Optimalizace

Konvexní optimalizační úlohy

Tomáš Kroupa Tomáš Werner

2024

Fakulta elektrotechnická
ČVUT v Praze

Obsah

1. Konvexní úlohy obecně a příklady
2. Jak řešit konvexní úlohy
3. Nekonvexní úlohy

Konvexní úlohy obecně a příklady

Konvexní optimalizace (1)

Definice

Nechť $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$.
Úloha konvexní optimalizace je optimalizační úloha

$$\min \; f(\mathbf{x}) \qquad \text{za podmínky } \mathbf{x} \in X.$$

- Různé třídy úloh lišící se tvarem účelové funkce a omezeními
- Mnoho numerických metod i specializovaných řešičů
- Výhodné matematické vlastnosti

Konvexní optimalizace (2)

Vlastnostni konvexní množiny přípustných řešení X

- otevřená/uzavřená
- omezená/neomezená
- kompaktní (omezená a uzavřená)

Vlastnosti konvexní účelové funkce f na X

- omezená/neomezená
- nemusí být diferencovatelná
- spojitá na vnitřku X, ale ne nutně na hranici množiny X

Vlastnosti konvexních úloh

$$\min \ f(\mathbf{x}) \qquad \text{za podmínky } \mathbf{x} \in X$$

Věta

Pro konvexní úlohu platí:

1. Každé lokální minimum funkce f na X je globální.
2. Množina všech optimálních řešení je konvexní.
3. Pokud je f ryze konvexní, existuje nejvýše jedno optimum.

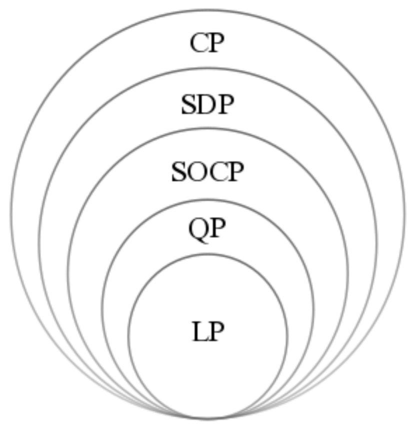
Konvexní úloha ve standardním tvaru

- $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní
- $g_1, \ldots, g_m\colon \mathbb{R}^n \to \mathbb{R}$ jsou konvexní
- $h_1, \ldots, h_\ell\colon \mathbb{R}^n \to \mathbb{R}$ jsou afinní

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

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

Třídy konvexních úloh - schéma



Wikipedia

Třídy konvexních úloh

- Lineární programování
  f, g_i, h_i afinní
- Kvadratické programování (QP)
  f konvexní kvadratická a g_i, h_i afinní
- Kvadratické programování s kvadrat. omezeními (QCQP)
  f, g_i konvexní kvadratické a h_i afinní
- Programování na kuželu druhého řádu
- Semidefinitní programování
- Kónické programování

Kvadratické programování

- f je kvadratická
- g_i, h_i jsou afinní

Maticově

$$\begin{aligned} \min \quad & \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \mathbf{b}^\mathsf{T}\mathbf{x} \\ \text{za podmínek} \quad & \mathbf{C}\mathbf{x} \le \mathbf{d} \\ & \mathbf{E}\mathbf{x} = \mathbf{f} \end{aligned}$$

- $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^n$
- $\mathbf{C} \in \mathbb{R}^{m\times n}$, $\mathbf{d} \in \mathbb{R}^m$
- $\mathbf{E} \in \mathbb{R}^{\ell\times n}$, $\mathbf{f} \in \mathbb{R}^\ell$

Je to konvexní úloha, právě když **A** je pozitivně semidefinitní.

Kvadratické programování – příklady konvexních úloh

Řešení přeurčené soustavy $\mathbf{Ax} = \mathbf{b}$

$$\min\{\|\mathbf{Ax} - \mathbf{b}\|_2^2 \mid \mathbf{x} \in \mathbb{R}^n\}$$

Řešíme převodem na lineární rovnice (normální rovnice).

Řešení přeurčené soustavy $\mathbf{Ax} = \mathbf{b}$ s lineárními omezeními

$$\min\{\|\mathbf{Ax} - \mathbf{b}\|_2^2 \mid \mathbf{Cx} = \mathbf{d},\ \mathbf{x} \in \mathbb{R}^n\}$$

Lze převést na řešení soustavy lineárních rovnic (Lagrange).

Řešení přeurčené soustavy $\mathbf{Ax} = \mathbf{b}$ s intervalovými omezeními

$$\min\{\|\mathbf{Ax} - \mathbf{b}\|_2^2 \mid \mathbf{c} \leq \mathbf{x} \leq \mathbf{d},\ \mathbf{x} \in \mathbb{R}^n\}$$

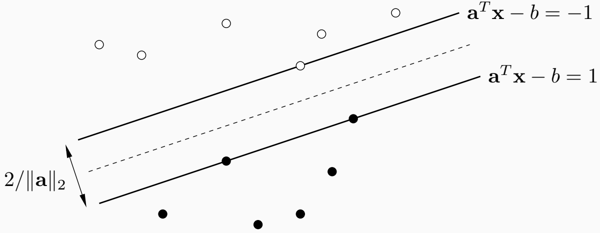
Kvadratické programování – Support vector machines (SVM)

Pro m bodů $(\mathbf{x}_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$ hledáme oddělující nadrovinu.
Tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ tak, aby

$$y_i(\mathbf{a}^\mathsf{T}\mathbf{x}_i - b) > 0, \qquad i = 1, \ldots, m.$$

Vydělením $(\mathbf{a}, b)$ vhodným kladným číslem je to ekvivalentní

$$y_i(\mathbf{a}^\mathsf{T}\mathbf{x}_i - b) \geq 1, \qquad i = 1, \ldots, m.$$



$$\min \{\|\mathbf{a}\|_2^2 \mid y_i(\mathbf{a}^\mathsf{T}\mathbf{x}_i - b) \geq 1,\ i = 1, \ldots, m\}$$

Kvadratické programování s kvadratickými omezeními

- f, g_i jsou kvadratické
- h_i jsou afinní

Je to konvexní úloha, právě když jsou funkce f a g_i konvexní.

