Optimalizace

8. Vázané extrémy (optimalizační úlohy s omezeními)

Tom Werner, Tom Kroupa

FEL ČVUT

Úloha s omezeními ve tvaru rovností

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

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

Tedy řešíme úlohu

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

neboli

$$\begin{array}{rl} \min & f(x_1, \dots, x_n) \\ \text{za podmínek} & g_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, m \end{array}$$

Extrémy vázané lineárními rovnostmi

Lineární omezení

Když $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ je afinní zobrazení ($\mathbf{A} \in \mathbb{R}^{m \times n}$), řešíme úlohu

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

Věta (Podmínka prvního řádu)

Nechť $\mathbf{x}$ je lokální extrém úlohy. Pak $\nabla f(\mathbf{x}) \in \operatorname{rng} \mathbf{A}^T$.

Důkaz: Nechť $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$ a sloupce $\mathbf{C} \in \mathbb{R}^{m \times k}$ tvoří bázi null $\mathbf{A}$. Pak

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

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

$$\frac{\mathrm{d}}{\mathrm{d}\mathbf{y}} f(\mathbf{x}_0 + \mathbf{C}\mathbf{y}) = f'(\mathbf{x}_0 + \mathbf{C}\mathbf{y})\mathbf{C} = f'(\mathbf{x})\mathbf{C} = \mathbf{0}.$$

Ale $f'(\mathbf{x})\mathbf{C} = \mathbf{0}$ znamená $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{A}$, tj. $\nabla f(\mathbf{x}) \in (\operatorname{null} \mathbf{A})^\perp = \operatorname{rng} \mathbf{A}^T$.

Použití věty: Je-li $\mathbf{x}$ lokální extrém, pak existuje $\boldsymbol{\lambda} \in \mathbb{R}^m$ tak, že

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

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

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

Příklad: Řešení lineární soustavy s nejmenší normou

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

Podmínky prvního řádu:

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

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

Má-li $\mathbf{A}$ 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

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

Podmínky prvního řádu:

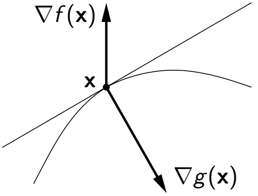
$$\begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

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.

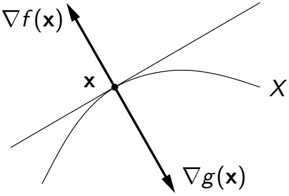
Extrémy vázané nelineárními rovnostmi

Neformální úvaha pro případ $m = 1$ (jedno omezení)

Pro $f, g \colon \mathbb{R}^n \to \mathbb{R}$ hledáme lok. extrémy f na množině $X = \{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = 0 \,\}$.



x není lokální extrém



x je lokální extrém

Nechť $\mathbf{x}$ je lokální extrém.

- Pak $\nabla f(\mathbf{x})$ je kolmý k množině X v bodě $\mathbf{x}$.
  Jinak by se malým pohybem body $\mathbf{x}$ po X hodota $f(\mathbf{x})$ zmenšila nebo zvětšila.
- Protože $\nabla g(\mathbf{x})$ je kolmý k X v bodě $\mathbf{x}$, jsou tedy $\nabla f(\mathbf{x})$ a $\nabla g(\mathbf{x})$ rovnoběžné.
  Tedy $\nabla f(\mathbf{x}) + \lambda \nabla g(\mathbf{x}) = \mathbf{0}$ pro nějaké číslo λ.

Problém: Tato úvaha nefunguje, když $\nabla g(\mathbf{x}) = \mathbf{0}$.

Tečné vektory k množině X

Nechť $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$. Zkoumejme množinu $X = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \,\}$.

- **Hladká křivka** v $\mathbb{R}^n$ je diferencovatelné zobrazení $\varphi\colon (\alpha_1, \alpha_2) \to \mathbb{R}^n$ (zde (α_1, α_2) je otevřený interval).
- Vektor $\mathbf{v} = \varphi'(\alpha)$ je **tečný** ke křivce v bodě $\mathbf{x} = \varphi(\alpha)$.
- Křivka leží na množině X, jestliže $\mathbf{g}(\varphi(\alpha)) = \mathbf{0}$ pro všechna $\alpha \in (\alpha_1, \alpha_2)$.

