F dvou proměnných je definována v jistém okolí Ω bodu $A(a, b)$ a má v Ω spojité parciální derivace druhého řádu. Nechť $F'_x(A) = F'_y(A) = D(A) = 0$. Nechť pro všechny body $X \in \Omega$, $X \neq A$, platí $D(X) > 0$. Pak má F v bodě A ostrý lokální extrém a jeho charakter je určen znaménkem funkce F''_{x^2} na množině $\Omega - (A)$.*

Důkaz. Můžeme pro jednoduchost předpokládat, že Ω je kruhové okolí bodu A. Platí především, že na celé souvislé množině $\Omega - (A)$ zachovává funkce F''_{x^2} (a stejně i F''_{y^2}) znaménko. V opačném případě by totiž ze souvislosti množiny $\Omega - (A)$ a spojitosti funkce F''_{x^2} plynulo, že v nějakém bodě $Y \in \Omega - (A)$ jest $F''_{x^2}(Y) = 0$, tedy $D(Y) \leqq 0$, což je spor. Předpokládejme v dalším, že na množině $\Omega - (A)$ platí $F''_{x^2} > 0$.

Nechť $X(a + h, b + k)$ je libovolný bod množiny $\Omega - (A)$. Vzhledem k předpokladům věty platí Taylorova formule

$$F(X) = F(A) + h\,F'_x(A) + k\,F'_y(A) + \tfrac{1}{2}(h^2\,F''_{x^2}(\Theta) + 2hk\,F''_{xy}(\Theta) + k^2\,F''_{y^2}(\Theta)), \quad (1)$$

kde bod $\Theta(\vartheta_1, \vartheta_2)$ má souřadnice

$$\vartheta_1 = a + \vartheta h\,, \quad \vartheta_2 = b + \vartheta k\,, \quad 0 < \vartheta < 1\,, \qquad (2)$$

a tedy $\Theta \neq A$. Vzhledem k předpokladu

$$F'_x(A) = F'_y(A) = 0 \qquad (3)$$

plyne z (1)

$$F(X) - F(A) = \tfrac{1}{2}(h^2\,F''_{x^2}(\Theta) + 2hk\,F''_{xy}(\Theta) + k^2\,F''_{y^2}(\Theta))\,. \qquad (4)$$

Protože $\Theta \in \Omega - (A)$, je $F''_{x^2}(\Theta) > 0$ a (4) můžeme přepsat takto:

$$F(X) - F(A) = \frac{1}{2F''_{x^2}(\Theta)}\,[(h\,F''_{x^2}(\Theta) + k\,F''_{xy}(\Theta))^2 + k^2\,D(\Theta)]\,. \qquad (5)$$

Ježto $D(\Theta) > 0$ a čísla h, k nejsou současně rovna nule, plyne odtud, že $F(X) - F(A) > 0$. Bod X byl libovolný bod množiny $\Omega - (A)$, tedy F má v bodě A ostré lokální minimum.

Jestliže $F''_{x^2} < 0$ v $\Omega - (A)$, pak přechodem k funkci $-F$ a užitím předchozího výsledku dostaneme, že funkce F má v bodě A ostré lokální maximum. Tím je věta dokázána.

Věta 1 je v jistém smyslu speciálním případem obecnější věty, kterou teď uvedeme pro případ n proměnných. Připomeňme definici:

Jestliže funkce F n proměnných ($n \geqq 1$) je definována v nějakém okolí O bodu $A(a_1, \ldots, a_n)$ a v bodě A je diferencovatelná, pak říkáme, že F je v bodě A ryze konvexní, existuje-li okolí $\Omega \subset O$ bodu A takové, že pro každý bod $X \in \Omega$, $X \neq A$, je hodnota $F(X)$ větší než hodnota, odpovídající tečné nadrovině (resp. tečně pro $n = 1$ resp. tečné rovině pro $n = 2$), zkonstruované ke grafu funkce F v bodě $(a_1, \ldots, a_n, F(A))$. Podobně se definují pojmy „ryze konkávní", „konvexní", „konkávní" (vše s dodatkem: „v bodě A"). Naše věta potom zní:

Věta 2. *Nechť funkce F n proměnných je definována a diferencovatelná*) v nějakém okolí Ω bodu $A(a_1, \ldots, a_n)$. Nechť F je ryze konvexní (ryze konkávní) v každém bodě množiny $\Omega - (A)$ a platí $F'_{x_1}(A) = F'_{x_2}(A) = \ldots = F'_{x_n}(A) = 0$. Pak má F v bodě A ostré lokální minimum (ostré lokální maximum).*

Důkaz. Vyšetříme případ, že F je v $\Omega - (A)$ ryze konvexní. Můžeme předpokládat, že Ω je sférické okolí. Dále pro jednoduchost předpokládejme, že $F(A) = 0$.

---

*) T. j. má totální diferenciál.