Kam umístit záchrannou stanici?

Hledáme lokaci $\mathbf{x} \in \mathbb{R}^2$ pro záchrannou službu tak, aby *nejdelší vzdálenost* k místům $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^2$ byla minimální:

$$\text{Minimalizuj} \quad \max_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|_2 \quad \text{z.p. } \mathbf{x} \in \mathbb{R}^2$$

Ekvivalentně – hledáme nejmenší kruh obsahující zadané body:

$$\text{Minimalizuj} \quad y \quad \text{z.p. } \|\mathbf{x} - \mathbf{a}_i\|_2^2 \leq y, \quad i = 1, \dots, m.$$

Jak řešit konvexní úlohy

Od gradientní metody k projektovanému gradientu

- Pokud je $X = \mathbb{R}^n$, úloha konvexní optimalizace je bez omezení
- Globální minimum $\mathbf{x}$ pro konvexní diferencovatelnou funkci f splňuje podmínku $\nabla f(\mathbf{x}) = \mathbf{0}$
- Lze použít gradientní metodu s iterací $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k)$

Modifikace pro konvexní úlohu s omezeními

- Globální minimum $\mathbf{x}$ na X typicky nesplňuje $\nabla f(\mathbf{x}) = \mathbf{0}$
- Přímočaré použití gradientní metody není možné, protože typicky $\mathbf{x}_{k+1} \notin X$
- Nabízí se však nalézt bod z X nejblíže $\mathbf{x}_{k+1}$

Eukleidovská projekce

Definice

Nechť $X \subseteq \mathbb{R}^n$ je uzavřená konvexní množina. Eukleidovská projekce bodu $\mathbf{z} \in \mathbb{R}^n$ na X je optimální řešení úlohy

$$\min_{\mathbf{x} \in X} \|\mathbf{z} - \mathbf{x}\|_2^2.$$

- Úloha má jediné optimální řešení, které značíme $P_X(\mathbf{z})$
- Navíc je optimální řešení charakterizováno podmínkou

$$(\mathbf{z} - P_X(\mathbf{z}))^T(\mathbf{x} - P_X(\mathbf{z})) \leq 0 \qquad \text{pro každé } \mathbf{x} \in X$$

- Ekvivalentně: v bodě $P_X(\mathbf{z})$ existuje opěrná nadrovina H množiny X s normálovým vektorem $P_X(\mathbf{z}) - \mathbf{z}$

Eukleidovská projekce – příklady

Lineární podprostor X

Eukleidovská projekce $P_X(\mathbf{z})$ bodu $\mathbf{z}$ je ortogonální projekce na X.

Jednotková koule $X = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq 1\}$

$$P_X(\mathbf{z}) = \begin{cases} \mathbf{z} & \|\mathbf{z}\|_2 \leq 1, \\ \frac{\mathbf{z}}{\|\mathbf{z}\|_2} & \text{jinak.} \end{cases}$$

Hyperoktant $X = \{\mathbf{x} \in \mathbb{R}^n \mid x_1, \ldots, x_n \geq 0\}$

$$P_X(\mathbf{z}) = \max(\mathbf{z}, \mathbf{0}).$$

Projektovaný gradient

- Gradientní metoda slouží k minimalizaci funkce bez omezení
- Při jejím použití na konvexní úlohu s omezeními můžeme v některé iteraci dostat nepřípustné řešení
- To však můžeme promítnout a získat tak přípustné řešení

Iterace metody projektovaného gradientu

$$\mathbf{x}_{k+1} = P_X\left(\mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k)\right)$$

kde $\alpha_k > 0$ je délka kroku

Nekonvexní úlohy

Různé nekonvexní úlohy (1)

Úloha vedoucí na spektrální rozklad

Minimalizuj $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ za podmínky $\|\mathbf{x}\|_2^2 = 1$, kde $\mathbf{x} \in \mathbb{R}^n$.

Úloha celočíselného lineárního programování

Minimalizuj $\mathbf{a}^\mathsf{T}\mathbf{x}$ za podmínek $\mathbf{A}\mathbf{x} \geq \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$, kde $\mathbf{x} \in \mathbb{Z}^n$.

Úloha s omezeními ve tvaru rovností (obecná)

Minimalizuj $f(\mathbf{x})$ za podmínky $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{x} \in \mathbb{R}^n$.

Obtížná úloha

Maximalizuj konvexní $f(\mathbf{x})$ na konvexní množině X.

Různé nekonvexní úlohy (2)

Shlukování

- K bodům $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ hledáme body $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ tak, aby součet vzdáleností bodů $\mathbf{a}_i$ k nejbližším bodům $\mathbf{x}_j$ byl minimální
- Minimalizujeme tak funkci

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^{m} \min_{j=1}^{k} \|\mathbf{a}_i - \mathbf{x}_j\|$$

na množině vektorů $(\mathbf{x}_1, \ldots, \mathbf{x}_k) \in \mathbb{R}^{kn}$