Optimalizace a teorie her

Úvod

Martin Bohata

Katedra matematiky
FEL ČVUT v Praze
bohata@math.feld.cvut.cz

Základní informace

Stránky předmětu:
https://moodle.fel.cvut.cz/courses/A8B01OGT

Obsah kurzu:

1. Základy konvexní analýzy v $\mathbb{R}^n$
2. Podmínky optimality
3. Dualita
4. Lineární a kvadratické programování
5. Vybrané numerické metody v optimalizaci
6. Úvod do teorie her

Kde lze potkat optimalizaci nebo teorii her?

L. Euler: „Nothing takes place in the world whose meaning is not that of some maximum or minimum.“

- Matematika
- Fyzika
- Řízení
- Zpracování signálu
- Zpracování obrazu
- Komunikace
- Elektronika
- Energetika
- Umělá inteligence
- Ekonomie

$\vdots$

Formulace úlohy

Definice

Nechť $f : D \to \mathbb{R}$ a $M \subseteq D$.

- Řekneme, že f nabývá v $\hat{x} \in M$ minima (resp. maxima) na M, jestliže pro každé $x \in M$ je $f(\hat{x}) \leq f(x)$ (resp. $f(x) \leq f(\hat{x})$).
- Nabývá-li f v $\hat{x}$ minima (resp. maxima), pak $f(\hat{x})$ se nazývá minimum (resp. maximum) funkce f na M a $\hat{x}$ se nazývá *bod minima* (resp. bod maxima) funkce f na M.
- Extrémem funkce f na M rozumíme její minimum nebo maximum na M. Body, ve kterých funkce f nabývá extrému na M, nazýváme body extrému funkce f na M.

- Úmluva: U pojmů z předchozí definice budeme vynechávat „na M“, jestliže $M = D$.

Formulace úlohy

Značení:

- $\min_{x \in M} f(x)$ ... minimum funkce f na M;
- $\max_{x \in M} f(x)$ ... maximum funkce f na M;
- $\operatorname{argmin}_{x \in M} f(x)$ ... množina všech bodů minima f na M;
- $\operatorname{argmax}_{x \in M} f(x)$ ... množina všech bodů maxima f na M.

Definice

Pod *optimalizační úlohou* rozumíme jakoukoli z následujících úloh:

(U1) Pro $f : D \to \mathbb{R}$ a $M \subseteq D$ nalezněte $\operatorname{argmin}_{x \in M} f(x)$.

(U2) Pro $f : D \to \mathbb{R}$ a $M \subseteq D$ nalezněte $\operatorname{argmax}_{x \in M} f(x)$.

- Zápis úlohy (U1): minimalizujte f na M.
- Zápis úlohy (U2): maximalizujte f na M.
- Nalézt všechna řešení optimalizační úlohy je většinou obtížné. Často se tak spokojíme i s nalezením jediného řešení.

Formulace úlohy

Terminologie:

- f ... cílová funkce;
- M ... přípustná množina;
- Prvky z M ... přípustné body;
- Prvky z $\operatorname{argmin}_{x \in M} f(x)$ ... řešení úlohy (U1);
- Prvky z $\operatorname{argmax}_{x \in M} f(x)$ ... řešení úlohy (U2).

Příklad

1. $\operatorname{argmin}_{x \in [0,1]} x = \{0\}$.
2. $\operatorname{argmin}_{x \in (0,1)} x = \emptyset$.
3. $\operatorname{argmin}_{x \in [-\pi,\pi]} |\sin x| = \{-\pi, 0, \pi\}$.
4. $\operatorname{argmin}_{x \in M} 1 = M$.

Souvislost (U1) a (U2)

Tvrzení

Nechť $f : D \to \mathbb{R}$, $M \subseteq D$ *a* $\hat{x} \in M$.

1. $\hat{x} \in \operatorname{argmin}_{x \in M} f(x)$ *právě tehdy, když* $\hat{x} \in \operatorname{argmax}_{x \in M} (-f(x))$.
2. *Nabývá-li* f *v nějakém bodě z* M *minima na* M, *pak*

$$\min_{x \in M} f(x) = -\max_{x \in M}(-f(x)).$$

Důkaz: Viz přednáška. ■

Dvě důležité oblasti optimalizace jsou:

- Optimalizace v $\mathbb{R}^n$.
- Variační počet.

Optimalizační úlohy v $\mathbb{R}^n$

Je dána cílová funkce $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, množina $\Omega \subseteq D$ a množina

$$M = \{x \in \Omega \,|\, g_1(x) \leq 0, \ldots, g_k(x) \leq 0, h_1(x) = 0, \ldots, h_l(x) = 0\},$$

kde $g_1, \ldots, g_k, h_1, \ldots, h_l$ jsou reálné funkce definované na Ω.
Úlohu (U1) s cílovou funkcí f a přípustnou množinou M budeme zapisovat ve tvaru:

$$\begin{array}{ll} \text{minimalizujte} & f(x) \\ \text{za podmínek} & g_i(x) \leq 0,\ i = 1, \ldots, k, \\ & h_j(x) = 0,\ j = 1, \ldots, l, \\ & x \in \Omega. \end{array}$$

Obdobně zapisujeme odpovídající úlohu (U2).