Definice

Vektor je **tečný k množině** X v bodě $\mathbf{x} \in X$,
jestliže je v bodě $\mathbf{x}$ tečným vektorem nějaké hladké křivky ležící na X.

Vztah tečných vektorů a gradientů

Tvrzení

Nechť je vektor $\mathbf{v} \in \mathbb{R}^n$ tečný k množině X v bodě $\mathbf{x}$.
Pak $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$ (tj. gradienty $\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})$ jsou kolmé na $\mathbf{v}$).

Důkaz: Existuje hladká křivka φ na X taková, že $\mathbf{x} = \varphi(\alpha)$ a $\mathbf{v} = \varphi'(\alpha)$. Protože zobrazení $\mathbf{g}(\varphi(\alpha))$ je nulové na intervalu (α_1, α_2), je i jeho derivace nulová. Z řetízkového pravidla

$$\frac{\mathrm{d}\mathbf{g}(\varphi(\alpha))}{\mathrm{d}\alpha} = \mathbf{g}'(\varphi(\alpha))\varphi'(\alpha) = \mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}.$$

Definice

Bod $\mathbf{x} \in \mathbb{R}^n$ je **regulární bod** zobrazení $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$, když $\operatorname{rank} \mathbf{g}'(\mathbf{x}) = m$ (tj. gradienty $\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})$ jsou lin. nezávislé).

Věta

Nechť $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$. Nechť $\mathbf{x}$ je regulární bod.
Pak je vektor $\mathbf{v}$ tečný k množině X v bodě $\mathbf{x}$.

Důkaz: Složitý, z tzv. věty o implicitní funkci.

Tečný a ortogonální prostor

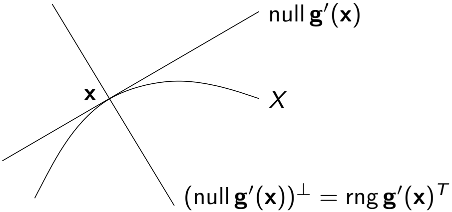
Nechť $\mathbf{x} \in X$ je regulární bod. Dokázali jsme:

$$\mathbf{v} \text{ je tečný v bodě } \mathbf{x} \quad \Longleftrightarrow \quad \mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$$

- Tedy všechny tečné vektory v bodě $\mathbf{x}$ tvoří podprostor $\operatorname{null} \mathbf{g}'(\mathbf{x})$. Je to **tečný prostor** k množině X v bodě $\mathbf{x}$. Má dimenzi $n - m$.
- Jeho ortogonální doplněk

$$(\operatorname{null} \mathbf{g}'(\mathbf{x}))^\perp = \operatorname{rng} \mathbf{g}'(\mathbf{x})^T = \operatorname{span}\{\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})\}$$

  je **ortogonální prostor** k množině X v bodě $\mathbf{x}$. Má dimenzi m.



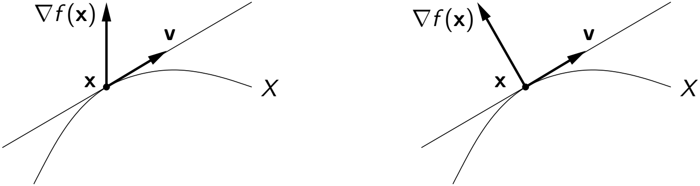
Podmínka prvního řádu

Tvrzení

Nechť $\mathbf{x} \in X$ je lokální extrém funkce f na množině X.
Nechť $\mathbf{v} \in \mathbb{R}^n$ je tečný vektor k množině X v bodě $\mathbf{x}$.
Pak $f'(\mathbf{x})\mathbf{v} = 0$ (tj. $\nabla f(\mathbf{x}) \perp \mathbf{v}$).

