Optimalizace

Lokální extrémy vázané rovnostmi

Tomáš Kroupa, Mirko Navara

21. 5. 2024

FEL ČVUT

Úloha s omezeními ve tvaru rovností

Hledáme extrémy funkce f vázané rovnostmi

$$g_i(x_1, \ldots, x_n) = 0\,, \quad i = 1, \ldots, m\,.$$

Úloha

$$\begin{aligned} \min \quad & f(x_1, \ldots, x_n) \\ \text{za podmínek} \quad & \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

Předpokládáme f, $\mathbf{g}$ spojitě diferencovatelné.

Speciální případy

- afinní kritérium, afinní omezení - není co řešit,
- obecné kritérium, afinní omezení - to bude rozcvička a inspirace,
- afinní kritérium, obecné omezení - ponecháme jako speciální případ následujícího:
- obecné kritérium, obecné omezení - to je cíl.

Speciální případ: afinní omezení

Úloha

$$\begin{aligned} \min \quad & f(x_1, \dots, x_n) \\ \text{za podmínek} \quad & \mathbf{g}(\mathbf{x}) = \mathbf{C}\,\mathbf{x} - \mathbf{d} = \mathbf{0}\,, \end{aligned}$$

kde $\mathbf{C} \in \mathbb{R}^{m \times n}$, $\mathbf{d} \in \mathbb{R}^m$.

$$\mathbf{g}'(\mathbf{x}) = \mathbf{C}\,, \quad \text{neboli } \nabla \mathbf{g}(\mathbf{x}) = \mathbf{C}^\mathsf{T}\,.$$

Tvrzení

Pro každý lokální extrém $\mathbf{x}$ této úlohy existuje $\boldsymbol{\lambda} \in \mathbb{R}^m$ tak, že

$$f'(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\,\mathbf{C} = \mathbf{0}\,, \qquad \text{tj.}$$

$$f'(\mathbf{x}) = -\boldsymbol{\lambda}^\mathsf{T}\,\mathbf{C} = -\boldsymbol{\lambda}^\mathsf{T}\,\mathbf{g}'(\mathbf{x})\,, \quad \text{neboli } \nabla f(\mathbf{x}) = -\mathbf{C}^\mathsf{T}\,\boldsymbol{\lambda}\,.$$

Lagrangeovy multiplikátory

Vektor $\boldsymbol{\lambda}$ bude (i v obecnějším případě) vektor Lagrangeových multiplikátorů.

Na jejich znaménku nezáleží; můžeme si je vynásobit jakýmkoli nenulovým číslem (i každý jiným).

Mohli bychom je i vynásobit regulární maticí a použít $\mathbf{R}\boldsymbol{\lambda}$ místo $\boldsymbol{\lambda}$; množina řešení takto upravených omezení se tím nezmění.

Jejich hodnoty poslouží jen během výpočtu, na jeho konci nás nezajímají.

Ještě speciálnější případ:
Úloha na nejmenší normu řešení nehomogenní soustavy

Hledáme řešení soustavy $\mathbf{C}\mathbf{x} = \mathbf{d}$ s nejmenší normou. Řešíme úlohu

$$\min \left\{ \tfrac{1}{2} \|\mathbf{x}\|^2 \mid \mathbf{C}\mathbf{x} = \mathbf{d} \right\}.$$

$$f(\mathbf{x}) = \tfrac{1}{2} \|\mathbf{x}\|^2 = \tfrac{1}{2} \mathbf{x}^\mathsf{T} \mathbf{x}, \qquad f'(\mathbf{x}) = \mathbf{x}^\mathsf{T}.$$

Podmínky optimality (s využitím předchozího)

$$\mathbf{x} = f'(\mathbf{x})^\mathsf{T} = -\mathbf{C}^\mathsf{T}\boldsymbol{\lambda}, \qquad \mathbf{C}\mathbf{x} = \mathbf{d}.$$

Má-li $\mathbf{C}$ lineárně nezávislé řádky, pak optimálním řešením je

$$\mathbf{x} = \mathbf{C}^\mathsf{T}(\mathbf{C}\mathbf{C}^\mathsf{T})^{-1}\mathbf{d}.$$

