

Optimalizace

8. Lokální extrémy vázané rovnostmi (úlohy s omezeními typu rovnosti)

Tomáš Werner, Tomáš Kroupa

28. listopadu 2025

FEL ČVUT

Úlohy na extrémý vázané rovnostmi (= úlohy s omezeními typu rovnosti)

Hledáme extrémý funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na množině

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \}$$

kde $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$. To také zapisujeme takto:

$$\begin{array}{ll} \min/\max & f(\mathbf{x}) \\ \text{za podmíněk} & \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{x} \in \mathbb{R}^n \end{array}$$

Příklad (jednoduchý):

$$\begin{array}{ll} \min & x + y \\ \text{za podmínky} & x^2 + y^2 = 1 \\ & x, y \in \mathbb{R} \end{array}$$

- Máme $n = 2$ proměnné a $m = 1$ omezení.
- $f(x, y) = x + y$
- $g(x, y) = x^2 + y^2 - 1$ nebo $g(x, y) = 1 - x^2 - y^2$

Lokální extrémy vázané **lineárními** rovnostmi

Hledáme extrémy diferencovatelné funkce $f: \mathbb{R}^n \rightarrow \mathbb{R}$ na afinním podprostoru

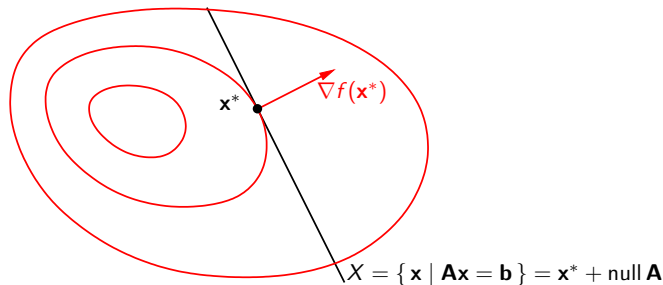
$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b} \}$$

kde $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. Neboli řešíme úlohu

$$\begin{array}{ll} \min/\max & f(\mathbf{x}) \\ \text{za podmínek} & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}^n \end{array}$$

Tedy $\mathbf{g}(\mathbf{x}) = \mathbf{Ax} - \mathbf{b}$ je afinní zobrazení.

Podmínka prvního řádu



Věta

Nechť x^* je lokální extrém funkce f za podmínky $Ax = b$. Pak $\nabla f(x^*) \perp \text{null } A$.

Důkaz: Protože $Ax^* = b$, množinu řešení soustavy $Ax = b$ lze psát jako

$$x^* + \text{null } A = \{x^* + By \mid y \in \mathbb{R}^k\}$$

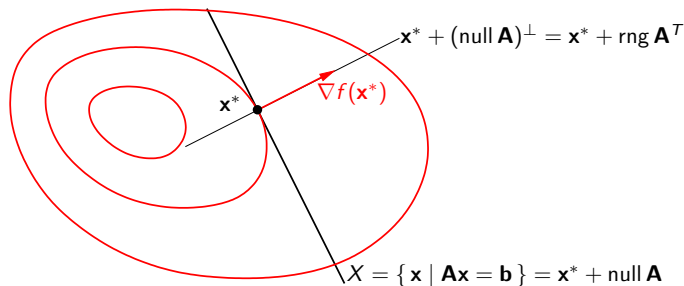
kde sloupce matice $B \in \mathbb{R}^{m \times k}$ tvoří bázi podprostoru $\text{null } A$.

Tedy bod $y = 0$ musí být **volný** lokální extrém funkce $\varphi(y) = f(x^* + By)$. Tedy

$$\varphi'(y) = f'(x^* + By)B = f'(x^*)B = 0.$$

Ale $f'(x^*)B = 0$ je totéž jako $\nabla f(x^*) \perp \text{null } A$.