Optimalizační úlohy v $\mathbb{R}^n$

- Terminologie:
  $x \in \Omega$ ... přímé omezení;
  $g_i(x) \leq 0$ ... omezení ve tvaru nerovnosti;
  $h_j(x) = 0$ ... omezení ve tvaru rovnosti.
- Úmluva: Je-li $\Omega = D$, pak přímé omezení budeme v zápisu vynechávat.
- Omezení $g(x) = h(x)$ lze vždy zapsat pomocí dvou omezení $g(x) \leq h(x)$ a $g(x) \geq h(x)$.
- Omezení $g(x) \geq h(x)$ můžeme přepsat do tvaru $-g(x) \leq -h(x)$.
- Omezení $g(x) \leq h(x)$ lze psát ve tvaru $G(x) \leq 0$, kde $G(x) = g(x) - h(x)$.

Optimalizační úlohy v $\mathbb{R}^n$

Příklad

$$\begin{aligned} \text{minimalizujte} \quad & x^2 + 1 \\ \text{za podmínek} \quad & \frac{3}{x} \leq 1, \\ & x \in \mathbb{N}. \end{aligned}$$

Vidíme, že přípustná množina je $M = \mathbb{N} \setminus \{1, 2\}$.

(Je snadné ukázat, že $\operatorname{argmin}_{x \in M} x^2 + 1 = \{3\}$.)

Optimalizační úlohy v $\mathbb{R}^n$

Příklad

$$\begin{aligned} \text{maximalizujte} \quad & \ln x \\ \text{za podmínek} \quad & x \leq 5, \\ & \cos(\pi x) = 1. \end{aligned}$$

Přípustná množina tedy je $M = \{2, 4\}$.

(Zřejmě $\operatorname{argmax}_{x \in M} \ln x = \{4\}$.)

Základní klasifikace oblastí optimalizace v $\mathbb{R}^n$

- Nepodmíněná optimalizace ... přípustná množina je $\mathbb{R}^n$.
- Podmíněná optimalizace ... přípustná množina je vlastní podmnožina množiny $\mathbb{R}^n$.
- Konvexní optimalizace ... přípustná množina M je konvexní a cílová funkce je konvexní (resp. konkávní) na M v případě minimalizační (resp. maximalizační) úlohy.
- Lineární programování ... přípustná množina je konvexní polyedrická množina a cílová funkce je afinní.
- Kvadratické programování ... přípustná množina je konvexní polyedrická množina a cílová funkce je kvadratická.
- Celočíselné programování ... přípustná množina je průnik konvexní polyedrické množiny a množiny $\mathbb{Z}^n$ a cílová funkce je afinní.

Opakování

- Prostor $\mathbb{R}^n$ ... lineární prostor všech uspořádaných n-tic reálných čísel. Prvky $\mathbb{R}^n$ budeme považovat za „sloupcové vektory“, tj. je-li $x \in \mathbb{R}^n$, pak píšeme
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$
- Skalární součin na $\mathbb{R}^n$:
$$\langle x, y \rangle = y^T x = \sum_{i=1}^{n} x_i y_i.$$
- Eukleidovská norma na $\mathbb{R}^n$:
$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Tvrzení (Cauchyova-Schwarzova nerovnost)

Pro každé $x, y \in \mathbb{R}^n$ platí

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| \, .$$

Důkaz: Vynecháváme. ∎

- Pro $\Omega \subseteq \mathbb{R}^n$ budeme symbolem $C(\Omega)$ označovat množinu všech reálných spojitých funkcí na Ω.
- Je-li $\Omega \subseteq \mathbb{R}^n$ otevřená a $k \in \mathbb{N}$, pak $C^k(\Omega)$ označuje množinu všech reálných funkcí, které mají na Ω spojité všechny parciální derivace řádu k.

Věta (Weierstrassova věta)

Je-li neprázdná množina $M \subseteq \mathbb{R}^n$ kompaktní (tj. omezená a uzavřená) a $f \in C(M)$, pak existuje bod minima a bod maxima funkce f na M.

Důkaz: Vynecháváme. ∎

Opakování

Definice

Nechť $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $M \subseteq D$.