Méně speciální případ:
Úloha nejmenších čtverců s lineárními omezeními

Hledáme řešení úlohy nejmenších čtverců pro soustavu $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ s lineárními omezeními $\mathbf{C}\,\mathbf{x} = \mathbf{d}$. Tedy řešíme úlohu

$$\min\left\{\tfrac{1}{2}\,\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \mid \mathbf{C}\,\mathbf{x} = \mathbf{d}\right\}.$$

$$f(\mathbf{x}) = \tfrac{1}{2}\,\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 = \tfrac{1}{2}\,(\mathbf{A}\,\mathbf{x} - \mathbf{b})^{\mathsf{T}}\,(\mathbf{A}\,\mathbf{x} - \mathbf{b}),$$
$$f'(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\,\mathbf{A}^{\mathsf{T}}\,\mathbf{A} - \mathbf{b}^{\mathsf{T}}\,\mathbf{A} = -\lambda^{\mathsf{T}}\,\mathbf{C},$$

$$\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{x} + \mathbf{C}^{\mathsf{T}}\lambda = \mathbf{A}^{\mathsf{T}}\mathbf{b}.$$

Podmínky optimality

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

Obecně: Metoda Lagrangeových multiplikátorů

Vyjdeme z lineární aproximace

$$\mathbf{g}(\mathbf{x}) \approx \underbrace{\mathbf{g}(\mathbf{x}_0)}_{\mathbf{0}} + \mathbf{g}'(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0) = \mathbf{g}'(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0)$$

$$= \underbrace{\mathbf{g}'(\mathbf{x}_0)}_{\mathbf{C}}\,\mathbf{x} - \underbrace{\mathbf{g}'(\mathbf{x}_0)\,\mathbf{x}_0}_{\mathbf{d}} = \mathbf{C}\,\mathbf{x} - \mathbf{d}\,.$$

Jako dříve, zavedeme Lagrangeovy multiplikátory $\boldsymbol{\lambda} := (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ a budeme řešit soustavu rovnic

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

kde $\mathbf{C} := \mathbf{g}'(\mathbf{x})$ tedy

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

To lze interpretovat také jako hledání stacionárních bodů Lagrangeovy funkce $L\colon\ \mathbb{R}^{n+m} \to \mathbb{R}$

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

Tečné vektory

Množina všech přípustných řešení úlohy s omezeními ve tvaru rovností:

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

Definice

Vektor $\mathbf{t} \in \mathbb{R}^n$ je tečný k množině X v bodě $\mathbf{x} \in X$, pokud je v $\mathbf{x}$ tečným vektorem nějaké hladké křivky ležící v X.
Tečný prostor $T_\mathbf{x}$:= prostor vektorů tečných k X v bodě $\mathbf{x}$.

Často je $T_\mathbf{x}$ lineární (a obvykle dim $T_\mathbf{x} = n - m$), ale nemusí být.

$$\forall \mathbf{t} \in T_\mathbf{x}\ \forall i : \mathbf{t} \perp g_i'(\mathbf{x})^\mathsf{T} = \nabla g_i(\mathbf{x})\,,$$

jinak by $\mathbf{t}$ nebyl tečný k vrstevnici funkce g_i. Stručněji

$$T_\mathbf{x} \subseteq \operatorname{null} \mathbf{g}'(\mathbf{x})\,.$$

Prostor generovaný gradienty omezení

$$
\begin{aligned}
G_{\mathbf{x}} &:= \operatorname{span}\{g_1'(\mathbf{x})^{\mathsf{T}}, \dots, g_m'(\mathbf{x})^{\mathsf{T}}\} = \operatorname{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\} = \\
&= \operatorname{rng} \mathbf{g}'(\mathbf{x})^{\mathsf{T}} = \operatorname{rng} \nabla \mathbf{g}(\mathbf{x}) = \\
&= \{\mathbf{g}'(\mathbf{x})^{\mathsf{T}} \boldsymbol{\lambda} \mid \boldsymbol{\lambda} \in \mathbb{R}^m\} \subseteq T_{\mathbf{x}}^{\perp} \,.
\end{aligned}
$$

$G_{\mathbf{x}}$ je lineární prostor parametrizovaný Lagrangeovými multiplikátory.

$$\forall \mathbf{v} \in G_{\mathbf{x}} \ \exists \boldsymbol{\lambda} \in \mathbb{R}^m : \mathbf{v} = \mathbf{g}'(\mathbf{x})^{\mathsf{T}} \boldsymbol{\lambda} \,.$$

$\dim G_{\mathbf{x}} \leq m$, $G_{\mathbf{x}} \perp T_{\mathbf{x}}$, může být $G_{\mathbf{x}} \subsetneq T_{\mathbf{x}}^{\perp}$ (a to bude problém).