Formulace podmínky soustavou rovnic



$$\nabla f(\mathbf{x}^*) \perp \text{null } \mathbf{A} \iff \nabla f(\mathbf{x}^*) \in (\text{null } \mathbf{A})^\perp = \text{rng}(\mathbf{A}^T)$$

$$\iff \nabla f(\mathbf{x}^*) \text{ je lin. kombinace sloupců } \mathbf{A}^T \text{ (tj. řádků } \mathbf{A})$$

$$\iff f'(\mathbf{x}^*) = -\boldsymbol{\lambda}^T \mathbf{A} \text{ pro nějaké } \boldsymbol{\lambda} \in \mathbb{R}^m$$

Podmínka prvního řádu (soustava $m + n$ rovnic s $m + n$ neznámými):

$$f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{A} = \mathbf{0}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

Příklad: Řešení lineární soustavy s nejmenší normou

$$\begin{array}{ll} \min & \frac{1}{2} \|\mathbf{x}\|^2 \\ \text{za podmínek} & \mathbf{Ax} = \mathbf{b} \end{array}$$

Podmínky prvního řádu:

$$\mathbf{x} + \mathbf{A}^T \boldsymbol{\lambda} = \mathbf{0}$$

$$\mathbf{Ax} = \mathbf{b}$$

Má-li $\mathbf{A}$ lineárně nezávislé řádky, pak optimálním řešením je $\mathbf{x} = \underbrace{\mathbf{A}^T (\mathbf{AA}^T)^{-1}}_{\mathbf{A}^+} \mathbf{b}$.

Příklad: Úloha nejmenších čtverců s lineárními omezeními

$$\begin{array}{ll} \min & \frac{1}{2} \|\mathbf{Ax} - \mathbf{b}\|^2 \\ \text{za podmínek} & \mathbf{Cx} = \mathbf{d} \end{array}$$

Podmínky prvního řádu:

$$\begin{aligned} \mathbf{A}^T(\mathbf{Ax} - \mathbf{b}) + \mathbf{C}^T\lambda &= \mathbf{0} \\ \mathbf{Cx} &= \mathbf{d} \end{aligned}$$

neboli

$$\begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

Tvrzení: Matice $\begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}$ je regulární, právě když matice $\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix}$ má l.n. sloupce a matice $\mathbf{C}$ má l.n. řádky.

Lokální extrémy vázané **nelineárními** rovnostmi

Jsou dány spojitě diferencovatelné $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Hledáme extrémy funkce f na množině

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \}.$$

Neboli řešíme úlohu

$$\begin{array}{ll} \min/\max & f(\mathbf{x}) \\ \text{za podmíněk} & \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{x} \in \mathbb{R}^n \end{array}$$

‘Odvození’ (neformální!) podmínky na lokální extrémy:

1. Množinu X **linearizujeme** v okolí lokálního extrému.
2. Tím úlohu **převédeme na případ lineárních omezení**, který už známe.

Tečný prostor k množině přípustných řešení

Jak 'linearizovat' množinu $X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \}$ v okolí daného bodu $\mathbf{x}^* \in X$?

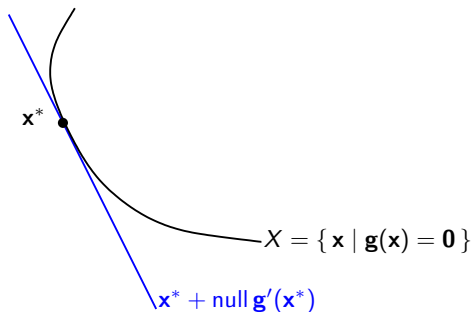
Hledáme 'tečný prostor' k množině X v bodě $\mathbf{x}^*$.

Aproximujme $\mathbf{g}$ v okolí bodu $\mathbf{x}^*$ afinním zobrazením (Taylorovým polynomem 1. stupně)

$$\mathbf{g}(\mathbf{x}) \approx \mathbf{T}_{\mathbf{x}^*}^1(\mathbf{x}) = \underbrace{\mathbf{g}(\mathbf{x}^*)}_0 + \mathbf{g}'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$$

Množina X se tím změní na (afinní) podprostor

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = \mathbf{0} \} = \mathbf{x}^* + \text{null } \mathbf{g}'(\mathbf{x}^*)$$



Podmínka regularity

Potíž: Ne vždy se množina X v okolí bodu $\mathbf{x}^*$ podobá afinnímu podprostoru $\mathbf{x}^* + \text{null } \mathbf{g}'(\mathbf{x}^*)$ (**přestože** se zobrazení $\mathbf{g}$ v okolí bodu $\mathbf{x}^*$ podobá afinnímu zobrazení $\mathbf{T}_{\mathbf{x}^*}^1$).

