Optimalizace

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

Tom Werner, Tom Kroupa

FEL ČVUT

Lokální extrémy vázané lineárními rovnostmi

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

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

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

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

Podmínka prvního řádu

Věta

Jestliže $\mathbf{x}$ je lokální extrém funkce f za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Důkaz: Nechť $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$ a $\mathbf{B} \in \mathbb{R}^{m \times k}$ splňuje $\operatorname{rng} \mathbf{B} = \operatorname{null} \mathbf{A}$. Pak

$$\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \} = \mathbf{x}_0 + \operatorname{null} \mathbf{A} = \{ \mathbf{x}_0 + \mathbf{B}\mathbf{y} \mid \mathbf{y} \in \mathbb{R}^k \}.$$

Teď minimalizujeme funkci $\varphi(\mathbf{y}) = f(\mathbf{x}_0 + \mathbf{B}\mathbf{y})$ bez omezení. Z řetízkového pravidla je

$$\varphi'(\mathbf{y}) = f'(\mathbf{x}_0 + \mathbf{B}\mathbf{y})\mathbf{B} = f'(\mathbf{x})\mathbf{B} = \mathbf{0}.$$

Ale $f'(\mathbf{x})\mathbf{B} = \mathbf{0}$ znamená $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Podmínka prvního řádu

Věta

Jestliže $\mathbf{x}$ je lokální extrém funkce f za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Důkaz: Nechť $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$ a $\mathbf{B} \in \mathbb{R}^{m\times k}$ splňuje $\operatorname{rng} \mathbf{B} = \operatorname{null} \mathbf{A}$. Pak

$$\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b} \,\} = \mathbf{x}_0 + \operatorname{null} \mathbf{A} = \{\, \mathbf{x}_0 + \mathbf{B}\mathbf{y} \mid \mathbf{y} \in \mathbb{R}^k \,\}.$$

Teď minimalizujeme funkci $\varphi(\mathbf{y}) = f(\mathbf{x}_0 + \mathbf{B}\mathbf{y})$ bez omezení. Z řetízkového pravidla je

$$\varphi'(\mathbf{y}) = f'(\mathbf{x}_0 + \mathbf{B}\mathbf{y})\mathbf{B} = f'(\mathbf{x})\mathbf{B} = \mathbf{0}.$$

Ale $f'(\mathbf{x})\mathbf{B} = \mathbf{0}$ znamená $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Ale $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$ je ekvivalentní $\nabla f(\mathbf{x}) \in (\operatorname{null} \mathbf{A})^\perp = \operatorname{rng}(\mathbf{A}^T)$.
Tedy $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ je lineární kombinací řádků $\mathbf{A}$.
Tedy $f'(\mathbf{x}) = -\boldsymbol{\lambda}^T\mathbf{A}$ (mínus je konvence) pro nějaké $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$.

Závěr: Jestliže $\mathbf{x}$ je lokální extrém, pak existuje $\boldsymbol{\lambda}$ tak, že

$$\begin{aligned} f'(\mathbf{x}) + \boldsymbol{\lambda}^T\mathbf{A} &= \mathbf{0} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

To je soustava $m + n$ rovnic s $m + n$ neznámými.

Příklad: Řešení lineární soustavy s nejmenší normou

$$\begin{aligned} \text{min} \quad & \tfrac{1}{2}\|\mathbf{x}\|^2 \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

Příklad: Řešení lineární soustavy s nejmenší normou

$$
\begin{aligned}
\min \quad & \tfrac{1}{2}\|\mathbf{x}\|^2 \\
\text{za podmínek} \quad & \mathbf{Ax} = \mathbf{b}
\end{aligned}
$$

Podmínky prvního řádu:

$$
\begin{aligned}
\mathbf{x} - \mathbf{A}^T \boldsymbol{\lambda} &= \mathbf{0} \\
\mathbf{Ax} &= \mathbf{b}
\end{aligned}
$$

Příklad: Řešení lineární soustavy s nejmenší normou

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|\mathbf{x}\|^2 \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

Podmínky prvního řádu:

$$\begin{aligned} \mathbf{x} - \mathbf{A}^T \boldsymbol{\lambda} &= \mathbf{0} \\ \mathbf{A}\mathbf{x} &= \mathbf{b} \end{aligned}$$