- Řekneme, že f nabývá v $\hat{x} \in M$ lokálního minima (resp. ostrého lokálního minima) na M, jestliže existuje $\delta > 0$ tak, že $f(\hat{x}) \leq f(x)$ pro všechna $x \in M \cap U(\hat{x}, \delta)$ (resp. $f(\hat{x}) < f(x)$ pro všechna $x \in M \cap P(\hat{x}, \delta)$).
- Nabývá-li f v $\hat{x}$ lokálního minima (resp. ostrého lokálního minima), pak $f(\hat{x})$ se nazývá lokální minimum (resp. ostré lokální minimum) funkce f na M a $\hat{x}$ se nazývá *bod lokálního minima* (resp. bod ostrého lokálního minima) funkce f na M.

- Úmluva: U pojmů z předchozí definice budeme vynechávat „na M“, jestliže $M = D$.
- Analogicky definujeme pojmy v případě lokálního maxima.

Opakování

- $$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$
- $$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}.$$
- Symetrická matice $A \in \mathbb{M}_n(\mathbb{R})$ je pozitivně semidefinitní, jestliže $\langle Ax, x \rangle \geq 0$ pro každé $x \in \mathbb{R}^n$.
- Symetrická matice $A \in \mathbb{M}_n(\mathbb{R})$ je pozitivně definitní, jestliže $\langle Ax, x \rangle > 0$ pro každé nenulové $x \in \mathbb{R}^n$.

Opakování

Věta (O Taylorově polynomu 1. a 2. řádu)

Nechť $\Omega \subseteq \mathbb{R}^n$ *je otevřená a* $x \in \Omega$. *Potom platí:*

1. *Jestliže* $f \in C^1(\Omega)$, *pak existuje* $\omega : \mathbb{R}^n \to \mathbb{R}$ *tak, že* $\lim\limits_{h \to 0} \omega(h) = 0$ *a pro každé* $h \in \mathbb{R}^n$ *splňující* $x + h \in \Omega$ *je*
$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \|h\| \, \omega(h)$.
2. *Jestliže* $f \in C^2(\Omega)$ *a* $h \in \mathbb{R}^n$ *splňuje* $[x, x+h] \subseteq \Omega$, *pak existuje* $\xi \in \{x + \lambda h \,|\, \lambda \in (0,1)\}$ *tak, že*
$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \left\langle \nabla^2 f(\xi) h, h \right\rangle$.
3. *Jestliže* $f \in C^2(\Omega)$, *pak existuje* $\omega : \mathbb{R}^n \to \mathbb{R}$ *tak, že* $\lim\limits_{h \to 0} \omega(h) = 0$ *a pro každé* $h \in \mathbb{R}^n$ *splňující* $x + h \in \Omega$ *je*
$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \left\langle \nabla^2 f(x) h, h \right\rangle + \|h\|^2 \omega(h)$.

Důkaz: Vynecháváme. ∎

Motivační příklady – optimalizace v $\mathbb{R}^n$

Příklad (Proložení bodů přímkou)

Naměřené hodnoty závislosti napětí na rezistoru na protékajícím proudu jsou:

Proud [A]	0,02	0,04	0,06	0,08
Napětí [V]	5,0	10,1	15,8	21,1

Chceme-li z naměřených hodnot určit odpor R metodou nejmenších čtverců, musíme vyřešit úlohu:

minimalizujte

$$f(R) = (0,02R-5)^2+(0,04R-10,1)^2+(0,06R-15,8)^2+(0,08R-21,1)^2$$

na $\mathbb{R}$.

Příklad (Úloha o dietě – příprava salátu)

Jedna porce salátu musí obsahovat alespoň 4g vlákniny, 15mg vitamínu C a 20mg hořčíku. Suroviny na přípravu salátu jsou mrkev, okurka a rajče. Ceny příslušných surovin a zastoupení vyžadovaných látek v těchto surovinách jsou:

	Mrkev	Okurka	Rajče
Vláknina [g/kg]	29	7	16
Vitamín C [mg/kg]	45	137	187
Hořčík [mg/kg]	180	90	80
Cena [Kč/kg]	25	20	30

Cílem je najít takové složení salátu, aby cena za porci byla co nejnižší při splnění předepsaných podmínek.

Motivační příklady – optimalizace v $\mathbb{R}^n$

Příklad (Úloha o dietě – příprava salátu)

Označme x_M množství mrkve, x_O množství okurek a x_R množství rajčat (vše v kilogramech), která jsou potřebná k přípravě jedné porce salátu. Potom zadaný problém vede na optimalizační úlohu

$$\begin{aligned}
\text{minimalizujte} \quad & 25x_M + 20x_O + 30x_R \\
\text{za podmínek} \quad & 29x_M + 7x_O + 16x_R \geq 4, \\
& 45x_M + 137x_O + 187x_R \geq 15, \\
& 180x_M + 90x_O + 80x_R \geq 20, \\
& x_M, x_O, x_R \geq 0.
\end{aligned}$$