Vázané extrémy funkce více proměnných

text pro studenty FP TUL

20. prosince 2024

Martina Šimůnková

Definice vázaného extrému. Nechť $\Omega \subseteq \mathbb{R}^d$ je otevřená množina, $a \in \Omega$, $f : \Omega \to \mathbb{R}$, $g : \Omega \to \mathbb{R}$, $M_g := \{x \in \mathbb{R}^d : g(x) = 0\}$.

Řekneme, že funkce f má v bodě $a \in M_g$ lokální minimum (respektive maximum) vázané na množinu M_g, pokud existuje $\delta > 0$, že pro všechna $x \in M_g \cap U_\delta(a)$, $x \neq a$ platí $f(x) > f(a)$ (respektive $f(x) < f(a)$).

Příklad.
$\Omega = \mathbb{R}^2$, $f(x, y) = x^2 + x - y^2$, $g(x, y) = 2x + y$.
Z rovnice $g(x, y) = 0$ vyjádříme y a úlohu převedeme na hledání extrému funkce jedné proměnné

$$\hat{f}(x) = f(x, -2x), \quad x \in \mathbb{R}$$

Po chvíli počítání dostaneme $\hat{f}(x) = x - 3x^2$, která má minimum lokální i globální rovné $1/12$ v bodě $x = 1/6$. Dopočítáme $y = -2x = -1/3$ a dostaneme: funkce f má lokální i globální maximum rovné $1/12$ v bodě $(1/6, -1/3)$.

Příklad.
$g(x, y) = x^2 + y^2 - 1$, množina M_g je kružnice. Body $(x, y) \in M_g$ rozdělíme do čtyř skupin

1. $x \in (-1, 1)$, $y(x) = \sqrt{1 - x^2}$,
2. $x \in (-1, 1)$, $y(x) = -\sqrt{1 - x^2}$,
3. $y \in (-\delta, \delta)$, $x(y) = \sqrt{1 - y^2}$,
4. $y \in (-\delta, \delta)$, $x(y) = -\sqrt{1 - y^2}$,

Případy 1, 2 převedeme na hledání extrému funkce jedné proměnné (a analogicky i případy 3, 4)

$$\hat{f}(t) = f(t, y(t)), \quad t \in (-1, 1)$$

Řetízkové pravidlo:

$$\hat{f}'(t) = f_x(t, y(t)) + y'(t) f_y(t, y(t)) \tag{1}$$
$$0 = g_x(t, y(t)) + y'(t) g_y(t, y(t)) \tag{2}$$

Pokud je $g_y(t, y(t)) \neq 0$, vyjádříme $y'(t)$ z rovnice (2) a dosadíme do (1). Budeme řešit soustavu dvou rovnic pro dvě neznámé

$$f_x(x, y) - \frac{g_x(x, y)}{g_y(x, y)} f_y(x, y) = 0 \tag{3}$$
$$g(x, y) = 0 \tag{4}$$

Věta o implicitně zadané funkci. Nechť $\Omega \subseteq \mathbb{R}^2$ je otevřená množina, $(a, b) \in \Omega$, $g \in \mathcal{C}^1(\Omega)$, $g(a, b) = 0$, $g_y(a, b) \neq 0$.

Pak existují $\delta_x > 0$, $\delta_y > 0$ taková že pro každé $x \in U_{\delta_x}(a)$ existuje právě jedno $y(x) \in U_{\delta_y}(b)$ splňující

$$g(x, y(x)) = 0 \tag{5}$$

Funkce $y : U_{\delta_x}(a) \to \mathbb{R}$ definovaná vztahem (5) je diferencovatelná v okolí bodu a a pro $t \in U_{\delta_x}(a)$ platí

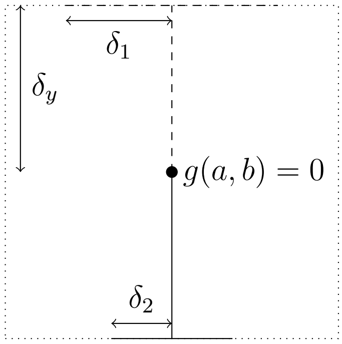
$$y'(t) = -\frac{g_x(t, y(t))}{g_y(t, y(t))} \tag{6}$$

Důkaz. Bez újmy na obecnosti budeme předpokládat, že $g_y(a, b) > 0$ (jinak přejdeme k funkci $-g$).