Má-li **A** lineárně nezávislé řádky, pak optimálním řešením je

$$\mathbf{x} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b}.$$

Příklad: Úloha nejmenších čtverců s lineárními omezeními

$$\begin{aligned} \text{min} \quad & \tfrac{1}{2}\|\mathbf{Ax}-\mathbf{b}\|^2 \\ \text{za podmínek} \quad & \mathbf{Cx}=\mathbf{d} \end{aligned}$$

Příklad: Úloha nejmenších čtverců s lineárními omezeními

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|^2 \\ \text{za podmínek} \quad & \mathbf{Cx} = \mathbf{d} \end{aligned}$$

Podmínky prvního řádu:

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

Příklad: Úloha nejmenších čtverců s lineárními omezeními

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|^2 \\ \text{za podmínek} \quad & \mathbf{Cx} = \mathbf{d} \end{aligned}$$

Podmínky prvního řádu:

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

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

Interpretace λ jako stínových cen omezení

Označ

$$F(\mathbf{b}) = \min\{\, f(\mathbf{x}) \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \in \mathbb{R}^n \,\}$$

optimální hodnotu úlohy jako funkci pravých stran omezení $\mathbf{b}$.

Interpretace λ jako stínových cen omezení

Označ

$$F(\mathbf{b}) = \min\{\, f(\mathbf{x}) \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \in \mathbb{R}^n \,\}$$

optimální hodnotu úlohy jako funkci pravých stran omezení $\mathbf{b}$.

Věta (neformálně)

Nechť $\mathbf{x}$ je minimum funkce f za podmínek $\mathbf{Ax} = \mathbf{b}$.
Nechť $(\mathbf{x}, \boldsymbol{\lambda})$ je řešení soustavy

$$\begin{aligned} \mathbf{x} - \mathbf{A}^T\boldsymbol{\lambda} &= \mathbf{0} \\ \mathbf{Ax} &= \mathbf{b} \end{aligned}$$

Pak $\nabla F(\mathbf{b}) = F'(\mathbf{b})^T = \boldsymbol{\lambda}$.

Přesněji (ale stále neformálně):

Nechť $\mathbf{x}(\mathbf{b}), \boldsymbol{\lambda}(\mathbf{b})$ je řešení soustavy jako funkce $\mathbf{b}$.
Označ $F(\mathbf{b}) = f(\mathbf{x}(\mathbf{b}))$. Pak $\nabla F(\mathbf{b}) = F'(\mathbf{b})^T = \boldsymbol{\lambda}(\mathbf{b})$.

Podmínky druhého řádu

Definice: Matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ je *positivně semidefinitní na podprostoru* $Y \subseteq \mathbb{R}^n$, jestliže $\mathbf{x}^T \mathbf{C} \mathbf{x} \geq 0$ pro každé $\mathbf{x} \in Y$.
To znamená, že matice $\mathbf{B}^T \mathbf{C} \mathbf{B}$ je pos. semidefinitní, kde $Y = \operatorname{rng} \mathbf{B}$.

Podmínky druhého řádu

Definice: Matice $\mathbf{C} \in \mathbb{R}^{n \times n}$ je *positivně semidefinitní na podprostoru* $Y \subseteq \mathbb{R}^n$, jestliže $\mathbf{x}^T \mathbf{C} \mathbf{x} \geq 0$ pro každé $\mathbf{x} \in Y$.
To znamená, že matice $\mathbf{B}^T \mathbf{C} \mathbf{B}$ je pos. semidefinitní, kde $Y = \operatorname{rng} \mathbf{B}$.

Věta

Nechť $f : \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná v bodě $\mathbf{x} \in \mathbb{R}^n$.