Dvě interpretace Lagrangeových multiplikátorů

Lagrangeovy multiplikátory můžeme chápat jako

- argumenty Lagrangeovy funkce $L\colon \mathbb{R}^{n+m} \to \mathbb{R}$, $(\mathbf{x}, \boldsymbol{\lambda}) \mapsto L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T} \mathbf{g}(\mathbf{x})$,
- parametry funkce ${}^{\boldsymbol{\lambda}}L\colon \mathbb{R}^{n} \to \mathbb{R}$, $\quad \mathbf{x} \mapsto {}^{\boldsymbol{\lambda}}L(\mathbf{x}) = L(\mathbf{x}, \boldsymbol{\lambda})$.

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

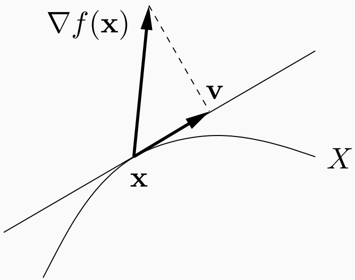
$${}^{\boldsymbol{\lambda}}L'(\mathbf{x}) = f'(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T} \mathbf{g}'(\mathbf{x}).$$

Lagrangeova funkce na tečném prostoru

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ má v $\mathbf{x}$ vázaný extrém na X.

$$\forall \mathbf{v} \in T_{\mathbf{x}} : \mathbf{v} \perp f'(\mathbf{x})^{\mathsf{T}} = \nabla f(\mathbf{x})\,,$$

jinak by ze směrů $\pm\mathbf{v}$ byl jeden vzestupný a jeden sestupný.



Lagrangeova funkce na tečném prostoru

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ má v $\mathbf{x}$ vázaný extrém na X.

$$\forall \mathbf{v} \in T_{\mathbf{x}} : \mathbf{v} \perp f'(\mathbf{x})^{\mathsf{T}} = \nabla f(\mathbf{x})\,,$$

jinak by ze směrů $\pm\mathbf{v}$ byl jeden vzestupný a jeden sestupný.
Stručněji $T_{\mathbf{x}} \perp f'(\mathbf{x})^{\mathsf{T}}$, též

$$T_{\mathbf{x}} \subseteq \operatorname{null} f'(\mathbf{x})\,,$$

což spolu s předchozím

$$T_{\mathbf{x}} \subseteq \operatorname{null} \mathbf{g}'(\mathbf{x})\,,$$

dává

$$T_{\mathbf{x}} \subseteq \operatorname{null}(f'(\mathbf{x}) + \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{g}'(\mathbf{x})) = \operatorname{null} {}^{\boldsymbol{\lambda}}L'(\mathbf{x})$$

pro všechna $\boldsymbol{\lambda} \in \mathbb{R}^m$.

Lagrangeova funkce na prostoru generovaném gradienty omezení

$$G_{\mathbf{x}} := \operatorname{span}\{g_1'(\mathbf{x})^{\mathsf{T}}, \ldots, g_m'(\mathbf{x})^{\mathsf{T}}\} = \{\mathbf{g}'(\mathbf{x})^{\mathsf{T}}\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \in \mathbb{R}^m\} \subseteq T_{\mathbf{x}}^{\perp},$$

$$\forall \mathbf{v} \in G_{\mathbf{x}}\ \exists \boldsymbol{\lambda} \in \mathbb{R}^m : \mathbf{v} = \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{g}'(\mathbf{x}).$$

Jestliže $f'(\mathbf{x})^{\mathsf{T}} \in G_{\mathbf{x}}$, pak pro $\mathbf{v} := -f'(\mathbf{x})^{\mathsf{T}} \in G_{\mathbf{x}}$