Pro libovolný bod $X(x_1, \ldots, x_n)$ množiny $\Omega - (A)$ definujme funkci f takto:

$$f(t) = F(\ldots, a_i + t(x_i - a_i), \ldots)\,, \quad t \in \langle 0, 1\rangle\,.$$

Je $f(0) = F(A) = 0$, $f(1) = F(X)$, f má v intervalu $\langle 0, 1\rangle$ derivaci a platí $f'(0) = 0$. Snadno se zjistí, že vzhledem k předpokladům věty je f ryze konvexní pro všechna $t \in (0, 1)$.

Definujme ještě funkci g předpisem

$$g(t) = f(t) - t\,f(1)\,, \quad t \in \langle 0, 1\rangle\,. \tag{6}$$

Funkce g má v intervalu $\langle 0, 1\rangle$ derivaci a je zřejmě zase pro všechna $t \in (0, 1)$ ryze konvexní. V intervalu $\langle 0, 1\rangle$ nabývá maxima. Nemůže ho však nabýt ve vnitřním bodě, neboť pak by zřejmě v tomto bodě byla konkávní, což je spor s ryzí konvexitou. Odtud vzhledem ke vztahům $g(0) = g(1) = 0$ plyne, že je

$$g(t) < 0 \text{ pro všechna } t \in (0, 1)\,, \tag{7}$$

což podle (6) znamená, že je

$$f(t) < t\,f(1) \text{ pro všechna } t \in (0, 1)\,. \tag{7'}$$

Odtud plyne, že $0 = f'(0) = \lim\limits_{t\to 0} \dfrac{f(t)}{t} \leqq f(1) = F(X)$, tedy F je v $\Omega - (A)$ nezáporná, neboť bod X byl libovolný. Speciálně tedy pro funkci f, příslušnou k libovolnému bodu $X \in \Omega - (A)$, platí zřejmě $f(t) \geqq 0$ pro všechna $t \in (0, 1)$. Odtud podle (7′) plyne, že (pro $t \in (0,1)$) je $F(X) = f(1) > \dfrac{f(t)}{t} \geqq 0$, tedy $F(X) > 0$. Protože $F(A) = 0$, je tím tvrzení věty pro případ konvexity dokázáno.

Přechodem k funkci $-F$ se vyšetří případ, že F je v každém bodě $X \in \Omega - (A)$ ryze konkávní.

Tím je věta dokázána.

Poznámka 1. Je snadno vidět, že jsou-li splněny předpoklady věty 2 s výjimkou předpokladu $F'_{x_1}(A) = \ldots = F'_{x_n}(A) = 0$, pak funkce F je i v bodě A ryze konvexní (ryze konkávní). Stačí totiž od F odečíst její tečnou nadrovinu v bodě A; pak jsou pro tuto novou funkci splněny předpoklady věty 2 v plném rozsahu. Tedy má v bodě A ostré lokální minimum (ostré lokální maximum), je tedy speciálně ryze konvexní (ryze konkávní) v bodě A. Tato vlastnost se zřejmě neporuší, jestliže tečnou nadrovinu opět přičteme.

Poznámka 2. Věta 2 platí i v modifikované formě, nahradíme-li předpoklad ryzí konvexity pouhou konvexitou a tvrdíme-li v bodě A existenci minima (neostrého); podobně s konkávitou a maximem.

Poznámka 3. Souvislost věty 2 s větou 1 je ta, že předpoklady věty 1 zaručují zřejmě ryzí konvexitu (resp. ryzí konkávitu) funkce F v celém okolí bodu A.

Obraťme se nyní k případu, kdy u funkce dvou proměnných je determinant D v bodě A roven nule, v jeho okolí je však záporný (s výjimkou bodu A). Na rozdíl od předchozího postupu formulujme však nyní hned obecnou větu, jejíhož výsledku pak užijeme i v uvedeném speciálním případě funkce dvou proměnných.

Věta 3. *Nechť funkce F n proměnných je definována a je diferencovatelná v okolí bodu A. Nechť F je v bodě A konvexní (konkávní). Pak v libovolném okolí bodu A existuje bod $Y \neq A$ takový, že F je v Y konvexní (konkávní).*

Důkaz. Můžeme se omezit na případ konvexity. Předpokládejme pro jednoduchost, že bod A je počátek: $A(0, \ldots, 0)$. Definujme novou funkci H pro každý bod $X(x_1, \ldots, x_n)$ z definičního oboru funkce F (odečtením tečné nadroviny) takto:

$$H(X) = F(X) - (F(A) + \sum_{i=1}^{n} x_i F'_{x_i}(A)) . \tag{8}$$

Funkce H zřejmě rovněž splňuje předpoklady naší věty, nadto má v bodě A lokální minimum, rovné nule, a platí

$$H'_{x_1}(A) = \ldots = H'_{x_n}(A) = 0 . \tag{9}$$

Existuje tedy okolí Ω bodu A takové, že platí $H(X) \geqq 0$ pro všechny body $X \in \Omega$. Dokažme, že funkce H splňuje tvrzení naší věty. Odtud pak ihned plyne, že je splňuje i funkce F.