Definice

Bod $\mathbf{x} \in \mathbb{R}^n$ je **regulární bod** zobrazení $\mathbf{g}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, jestliže řádky Jacobiho matice $\mathbf{g}'(\mathbf{x})$ (tj. gradienty $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$) jsou lineárně nezávislé.

Speciálně pro $m = 1$: Bod $\mathbf{x}$ je regulární bod funkce $g: \mathbb{R}^n \rightarrow \mathbb{R}$, jestliže $\nabla g(\mathbf{x}) \neq \mathbf{0}$.

Tvrzení (neformální)

Nechť $\mathbf{x}^* \in X$ je regulární bod zobrazení $\mathbf{g}$.

Pak množina X je v okolí bodu $\mathbf{x}^*$ 'podobná' podprostoru $\mathbf{x}^* + \text{null } \mathbf{g}'(\mathbf{x}^*)$.

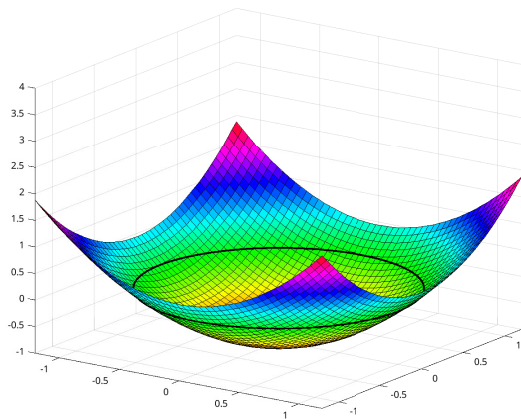
Ten má dimenzi $n - m$ a je 'tečný' k množině X v bodě $\mathbf{x}^*$.

(Tvrzení je neformální, protože jsme nedefinovali pojmy 'tečný' a 'podobná'.)

Příklad: dvě různé reprezentace kružnice

Kružnice $X = \{ (x, y) \in \mathbb{R}^2 \mid g(x, y) = 0 \}$ je vrstevnice výšky nula funkce g :

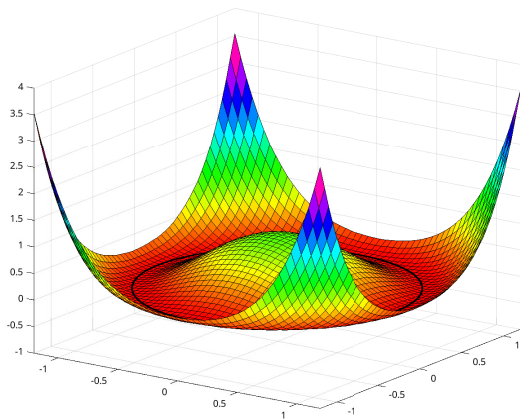
$$g(x, y) = x^2 + y^2 - 1 = 0$$



$$\nabla g(x, y) = 2(x, y)$$

Všechny body z X jsou regulární body g .

$$g(x, y) = (x^2 + y^2 - 1)^2 = 0$$

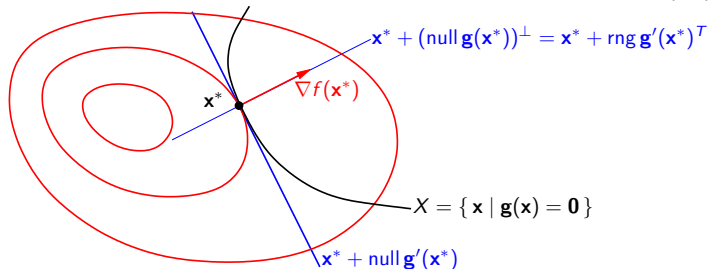


$$\nabla g(x, y) = 4(x^2 + y^2 - 1)(x, y) = (0, 0) \quad \forall (x, y) \in X$$

Žádný bod z X není regulární bod g .