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

pro nějaké $\boldsymbol{\lambda} \in \mathbb{R}^m$; $^{\lambda}L$ má nulovou derivaci ve všech směrech

$$\mathbf{t} \in T_{\mathbf{x}},$$

$$\mathbf{v} \in G_{\mathbf{x}}, \qquad \text{a tedy i}$$

$$\Uparrow \qquad c_1\,\mathbf{t} + c_2\,\mathbf{v} \in T_{\mathbf{x}} + G_{\mathbf{x}}, \qquad c_1, c_2 \in \mathbb{R}.$$

Jestliže $T_{\mathbf{x}} + G_{\mathbf{x}} = \mathbb{R}^n$, pak řešení $\mathbf{x}$ můžeme najít jako stacionární bod funkce $^{\lambda}L$.

Regulární body

Odvodili jsme postup pro hledání vázaných extrémů v bodech $\mathbf{x} \in X = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$, splňujících $T_\mathbf{x} + G_\mathbf{x} = \mathbb{R}^n$, kde

$$T_\mathbf{x} = \text{prostor vektorů tečných k } X \text{ v bodě } \mathbf{x}\,,$$
$$G_\mathbf{x} = \operatorname{span}\{g_1'(\mathbf{x})^\mathsf{T}, \dots, g_m'(\mathbf{x})^\mathsf{T}\}\,.$$

Potřebnou podmínku zformulujeme pro obecné body:

Definice

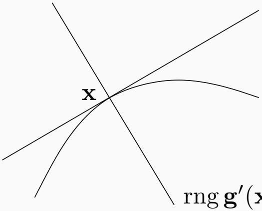
Bod $\mathbf{x} \in \mathbb{R}^n$ je regulární bod zobrazení $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$, jestliže zobrazení $\mathbf{g}$ je v bodě $\mathbf{x}$ spojitě diferencovatelné a Jacobiho matice $\mathbf{g}'(\mathbf{x})$ má lineárně nezávislé řádky.

- Stačí pouze $\mathbf{x}$, $\mathbf{g}$.
- Nepotřebujeme f.
- Nestačí $\mathbf{x}$, X.

Tečný a ortogonální prostor v regulárních bodech

V regulárním bodě $\mathbf{x} \in X$ zobrazení $\mathbf{g}$ je:

- tečný prostor $T_{\mathbf{x}} = \operatorname{null} \mathbf{g}'(\mathbf{x})$
- ortogonální prostor $T_{\mathbf{x}}^{\perp} = (\operatorname{null} \mathbf{g}'(\mathbf{x}))^{\perp} = \operatorname{rng} \mathbf{g}'(\mathbf{x})^{\mathsf{T}}$



$$\operatorname{null} \mathbf{g}'(\mathbf{x}) = (\operatorname{span}\{\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})\})^{\perp}$$

$$X = \{\, \mathbf{x} \mid \mathbf{g}(\mathbf{x}) = \mathbf{0} \,\} = \{\, \mathbf{x} \mid g_1(\mathbf{x}) = \cdots = g_m(\mathbf{x}) = 0 \,\}$$

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

Podmínky prvního řádu pro vázané extrémy

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

Věta

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$, $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ jsou spojitě diferencovatelné v regulárním bodě $\mathbf{x}$ zobrazení $\mathbf{g}$. Jestliže $\mathbf{x}$ je lokální extrém funkce f za podmínky $\mathbf{g}(\mathbf{x}) = \mathbf{0}_m$, pak existuje $\boldsymbol{\lambda} \in \mathbb{R}^m$, splňující

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

neboli

$$L'(\mathbf{x}, \boldsymbol{\lambda}) = \left(f'(\mathbf{x}) + \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{g}'(\mathbf{x}),\ \mathbf{g}(\mathbf{x})^{\mathsf{T}}\right) = \mathbf{0}_{n+m}\,.$$

Podmínky druhého řádu pro volné extrémy (opakování)