Protože je $g \in \mathcal{C}^1(\Omega)$, existuje $\delta_y > 0$, že platí $g_y(x, y) > 0$ pro $x \in U_{\delta_y}(a)$, $y \in U_{\delta_y}(b)$. Odtud plyne, že funkce $f(y) = g(x, y)$ je rostoucí na $U_{\delta_y}(b)$ pro $x \in U_{\delta_y}(a)$.

Odtud a ze spojitosti g v bodě $(a, b+\delta_y)$ plyne existence $\delta_1 > 0$, že $g(x, y)$ je kladná na čárkovaných úsečkách na obrázku. Analogicky zdůvodníme, že na plných úsečkách je g záporná.

Odtud a ze spojitosti g dále plyne, že pro $\delta_x := \min\{\delta_1, \delta_2\}$ a $x \in U_{\delta_x}(a)$ existuje právě jedno $y \in U_{\delta_y}(b)$ takové, že $g(x, y) = 0$.



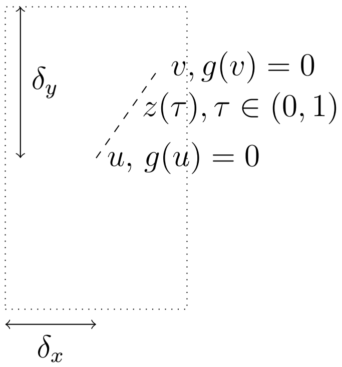
na čárkovaných úsečkách je $g(x, y) > 0$

na plných úsečkách je $g(x, y) < 0$

K důkazu (6) odvodíme vztah pro

$$\frac{y(a+t)-y(a)}{t}$$

Položíme
$u := (a, b)$
$v := (a+t, y(a+t))$
$z(\tau) := u + \tau(v-u)$
Platí: $g(u) = g(v) = 0$, $v - u = (t, y(a+t) - y(a))$.



Dále na úsečce uv definujeme funkci

$$\varphi(\tau) := g(z(\tau))$$

a ověříme, že splňuje na intervalu $[0,1]$ předpoklady Rolleovy věty o střední hodnotě: Platí $\varphi(0) = g(u) = 0$, $\varphi(1) = g(v) = 0$ a z řetízkového pravidla pro $\tau \in (0,1)$ plyne

$$\varphi'(\tau) = (v-u) \cdot \nabla g(z(\tau)) = t g_x(z(\tau)) + (y(a+t) - y(a)) g_y(z(\tau)) \tag{7}$$

Z Rolleovy věty pak plyne, že existuje $c \in (0,1)$ takové, že platí

$$\varphi'(c) = 0 \tag{8}$$

Dosazením do (7) dostaneme

$$0 = t g_x(z(c)) + (y(a+t) - y(a)) g_y(z(c))$$

Odtud vyjádříme

$$\frac{y(a+t)-y(a)}{t} = -\frac{g_x(z(c))}{g_y(z(c))} \tag{9}$$

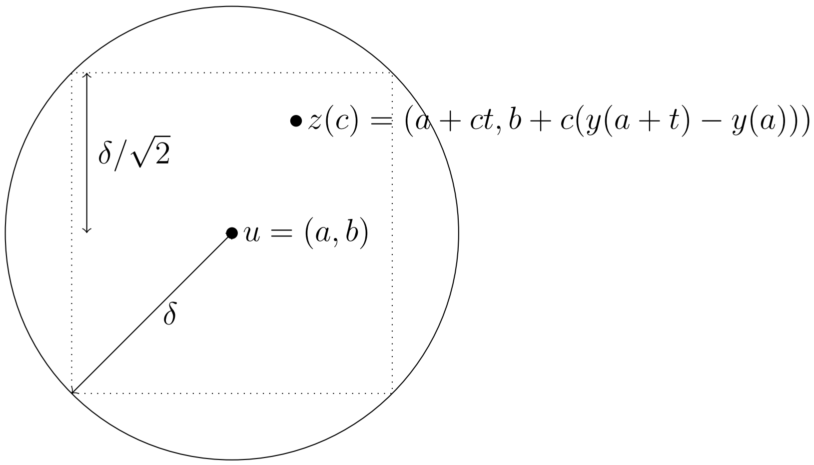
Nechť je $\varepsilon > 0$. Ze spojitosti funkce $(x,y) \mapsto g_x(x,y)/g_y(x,y)$ v bodě (a,b) plyne existence $\delta > 0$ takového, že pro $(x,y) \in U_\delta(a,b)$ platí

