Optimalizace

Lokální extrémy vázané rovnostmi

Tomáš Kroupa, Mirko Navara

2. 5. 2024

FEL ČVUT

Úloha s omezeními ve tvaru rovností

Hledáme extrémy funkce f vázané rovnostmi

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

Úloha

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

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 - probereme až po následujícím:
- obecné kritérium, obecné omezení - to je cíl.

Speciální případ: afinní omezení

Úloha

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

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

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

Tvrzení

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

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

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

$$\mathbf{x} = \mathbf{C}^\mathsf{T} \boldsymbol{\lambda}, \qquad \mathbf{C}\mathbf{x} = \mathbf{d}.$$

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

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

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

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

Lagrangeovy multiplikátory $\boldsymbol{\lambda} \coloneqq (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$

Lagrangeova funkce $L\colon \mathbb{R}^{n+m} \to \mathbb{R}$

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

Podmínky optimality

$$\frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \mathbf{x}} = f'(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T} \mathbf{g}'(\mathbf{x}) = \mathbf{0}\,, \text{ neboli } \nabla f(\mathbf{x}) + \nabla g(\mathbf{x})\, \boldsymbol{\lambda} = \mathbf{0}\,,$$

$$\frac{\partial L(\mathbf{x}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \mathbf{g}(\mathbf{x})^\mathsf{T} = \mathbf{0}\,.$$

Za příslušných předpokladů existuje $\boldsymbol{\lambda} \in \mathbb{R}^m$ splňující $L'(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$, tedy $(\mathbf{x}, \boldsymbol{\lambda})$ je stacionárním bodem funkce L.

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