Věta

Nechť $X \subseteq \mathbb{R}^n$, $\mathbf{x} \in X$ je vnitřní bod množiny X a $f\colon \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná v $\mathbf{x}$. Je-li $\mathbf{x}$ stacionární bod funkce f, pak:

$$
\begin{gathered}
f''(\mathbf{x}) \text{ je pozitivně definitní} \\
\Downarrow \quad \not\Uparrow \\
\mathbf{x} \text{ je lokální minimum} \\
\Downarrow \quad \not\Uparrow \\
f''(\mathbf{x}) \text{ je pozitivně semidefinitní}
\end{gathered}
$$

Matice $\mathbf{C}$ je pozitivně definitní na lineárním prostoru Y, jestliže $\forall \mathbf{x} \in Y \setminus \{\mathbf{0}\} : \mathbf{x}^T \mathbf{C} \mathbf{x} > 0$.

Podmínky druhého řádu pro vázané extrémy

Věta

Nechť $\mathbf{x} \in X$ a funkce $f\colon\ \mathbb{R}^n \to \mathbb{R}, \mathbf{g}\colon\ \mathbb{R}^n \to \mathbb{R}^m$ jsou dvakrát diferencovatelné v $\mathbf{x}$. Nechť $\mathbf{x}$ je regulární bod zobrazení $\mathbf{g}$ a stacionární bod Lagrangeovy funkce ${}^{\lambda}L(\mathbf{x}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x})$. Pak

$${}^{\lambda}L''(\mathbf{x}) \text{ je pozitivně definitní na null}\,\mathbf{g}'(\mathbf{x})$$

$$\Downarrow \quad \not\Uparrow$$

$$\mathbf{x} \text{ je vázané minimum } f \quad \cancel{(\text{a volné minimum } {}^{\lambda}L)}$$

$$\Downarrow \quad \not\Uparrow$$

$${}^{\lambda}L''(\mathbf{x}) \text{ je pozitivně semidefinitní na null}\,\mathbf{g}'(\mathbf{x})$$

kde ${}^{\lambda}L''(\mathbf{x}) = L_{\mathbf{x},\mathbf{x}}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\partial^2}{\partial \mathbf{x}^2} L(\mathbf{x}, \boldsymbol{\lambda}) = f''(\mathbf{x}) + \sum_{i=1}^m \lambda_i\, g_i''(\mathbf{x})\,.$

Na $L_{\boldsymbol{\lambda},\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\partial^2}{\partial \boldsymbol{\lambda}^2} L(\mathbf{x}, \boldsymbol{\lambda}),\ \ L_{\mathbf{x},\boldsymbol{\lambda}}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{\partial^2}{\partial \mathbf{x} \partial \boldsymbol{\lambda}} L(\mathbf{x}, \boldsymbol{\lambda})$ nezáleží.

Možný postup: sekvenční kvadratické programování (SQP)

Newtonova metoda minimalizace funkce (opakování)

Řešíme rovnici $f'(\mathbf{x}) = \mathbf{0}$ iterací $\mathbf{x}_{k+1} = \mathbf{x}_k - f''(\mathbf{x}_k)^{-1} f'(\mathbf{x}_k)^\mathsf{T}$.

Použijeme pro $f(\mathbf{x}) := L(\mathbf{x}, \boldsymbol{\lambda})$:

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

$$L'(\mathbf{x}, \boldsymbol{\lambda}) = \big(f'(\mathbf{x}) + \underbrace{\boldsymbol{\lambda}^\mathsf{T} \mathbf{g}'(\mathbf{x})}_{\sum_{i=1}^m \lambda_i g_i'(\mathbf{x})} \,,\; \mathbf{g}(\mathbf{x})^\mathsf{T} \big) \,,$$

$$L''(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} f''(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i''(\mathbf{x}) & \mathbf{g}'(\mathbf{x})^\mathsf{T} \\ \mathbf{g}'(\mathbf{x}) & \mathbf{0} \end{bmatrix} ,$$

$$\begin{bmatrix} \mathbf{x}_{k+1} - \mathbf{x}_k \\ \boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k \end{bmatrix} = -L''(\mathbf{x}_k, \boldsymbol{\lambda}_k)^{-1} L'(\mathbf{x}_k, \boldsymbol{\lambda}_k)^\mathsf{T} =$$

$$= - \begin{bmatrix} f''(\mathbf{x}_k) + \sum_{i=1}^m \lambda_i g_i''(\mathbf{x}_k) & \mathbf{g}'(\mathbf{x}_k)^\mathsf{T} \\ \mathbf{g}'(\mathbf{x}_k) & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} f'(\mathbf{x}_k)^\mathsf{T} + \mathbf{g}'(\mathbf{x}_k)^\mathsf{T} \boldsymbol{\lambda}_k \\ \mathbf{g}(\mathbf{x}_k) \end{bmatrix} .$$

Co s body, které nejsou regulární?