$$\left| \frac{g_x(x,y)}{g_y(x,y)} - \frac{g_x(a,b)}{g_y(a,b)} \right| < \varepsilon \tag{10}$$

Níže ukážeme, že existuje δ_t takové, že pro $t \in (-\delta_t, \delta_t)$ a $c \in (0,1)$ je $z(c) \in U_\delta(a,b)$. Odtud a z (9), (10) pro $t \in (-\delta_t, \delta_t)$ plyne

$$\left| \frac{y(a+t) - y(a)}{t} + \frac{g_x(a,b)}{g_y(a,b)} \right| < \varepsilon$$

a odtud plyne (6).



Dokažme existenci δ_t. Podobně jako na začátku důkazu ukážeme, že k $\delta_y \le \delta/\sqrt{2}$ existuje $\delta_x \in (0, \delta_y)$ takové, že ke každému $x \in U_{\delta_x}(a)$ existuje právě jedno $y \in U_{\delta_y}(b)$ takové, že platí $g(x,y) = 0$. Zvolíme-li $\delta_t := \delta_x$, pak pro $t \in (-\delta_t, \delta_t)$, je $z(c) = (a + ct, b + c(y(a+t) - y(a)) \in U_\delta(a,b)$.

Poznámka. Funkci y budeme nazývat funkcí implicitně zadanou rovnicí $g(t, y(t)) = 0$ na okolí bodu (a,b).

Věta (nutná podmínka pro vázaný extrém). Nechť $\Omega \subseteq \mathbb{R}^2$ je otevřená množina, $(a,b) \in \Omega$, $f, g \in \mathcal{C}^1(\Omega)$, $\nabla g(a,b) \neq (0,0)$. Nechť má funkce f v bodě (a,b) lokální vázaný extrém na množině M_g.

Pak existuje $\lambda \in \mathbb{R}$ takové, že platí

$$\nabla f(a,b) = \lambda \nabla g(a,b) \tag{11}$$

Důkaz. Protože je $\nabla g(a,b) \neq (0,0)$, bez újmy na obecnosti předpokládáme $g_y(a,b) \neq 0$ (jinak přejdeme k funkci $\hat{g}(x,y) = g(y,x)$).

Z věty o implicitně zadané funkci plyne existence $\delta > 0$ takového, že na okolí $U_\delta(a)$ je funkce y dána vztahem $g(x, y(x)) = 0$ a její derivace vztahem (6).

Položme

$$\lambda = \frac{f_y(a,b)}{g_y(a,b)} \tag{12}$$

Z rovnice (3) plyne $f_x(a,b) = \lambda g_x(a,b)$, a odtud plyne (11).

Lagrangeova funkce. Funkci

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

nazýváme Lagrangeovou funkcí a číslo λ nazýváme Lagrangeovým multiplikátorem.

Poznámka. Nutnou podmínku existence vázaného extrému můžeme pomocí Lagrangeovy funkce napsat ve tvaru

$$\nabla F(a,b,\lambda) = (0,0,0) \tag{13}$$

Věta (druhá derivace implicitně zadané funkce). Nechť $\Omega \subseteq \mathbb{R}^2$ je otevřená množina, $(a,b) \in \Omega$, $g \in \mathcal{C}^2(\Omega)$, $g(a,b) = 0$, $g_y(a,b) \neq 0$. Nechť $y : U(a) \to \mathbb{R}$ je funkce implicitně zadaná rovnicí $g(t, y(t)) = 0$ na okolí bodu (a,b). Pak existuje druhá derivace funkce y v bodě a a je rovna

$$y''(a) = -\frac{g_{xx}(a,b) + 2y'(a)g_{xy}(a,b) + (y'(a))^2 g_{yy}(a,b)}{g_y(a,b)} \tag{14}$$