- Jestliže $\mathbf{x}$ je lokální minimum [maximum] funkce f za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $f''(\mathbf{x})$ je positivně [negativně] semidefinitní na podprostoru $\operatorname{null} \mathbf{A}$.
- Jestliže $\mathbf{A}\mathbf{x} = \mathbf{b}$ a $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$ (= podmínka prvního řádu) a $f''(\mathbf{x})$ je positivně [negativně] definitní na podprostoru $\operatorname{null} \mathbf{A}$, pak $\mathbf{x}$ je ostré lokální minimum [maximum] funkce f za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Důkaz. Ukázali jsme, že $\mathbf{x} = \mathbf{B}\mathbf{y} + \mathbf{x}_0$ je lokální extrém funkce f vázaný podmínkou $\mathbf{A}\mathbf{x} = \mathbf{b}$, právě když $\mathbf{y}$ je lokální extrém funkce $\varphi(\mathbf{y}) = f(\mathbf{B}\mathbf{y} + \mathbf{x}_0)$ bez omezení. Je

$$\varphi''(\mathbf{y}) = \mathbf{B}^T f''(\mathbf{B}\mathbf{y} + \mathbf{x}_0)\mathbf{B} = \mathbf{B}^T f''(\mathbf{x})\mathbf{B}.$$

Teď můžeme použít podmínky druhého řádu na volné lok. extrémy.

Lokální extrémy vázané nelineárními rovnostmi

Jsou dána spojitě diferencovatelná $f\colon \mathbb{R}^n \to \mathbb{R}$ a $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$.
Hledáme extrémy funkce f na množině

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

Neboli řešíme úlohu

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

Lineární omezení jsou speciální případ pro $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}$.

Intuice pro nalezení podmínek:
Množinu X linearizujeme v okolí extrému, čímž situaci převedeme na případ lineárních omezení.

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 $\mathbf{x} \in X$?
Zkusíme $\mathbf{g}$ nahradit Taylorovým polynomem 1. stupně:

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

Množina X se tím změní na afinní podprostor

$$\tilde{X} = \{\, \mathbf{y} \in \mathbb{R}^n \mid \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{0} \,\} \;=\; \mathbf{x} + \operatorname{null} \mathbf{g}'(\mathbf{x})$$

Ten ovšem v okolí bodu $\mathbf{x}$ nemusí být 'podobný' množině X
(přestože zobrazení $\mathbf{T}^1_{\mathbf{x}}$ je v okolí bodu $\mathbf{x}$ podobné zobrazení $\mathbf{g}$)!

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 $\mathbf{x} \in X$?
Zkusíme $\mathbf{g}$ nahradit Taylorovým polynomem 1. stupně:

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

Množina X se tím změní na afinní podprostor

$$\tilde{X} = \{ \mathbf{y} \in \mathbb{R}^n \mid \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{0} \} \;=\; \mathbf{x} + \operatorname{null} \mathbf{g}'(\mathbf{x})$$

Ten ovšem v okolí bodu $\mathbf{x}$ nemusí být 'podobný' množině X
(přestože zobrazení $\mathbf{T}^1_{\mathbf{x}}$ je v okolí bodu $\mathbf{x}$ podobné zobrazení $\mathbf{g}$)!

Definice

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

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 $\mathbf{x} \in X$? Zkusíme $\mathbf{g}$ nahradit Taylorovým polynomem 1. stupně:

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

Množina X se tím změní na afinní podprostor

$$\tilde{X} = \{ \mathbf{y} \in \mathbb{R}^n \mid \mathbf{g}'(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{0} \} \;=\; \mathbf{x} + \operatorname{null} \mathbf{g}'(\mathbf{x})$$

Ten ovšem v okolí bodu $\mathbf{x}$ nemusí být 'podobný' množině X
(přestože zobrazení $\mathbf{T}^1_{\mathbf{x}}$ je v okolí bodu $\mathbf{x}$ podobné zobrazení $\mathbf{g}$)!

Definice

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

Tvrzení (neformální)

Je-li $\mathbf{x} \in X$ regulární bod zobrazení $\mathbf{g}$,
pak množina X je v okolí bodu $\mathbf{x}$ 'podobná' afinnímu podprostoru $\tilde{X}$.
Ten má dimenzi $n - m$ a je 'tečný' k množině X v bodě $\mathbf{x}$.

Podmínka prvního řádu

Zopakujme podmínku prvního řádu pro lineární omezení:

Věta

Jestliže $\mathbf{x}$ je lokální extrém funkce f za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Intuice: nahradíme $\operatorname{null} \mathbf{A}$ tečným podprostorem k X v okolí extrému:

Věta

Nechť $\mathbf{x}$ je regulární bod zobrazení $\mathbf{g}$.
Jestliže $\mathbf{x}$ je lok. extrém funkce f za podmínky $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{g}'(\mathbf{x})$.