- Pokud je jich málo, vyzkoušíme je všechny.
- Pokud je některé $g_i'(\mathbf{x}) = \mathbf{0}$, zkusme ho rozložit na součin a vyšetřit každý činitel zvlášť.
- Chybějící tečné vektory můžeme najít sami a zpracovat stejně jako gradienty omezujících funkcí.
- Místo toho můžeme odvodit jiné omezující funkce, určující stejný obor X.

Speciální případ: afinní omezení (opak.)

$$\mathbf{g}(\mathbf{x}) = \mathbf{C}\,\mathbf{x} - \mathbf{d}\,, \qquad \mathbf{g}'(\mathbf{x}) = \mathbf{C}\,, \qquad \nabla \mathbf{g}(\mathbf{x}) = \mathbf{C}^\top\,.$$

1. Jsou-li řádky matice $\mathbf{C}$ (neboli $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$) lineárně nezávislé, pak jsou všechny body regulární.
2. Jsou-li řádky matice $\mathbf{C}$ (neboli $\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})$) lineárně závislé, pak není *žádný* bod regulární a záleží na $\mathbf{d}$:
   - $X = \emptyset$: není co řešit.
   - $X \neq \emptyset$: můžeme z $[\mathbf{C} \;\; \mathbf{d}]$ vynechat LZ řádky a pokračovat dle bodu 1, nebo na to nedbat a řešit původní soustavu

$$\begin{aligned} f'(\mathbf{x}) &= \boldsymbol{\lambda}^\top \mathbf{C}\,, \;\; \text{neboli } \nabla f(\mathbf{x}) = \mathbf{C}^\top \boldsymbol{\lambda}\,, \\ \mathbf{C}\,\mathbf{x} &= \mathbf{d}\,. \end{aligned}$$

(To pro nelineární podmínky nelze.)

Prokletí dimenzionality

Pro velký počet omezujících funkcí je těžké zajistit nebo ověřit regularitu.

Můžeme doufat, že množina bodů, v nichž je regularita porušena, je „řídká“, takže i když přes ni projdeme, nezůstaneme v ní.

Alternativa: Metoda projektovaných gradientů

Někdy dovedeme k libovolnému bodu z $\mathbb{R}^n$ *snadno* najít „nejbližší“ přípustný z X („projekci“ na X), např. pro

$$\begin{aligned} X_1 &= \{\mathbf{x} \mid \mathbf{a} \le \mathbf{x} \le \mathbf{b}\}\,, \\ X_2 &= \{\mathbf{x} \mid \|\mathbf{x}\| \le r\}\,, \\ X_3 &= \{\mathbf{x} \mid \|\mathbf{x}\| = r\}\,. \end{aligned}$$

Pak můžeme střídat kroky:

1. Postup ve směru $\nabla f(\mathbf{x})$ bez respektování omezujících podmínek.
2. Projekce na X.

Problém: V blízkosti extrému nás 1. krok zavádí dál od cíle.

Reference

- T. Werner. Optimalizace (kapitola 8). Elektronická skripta. FEL ČVUT, 17. 10. 2022. `https://cw.fel.cvut.cz/wiki/_media/courses/b0b33opt/opt.pdf`
- S. Boyd, L. Vandenberghe. Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares. Cambridge University Press, 2018.
- J. Hamhalter a J. Tišer. Diferenciální počet funkcí více proměnných. FEL ČVUT, 2005. `https://math.fel.cvut.cz/en/people/tiser/dipocet.pdf`
- M. Navara. Matematika spojitého světa. Elektronická skripta. FEL ČVUT, 2011. `https://cmp.felk.cvut.cz/~navara/MSS/MSS_print.pdf`