Důkaz: Existuje hladká křivka φ na X taková, že $\mathbf{x} = \varphi(\alpha)$ a $\mathbf{v} = \varphi'(\alpha)$.
Protože $\mathbf{x}$ je lokální extrém f na množině X, je lokální extrém také na křivce. Tedy

$$\frac{\mathrm{d}f(\varphi(\alpha))}{\mathrm{d}\alpha} = f'(\varphi(\alpha))\varphi'(\alpha) = f'(\mathbf{x})\mathbf{v} = 0.$$



Věta (Podmínka prvního řádu)

Nechť $\mathbf{x} \in X$ je lokální extrém funkce f na množině X. Nechť $\mathbf{x}$ je regulární bod.
Pak $\nabla f(\mathbf{x}) \in \operatorname{rng} \mathbf{g}'(\mathbf{x})^T = \operatorname{span}\{\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})\}$.

Důkaz: Z Tvrzení máme $f'(\mathbf{x})\mathbf{v} = \mathbf{0}$ pro každý vektor $\mathbf{v}$ tečný k množině X v bodě $\mathbf{x}$.
Tedy $\nabla f(\mathbf{x}) \perp \operatorname{null} \mathbf{g}'(\mathbf{x})$, tedy $\nabla f(\mathbf{x}) \in (\operatorname{null} \mathbf{g}'(\mathbf{x}))^\perp = \operatorname{rng} \mathbf{g}'(\mathbf{x})^T$.

Lagrangeovy multiplikátory, Lagrangeova funkce

Nalezená podmínka

$$\nabla f(\mathbf{x}) \in \operatorname{rng} \mathbf{g}'(\mathbf{x})^T$$

říká, že existují $(\lambda_1, \ldots, \lambda_m) = \boldsymbol{\lambda} \in \mathbb{R}^m$ (**Lagrangeovy multiplikátory**) tak, že

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

Podmínka i s omezeními tvoří soustavu $m + n$ rovnic s $m + n$ neznámými

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

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 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 derivaci L podle $\mathbf{x}$ resp. $\boldsymbol{\lambda}$)

Příklad

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

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 $\operatorname{rank} g'(x) < 1$).

Lagrangeova funkce

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

Podmínky stacionarity

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

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.

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

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

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).

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.

Podmínky druhého řádu [nepovinné]

Věta

Nechť f a $\mathbf{g}$ jsou dvakrát diferencovatelné v bodě $\mathbf{x}$. Nechť $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$.

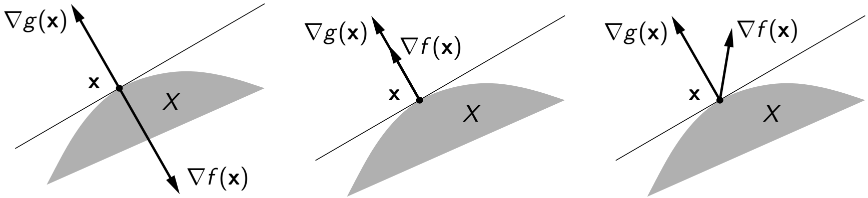
- Je-li $\mathbf{x}$ lokální minimum, pak $L_{\mathbf{xx}}(\mathbf{x}, \boldsymbol{\lambda})$ je positivně semidefinitní na podprostoru $\operatorname{null} \mathbf{g}'(\mathbf{x})$.
- Je-li $L_{\mathbf{xx}}(\mathbf{x}, \boldsymbol{\lambda})$ positivně definitní na podprostoru $\operatorname{null} \mathbf{g}'(\mathbf{x})$, pak $\mathbf{x}$ je ostré lokální minimum.
- Je-li $L_{\mathbf{xx}}(\mathbf{x}, \boldsymbol{\lambda})$ indefinitní na podprostoru $\operatorname{null} \mathbf{g}'(\mathbf{x})$, pak $\mathbf{x}$ není lokální extrém.

- Symbol
$$L_{\mathbf{xx}}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\partial^2 L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}^2} = f''(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \, g_i''(\mathbf{x}) \in \mathbb{R}^{n \times n}$$
značí Hessián funkce $L(\cdot, \boldsymbol{\lambda})$ v bodě $\mathbf{x}$ (tedy podle λ_i nederivujeme).
- Matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je **positivně definitní na podprostoru** $T \subseteq \mathbb{R}^n$, jestliže $\mathbf{v}^T \mathbf{A} \mathbf{v} > 0$ pro každé $\mathbf{v} \in T$. To platí právě když matice $\mathbf{B}^T \mathbf{A} \mathbf{B}$ je positivně definitní, kde sloupce $\mathbf{B}$ tvoří bázi podprostoru T.