Podmínka prvního řádu

V podmínce pro lineární omezení nahradíme matici $\mathbf{A}$ Jacobiho maticí $\mathbf{g}'(\mathbf{x}^*)$:



Věta

Nechť $\mathbf{x}^*$ je regulární bod zobrazení $\mathbf{g}$ a lokální extrém funkce f za podmínky $\mathbf{g}(\mathbf{x}) = \mathbf{0}$. Pak $\nabla f(\mathbf{x}^*) \perp \text{null } \mathbf{g}'(\mathbf{x}^*)$.

$$\nabla f(\mathbf{x}^*) \perp \text{null } \mathbf{g}'(\mathbf{x}^*) \iff f'(\mathbf{x}^*) = -\lambda^T \mathbf{g}'(\mathbf{x}^*) \text{ pro nějaké } \lambda \in \mathbb{R}^m$$

Podmínka prvního řádu (soustava $m + n$ rovnic s $m + n$ neznámými):

$$f'(\mathbf{x}) + \lambda^T \mathbf{g}'(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) = \mathbf{0}$$

Lagrangeova funkce

Zavedeme **Lagrangeovu funkci** $L: \mathbb{R}^{m+n} \rightarrow \mathbb{R}$

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x})$$

Pak soustava výše jde napsat jako

$$L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) = f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x}) = \mathbf{0}$$

$$L_{\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{g}(\mathbf{x})^T = \mathbf{0}$$

neboli $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$, tj. $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionární bod funkce L .

Čísla $(\lambda_1, \dots, \lambda_m) = \boldsymbol{\lambda}$ se nazývají **Lagrangeovy multiplikátory**.

$$\begin{array}{ll} \min & x + y \\ \text{za podmínky} & x^2 + y^2 = 1 \end{array}$$

Lagrangeova funkce

$$L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2)$$

Stacionární podmínky pro L :

$$L_x(x, y, \lambda) = 1 - 2\lambda x = 0$$

$$L_y(x, y, \lambda) = 1 - 2\lambda y = 0$$

$$L_\lambda(x, y, \lambda) = 1 - x^2 - y^2 = 0$$

Řešení: $(x, y, \lambda) = \pm(1, 1, 1)/\sqrt{2}$.

Našli jsme dva kandidáty na lok. extrémy: $(x, y) = \pm(1, 1)/\sqrt{2}$.

Příklad: Nesplněné podmínky regularity

$$\begin{array}{ll} \min & x + y \\ \text{za podmínky} & (1 - x^2 - y^2)^2 = 0 \end{array}$$

Množina přípustných řešení X je **stejná kružnice** jako minule, ale gradient

$$\nabla g(x, y) = -4(1 - x^2 - y^2)(x, y)$$

je na kružnici X nulový, tedy žádný bod X není regulární.

Lagrangeova funkce

$$L(x, y, \lambda) = x + y + \lambda(1 - x^2 - y^2)^2$$

Podmínky stacionarity

$$L_x(x, y, \lambda) = 1 - 4\lambda x(1 - x^2 - y^2) = 0$$

$$L_y(x, y, \lambda) = 1 - 4\lambda y(1 - x^2 - y^2) = 0$$

$$L_\lambda(x, y, \lambda) = (1 - x^2 - y^2)^2 = 0$$

nemají řešení. Lokální extrémy jsme nenašli, i když existují.

Příklad: Lineární omezení

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{za podmínky} & \mathbf{Ax} = \mathbf{b} \end{array}$$

Lagrangeova funkce

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T (\mathbf{Ax} - \mathbf{b})$$

Podmínky stacionarity

$$L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) = f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{A} = \mathbf{0}$$

$$L_{\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) = (\mathbf{Ax} - \mathbf{b})^T = \mathbf{0}$$

To jsme už odvodili dříve.

Pro lineární omezení ale **nemusíme** předpokládat $\text{rank } \mathbf{g}'(\mathbf{x}) = \text{rank } \mathbf{A} = m$.