Podmínka prvního řádu

Zopakujme podmínku prvního řádu pro lineární omezení:

Věta

Jestliže $\mathbf{x}$ je lokální extrém funkce f za podmínky $\mathbf{Ax} = \mathbf{b}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Intuice: nahradíme $\operatorname{null} \mathbf{A}$ tečným podprostorem k X v okolí extrému:

Věta

Nechť $\mathbf{x}$ je regulární bod zobrazení $\mathbf{g}$.
Jestliže $\mathbf{x}$ je lok. extrém funkce f za podmínky $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{g}'(\mathbf{x})$.

Ale $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{g}'(\mathbf{x})$ je ekvivalentní $\nabla f(\mathbf{x}) \in (\operatorname{null} \mathbf{g}'(\mathbf{x}))^\perp = \operatorname{rng}(\mathbf{g}'(\mathbf{x})^T)$.
Tedy $\nabla f(\mathbf{x})$ je lineární kombinací řádků $\mathbf{g}'(\mathbf{x})$ (tj. gradientů $\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})$).
Tedy $f'(\mathbf{x}) = -\boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x})$ pro nějaké $\boldsymbol{\lambda} \in \mathbb{R}^m$.

Podmínka prvního řádu

Zopakujme podmínku prvního řádu pro lineární omezení:

Věta

Jestliže $\mathbf{x}$ je lokální extrém funkce f za podmínky $\mathbf{A}\mathbf{x} = \mathbf{b}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$.

Intuice: nahradíme $\operatorname{null} \mathbf{A}$ tečným podprostorem k X v okolí extrému:

Věta

Nechť $\mathbf{x}$ je regulární bod zobrazení $\mathbf{g}$.
Jestliže $\mathbf{x}$ je lok. extrém funkce f za podmínky $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, pak $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{g}'(\mathbf{x})$.

Ale $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{g}'(\mathbf{x})$ je ekvivalentní $\nabla f(\mathbf{x}) \in (\operatorname{null} \mathbf{g}'(\mathbf{x}))^\perp = \operatorname{rng}(\mathbf{g}'(\mathbf{x})^T)$.
Tedy $\nabla f(\mathbf{x})$ je lineární kombinací řádků $\mathbf{g}'(\mathbf{x})$ (tj. gradientů $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$).
Tedy $f'(\mathbf{x}) = -\boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x})$ pro nějaké $\boldsymbol{\lambda} \in \mathbb{R}^m$.

Závěr: Jestliže $\mathbf{x}$ je lokální extrém, pak existuje $\boldsymbol{\lambda}$ tak, že

$$\begin{aligned} f'(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}'(\mathbf{x}) &= \mathbf{0} \\ \mathbf{g}(\mathbf{x}) &= \mathbf{0} \end{aligned}$$

To je soustava $m + n$ rovnic s $m + n$ neznámými.

Lagrangeova funkce

Zavedeme-li **Lagrangeovu funkci** $L\colon \mathbb{R}^{m+n} \to \mathbb{R}$

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

pak soustava výše jde napsat jako

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

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

($L_{\mathbf{x}}$ resp. $L_{\boldsymbol{\lambda}}$ značí parciální derivace L podle $\mathbf{x}$ resp. $\boldsymbol{\lambda}$.)

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

Příklad

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

Příklad

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

Lagrangeova funkce

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

Stacionární podmínky pro L:

$$\begin{aligned} L_x(x, y, \lambda) &= 1 - 2\lambda x &&= 0 \\ L_y(x, y, \lambda) &= 1 - 2\lambda y &&= 0 \\ L_\lambda(x, y, \lambda) &= 1 - x^2 - y^2 &&= 0 \end{aligned}$$

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

$$\begin{aligned} \min \quad & x \\ \text{za podmínky} \quad & x^2 = 0 \end{aligned}$$

Lokální extrém $x = 0$ není regulární bod, protože $g'(x) = 2x = 0$ (tedy $\operatorname{rank} g'(x) < 1$).