Důkaz. Podle věty o derivaci složené funkce (řetízkové pravidlo) má pravá strana (6) derivaci podle proměnné t. Odtud plyne existence $y''(t)$. Aplikací řetízkového pravidla dostaneme

$$\begin{aligned} y''(t) &= -\frac{d}{dt}\left(\frac{g_x(t,y(t))}{g_y(t,y(t))}\right) = -\frac{g_y \frac{d}{dt} g_x - g_x \frac{d}{dt} g_y}{(g_y)^2} \\ &= \frac{-g_y g_{xx} - y'(t) g_y g_{xy} + g_x g_{xy} + y'(t) g_x g_{yy}}{(g_y)^2} \end{aligned}$$

Dosadíme $g_x = -y'(t)g_y$, $t = a$, pokrátíme zlomek g_y a dostaneme (14).

Věta (druhá derivace ve stacionárním bodě). Nechť $f, g \in \mathcal{C}^2(\Omega)$, $(a,b) \in \Omega$, $g(a,b) = 0$ a existuje $\lambda \in \mathbb{R}$, že platí (11). Nechť y je funkce implicitně zadaná rovnicí $g(t, y(t)) = 0$ na okolí bodu (a,b). Funkce $\hat{f}$ nechť je definovaná na okolí bodu a vztahem

$$\hat{f}(t) = f(t, y(t)) \tag{15}$$

Pak

$$\begin{aligned} \hat{f}'(a) &= 0 \qquad (16) \\ \hat{f}''(a) &= (1, y'(a)) \begin{pmatrix} F_{xx}(a,b,\lambda) & F_{xy}(a,b,\lambda) \\ F_{xy}(a,b,\lambda) & F_{yy}(a,b,\lambda) \end{pmatrix} (1, y'(a))^T \qquad (17) \end{aligned}$$

Důkaz. Použijeme řetízkové pravidlo na (15)

$$\hat{f}'(t) = f_x(t, y(t)) + y'(t) f_y(t, y(t))$$

dosazením derivace $y'(t)$ z (6) a f_x, f_y z (11) dostaneme (16).

Druhým použitím řetízkového pravidla postupně dostaneme

$$\begin{aligned} \hat{f}''(t) &= \frac{d}{dt}\left(f_x(t, y(t)) + y'(t) f_y(t, y(t))\right) \\ &= f_{xx} + y'(t) f_{xy} + y''(t) f_y + y'(t)(f_{xy} + y'(t) f_{yy}) \\ &= f_{xx} + 2y'(t) f_{xy} + (y'(t))^2 f_{yy} + y''(t) f_y \end{aligned}$$

Označíme $k = y'(t)$, za $y''(t)$ dosadíme z (14)

$$\hat{f}''(a) = f_{xx} + 2k f_{xy} + k^2 f_{yy} - \frac{f_y}{g_y}(g_{xx} + 2k g_{xy} + k^2 g_{yy})$$

Za podíl f_y/g_y dosadíme z předpokladu (11)

$$\hat{f}''(a) = f_{xx} + 2k f_{xy} + k^2 f_{yy} - \lambda(g_{xx} + 2k g_{xy} + k^2 g_{yy})$$

a dostaneme

$$\hat{f}''(a) = F_{xx} + 2k F_{xy} + k^2 F_{yy}$$

Pomocí Hessovy matice $H := \nabla^2 F(a, b)$ definované v (18) upravíme na

$$\hat{f}''(a) = (1, k) H (1, k)^T$$

K odvození (17) stačí ukázat

$$(1, k) \cdot \nabla g(a, b) = 0,$$

což ověříme dosazením

$$k = y'(t) = -\frac{g_x(a, b)}{g_y(a, b)}, \quad \nabla g(a, b) = (g_x(a, b), g_y(a, b))$$

Věta (postačující podmínka pro vázaný extrém). Nechť $\Omega \subseteq \mathbb{R}^2$ je otevřená množina, $(a,b) \in \Omega$, $f, g \in \mathcal{C}^2(\Omega)$, $g(a,b) = 0$, $\nabla g(a,b) \neq (0,0)$. Nechť existuje $\lambda \in \mathbb{R}$ takové, že pro Lagrangeovu funkci platí (13). Nechť

$$\nabla^2 F(a,b) := \begin{pmatrix} F_{xx}(a,b,\lambda) & F_{xy}(a,b,\lambda) \\ F_{xy}(a,b,\lambda) & F_{yy}(a,b,\lambda) \end{pmatrix} \tag{18}$$

a $h \in \mathbb{R}^2 \setminus \{(0,0)\}$ je vektor kolmý na $\nabla g(a,b)$.