Extrémy vázané nerovnostmi

Neformální úvaha pro omezení jednou nerovnicí

Pro $f, g\colon \mathbb{R}^n \to \mathbb{R}$ hledáme lok. extrémy f na množině $X = \{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \leq 0 \,\}$.



x je lokální minimum | **x** je lokální maximum | **x** není lokální extrém

Nechť **x** je lokální minimum.

- Když $g(\mathbf{x}) < 0$, pak **x** je volné lokální minimum a tedy $\nabla f(\mathbf{x}) = \mathbf{0}$.
- Když $g(\mathbf{x}) = 0$, pak **x** je lokální minimum vázané rovností $g(\mathbf{x}) = 0$.
  Pokud $\nabla g(\mathbf{x}) \neq \mathbf{0}$ (regularita), existuje tedy λ tak, že $\nabla f(\mathbf{x}) + \lambda \nabla g(\mathbf{x}) = \mathbf{0}$.
  Musí být $\lambda > 0$, neboť jinak by se $f(\mathbf{x})$ dalo zmenšit pohybem **x** dovnitř X.

Shrnuto: Pokud buď $g(\mathbf{x}) < 0$ nebo $\nabla g(\mathbf{x}) \neq \mathbf{0}$, existuje $\lambda \geq 0$ tak, že

$$\nabla f(\mathbf{x}) + \lambda \nabla g(\mathbf{x}) = \mathbf{0}$$
$$\lambda g(\mathbf{x}) = 0$$

Úloha s omezeními ve tvaru rovností a nerovností

Jsou dána spojitě diferencovatelná $f\colon \mathbb{R}^n \to \mathbb{R}$, $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{h}\colon \mathbb{R}^n \to \mathbb{R}^\ell$.
Hledáme minimum funkce f na množině

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

kde $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ označuje, že $g_i(\mathbf{x}) \leq 0$ pro každé $i = 1, \dots, m$.

Tedy řešíme úlohu

$$\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0} \,\}$$

neboli

$$\begin{aligned} \text{min} \quad & f(x_1, \dots, x_n) \\ \text{za podmínek} \quad & g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m \\ & h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, \ell \end{aligned}$$

Názvosloví: Omezení $g_i(\mathbf{x}) \leq 0$ je v bodě $\mathbf{x} \in X$

- **aktivní**, jestliže $g_i(\mathbf{x}) = 0$,
- **neaktivní**, jestliže $g_i(\mathbf{x}) < 0$.

Karush-Kuhn-Tucker (KKT) podmínky prvního řádu

Pro $\mathbf{x} \in X$ označme množinu indexů omezení aktivních v bodě $\mathbf{x}$ jako

$$I(\mathbf{x}) = \{\, i = 1, \dots, m \mid g_i(\mathbf{x}) = 0 \,\}.$$

Definice

Bod $\mathbf{x}$ je **regulární bod omezení**, jestliže gradienty

$$\nabla h_1(\mathbf{x}), \dots, \nabla h_m(\mathbf{x}), \qquad \nabla g_i(\mathbf{x}),\ i \in I(\mathbf{x})$$

jsou lineárně nezávislé.

Věta (KKT podmínka prvního řádu)

Nechť $\mathbf{x} \in X$ je lokální minimum a navíc regulární bod.
Pak existují $\boldsymbol{\lambda} \geq \mathbf{0}$ a $\boldsymbol{\mu} \in \mathbb{R}^\ell$ tak, že

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

Důkaz vynecháme (vychází z neformální úvahy na minulém slajdu).

Všimněte si: $\boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = \mathbf{0}$ platí právě když $\lambda_i g_i(\mathbf{x}) = 0$ pro každé $i = 1, \dots, m$ (neboť $\boldsymbol{\lambda} \geq \mathbf{0}$ a $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$).