Rozeznávejme dva možné případy:

a) Existuje posloupnost bodů $\{X_k\}$ taková, že platí $X_k \to A$, $H(X_k) = 0$, a přitom pro všechna k je $X_k \in \Omega$, $X_k \neq A$. Odtud vzhledem k významu množiny Ω plyne, že v každém bodě X_k má funkce H lokální minimum a tedy je tam zřejmě konvexní. Ježto $X_k \neq A$, $X_k \to A$, je v tomto případě tvrzení věty pro funkci H dokázáno.

b) Existuje $\delta > 0$ tak, že příslušné uzavřené sférické okolí bodu A o poloměru δ je částí okolí Ω a pro všechny jeho body $X \neq A$ platí $H(X) > 0$. Nechť δ je libovolné takové číslo a $\overline{\Omega}_1$ příslušné uzavřené sférické okolí bodu A.

Funkce H je spojitá na hranici množiny $\overline{\Omega}_1$ a nabývá tam minima, které označme ε. Jest $\varepsilon > 0$. Definujme pro každý bod $X(x_1, \ldots, x_n) \in \Omega$ novou funkci G (odečtením jisté nadroviny od funkce H) takto:

$$G(X) = H(X) - \frac{\varepsilon}{2\delta} x_1 . \tag{10}$$

Je pak zřejmě

$$G(A) = G(0, \ldots, 0) = 0 , \tag{11}$$

$$G(X) \geqq \varepsilon - \frac{\varepsilon}{2\delta} x_1 \geqq \frac{\varepsilon}{2} \text{ na hranici množiny } \overline{\Omega}_1 . \tag{12}$$

Spojitá funkce G nabývá na uzavřené množině $\bar{\Omega}_1$ minima a vzhledem k (11), (12) ho může nabýt pouze ve vnitřním bodě množiny $\bar{\Omega}_1$. Označme takový bod Θ. Je pak $G'_{x_1}(\Theta) = \ldots = G'_{x_n}(\Theta) = 0$. Odtud vzhledem k (10) plyne

$$H'_{x_1}(\Theta) = \frac{\varepsilon}{2\delta}\,; \quad H'_{x_i}(\Theta) = 0\,, \quad i = 2, \ldots, n\,. \tag{13}$$

Z (13) a (9) plyne, že $\Theta \neq A$. Protože funkce G nabývá v bodě Θ minima, je tam zřejmě konvexní. Z (10) plyne, že v bodě Θ je zřejmě konvexní i funkce H. Ježto $\Theta \neq A$ a okolí Ω_1 bylo možno volit libovolně malé, je tím naše tvrzení pro funkci H dokázáno.

Tím je věta dokázána.

Jako důsledek věty 3 dostáváme tuto větu:

Věta 4. *Nechť funkce F dvou proměnných je definována v jistém okolí Ω bodu $A(a, b)$ a má v Ω spojité parciální derivace druhého řádu. Nechť $F'_x(A) = F'_y(A) = D(A) = 0$ (kde D je opět determinant z druhých parc. derivací). Nechť pro všechny body $X \in \Omega$, $X \neq A$ platí $D(X) < 0$. Pak F nemá v bodě A lokální extrém (ani neostrý).*

Důkaz. Kdyby funkce F měla v bodě A lokální extrém, byla by tam konvexní neb konkávní. Podle věty 3 by v libovolné blízkosti bodu A existovaly body, různé od bodu A, v nichž by F byla konvexní nebo konkávní. V těchto bodech však determinant D je záporný a odtud jak známo plyne, že tam F nemůže být ani konvexní, ani konkávní.

Poznamenejme nakonec toto: Má-li funkce F dvou proměnných v nějakém okolí bodu A spojité druhé parciální derivace, je-li $D(A) = 0$ a existují-li v libovolně malém okolí bodu A jak body, ve kterých je determinant D kladný, tak i body, v nichž je záporný, pak není vyloučen žádný z těchto dvou případů:

a) funkce F nemá v bodě A lokální extrém (ani neostrý),

b) funkce F má v bodě A ostrý lokální extrém.

První případ je ilustrován funkcí $F(x, y) = x^3 + y^3$, která v bodě $A = (0, 0)$ zřejmě nemá lokální extrém. Determinant $D(x, y) = 36xy$ je v bodě A roven nule a v libovolném jeho okolí nabývá jak kladných, tak i záporných hodnot. — Druhý případ nastává u funkce $F(x, y) = x^5 \sin \frac{1}{x} + x^4 + y^4$, kterou na ose y dodefinujeme rovnicí $F(0, y) = y^4$. Funkce F má v počátku zřejmě ostré lokální minimum a lze ukázat, že v celé rovině má spojité druhé parciální derivace. Jest $D(0, 0) = 0$ a pro body neležící na ose y máme $D(x, y) = 12y^2(20x^3 \sin \frac{1}{x} - 8x^2 \cos \frac{1}{x} - x \sin \frac{1}{x} + 12x^2)$. Odtud je vidět, že v libovolném okolí počátku nabývá D jak kladných, tak záporných hodnot.