Pak platí

(i) Pokud je $h\nabla^2 F(a,b)h^T > 0$, pak má funkce f v bodě (a,b) lokální minimum vázané na množinu M_g.

(ii) Pokud je $h\nabla^2 F(a,b)h^T < 0$, pak má funkce f v bodě (a,b) lokální maximum vázané na množinu M_g.

Důkaz. Stačí ukázat, že funkce $\hat{f}$ definovaná vztahem (15) má extrém v bodě a. Zvolme $h = (1, y'(a))$. Pak z (6) plyne $h \cdot \nabla g(a,b) = 0$ a z věty o druhé derivace ve stacionárním bodě plyne $\hat{f}'(a) = 0$, $\hat{f}''(a) = h\nabla^2 F(a,b)h^T$, odkud plyne tvrzení věty.

Příklad. Určete poměr poloměru podstavy válce a jeho výšky, chceme-li, aby povrch válce byl co nejmenší při zadaném objemu válce.

Označíme poloměr podstavy válce x a výšku válce y a budeme hledat minimum funkce $f(x,y) = 2\pi x^2 + 2\pi xy$ vázaný na množinu

$$M = \{[x,y] \in (0,\infty) \times (0,\infty) : \pi x^2 y - V = 0\}$$

Protože nechceme spočítat hodnotu extrému, ale jen poměr $x : y$, můžeme hledat extrém funkce $f(x,y) = x^2 + xy$ s vazbou $g(x,y) = x^2 y - c$.

Vypočteme gradienty $\nabla f(x,y) = (2x + y, x)$, $\nabla g(x,y) = (2xy, x^2)$ a napíšeme nutnou podmínku pro existenci extrému

$$\begin{aligned} 2x + y &= 2\lambda xy \\ x &= \lambda x^2 \end{aligned}$$

Protože je $x \neq 0$, vyjádříme z druhé rovnice $\lambda x = 1$ a dosadíme do první. Dostaneme

$$2x + y = 2y$$

Odtud dostaneme $y = 2x$ a hledaný poměr poloměru podstavy a výšky válce tedy je $r : v = 1 : 2$. Tj. výška je dvakrát větší než poloměr podstavy.

Ještě ověříme, že se jedná o minimum:

1. Lagrangeova funkce je $F(x,y) = x^2 + xy - \lambda(x^2y - c)$, její druhé derivace jsou $F_{xx} = 2 - 2\lambda y$, $F_{xy} = 1 - 2\lambda x$, $F_{yy} = 0$.

2. Z podmínky $\nabla f(r,v) = \lambda \nabla g(r,v)$ jsme odvodili $\lambda = 1/r$. Odtud dostaneme
$$\nabla^2 F(r,v) = \begin{pmatrix} 2 - 2v/r & -1 \\ -1 & 0 \end{pmatrix}$$

3. Vektor h: pro $\nabla g(r,v) = (2rv, 2r^2)$ dostaneme ze vztahu $h \cdot \nabla g(r,v) = 0$ vektor $h = (-r, v)$.

4. Výpočtem dostaneme $h\nabla^2 F(r,v)h^T = 2r(r-v)$.

5. Výše jsme spočítali $v = 2r$, proto je $h\nabla^2 F(r,v)h^T < 0$ a funkce f má v bodě (r,v) lokální vázané minimum.

Úlohy.

1. Ukažte, že pro symetrickou matici
$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$
a vektor $h = (x,y)$ je
$$hAh^T = ax^2 + 2bxy + cy^2$$
a pro vektor $h = (1,k)$ je
$$hAh^T = a + 2bk + ck^2.$$

2. Určete symetrickou matici A, tak aby platilo
$$(x,y)A(x,y)^T = 5x^2 - 6xy + y^2$$