Příklad: Nesplněné podmínky regularity

$$\begin{aligned} \min \quad & x \\ \text{za podmínky} \quad & x^2 = 0 \end{aligned}$$

Lokální extrém $x = 0$ není regulární bod, protože $g'(x) = 2x = 0$ (tedy $\operatorname{rank} g'(x) < 1$).

Lagrangeova funkce

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

Podmínky stacionarity

$$\begin{aligned} L_x(x, \lambda) &= 1 + 2x = 0 \\ L_\lambda(x, \lambda) &= x^2 \qquad = 0 \end{aligned}$$

nemají řešení. Lokální extrém jsme nenašli.

Příklad: Nesplněné podmínky regularity

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

Množina přípustných řešení X je stejná kružnice jako minule, ale gradient funkce $g(x, y) = (1 - x^2 - y^2)^2$ je na kružnici X nulový, tedy žádný bod X není íregulární.

Příklad: Nesplněné podmínky regularity

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

Množina přípustných řešení X je stejná kružnice jako minule, ale gradient funkce $g(x, y) = (1 - x^2 - y^2)^2$ je na kružnici X nulový, tedy žádný bod X není íregulární.

Lagrangeova funkce

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

Podmínky stacionarity

$$\begin{aligned} L_x(x, y, \lambda) &= 1 - 4\lambda x(1 - x^2 - y^2) = 0 \\ L_y(x, y, \lambda) &= 1 - 4\lambda y(1 - x^2 - y^2) = 0 \\ L_y(x, y, \lambda) &= (1 - x^2 - y^2)^2 \qquad\quad = 0 \end{aligned}$$

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

Příklad: Lineární omezení

$$
\begin{aligned}
\min \quad & f(\mathbf{x}) \\
\text{za podmínky} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Příklad: Lineární omezení

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínky} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

Lagrangeova funkce

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

Podmínky stacionarity

$$\begin{aligned} L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) &= f'(\mathbf{x}) - \boldsymbol{\lambda}^T\mathbf{A} = \mathbf{0} \\ L_{\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) &= (\mathbf{b} - \mathbf{A}\mathbf{x})^T \quad = \mathbf{0} \end{aligned}$$

To jsme už odvodili dříve.

Pro lineární omezení jsme ale nemuseli předpokládat $\operatorname{rank} \mathbf{g}'(\mathbf{x}) = \operatorname{rank} \mathbf{A} = m$.

Příklad z optiky: Zákon odrazu z Fermatova principu

Fermatův princip nejkratšího času v optice (1662)

Paprsek mezi dvěma body letí takovou dráhou, aby doba letu byla nejkratší.

Modernější upřesnění: Doba letu je *stacionární* (vzhledem k variacím dráhy).

Příklad z optiky: Zákon odrazu z Fermatova principu

Fermatův princip nejkratšího času v optice (1662)

Paprsek mezi dvěma body letí takovou dráhou, aby doba letu byla nejkratší.
Modernější upřesnění: Doba letu je *stacionární* (vzhledem k variacím dráhy).

Křivé zrcadlo je plocha $X = \{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = 0 \,\}$ kde $g\colon \mathbb{R}^n \to \mathbb{R}$.
Délka paprsku mezi dvěma body $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ s odrazem od zrcadla v bodě $\mathbf{x} \in X$ je

$$f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{b}\|.$$

Hledáme stacionární bod funkce f za podmínky $g(\mathbf{x}) = 0$.

Lagrangeova funkce

$$L(\mathbf{x}, \lambda) = \|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{b}\| + \lambda g(\mathbf{x})$$

Podmínka $L_{\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$ je (po transpozici)

$$(\mathbf{x} - \mathbf{a})^0 + (\mathbf{x} - \mathbf{b})^0 + \lambda \nabla g(\mathbf{x}) = \mathbf{0} \qquad \text{kde jsme označili } \mathbf{y}^0 = \frac{\mathbf{y}}{\|\mathbf{y}\|}.$$

To říká, že vektor $\nabla g(\mathbf{x})$ (normála k zrcadlu v bodě $\mathbf{x}$) leží v jedné rovině s jednotkovými vektory $(\mathbf{x} - \mathbf{a})^0$ a $(\mathbf{x} - \mathbf{b})^0$ a půlí úhel mezi nimi.