Optimalizace

9. Úvod do lineárního programování

Tom Werner, Tom Kroupa

FEL ČVUT

Úloha lineárního programování (LP)

Lineární program (LP) (neboli úloha **lineární optimalizace**) je hledání extrémů lineární funkce vázaných lineárními nerovnicemi a rovnicemi.

Zopakujme:

- Lineární funkce je $c_1x_1 + \cdots + c_nx_n = \mathbf{c}^T\mathbf{x}$
- Lineární rovnice je $a_1x_1 + \cdots + a_nx_n = \mathbf{a}^T\mathbf{x} = b$
- Lineární nerovnice je jedna ze dvou možností:

$$a_1x_1 + \cdots + a_nx_n = \mathbf{a}^T\mathbf{x} \le b$$
$$a_1x_1 + \cdots + a_nx_n = \mathbf{a}^T\mathbf{x} \ge b$$

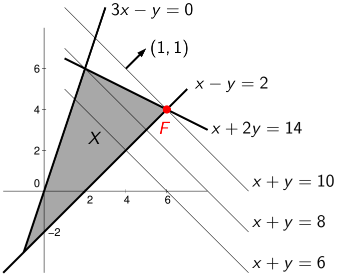
Obecná úloha spojité optimalizace

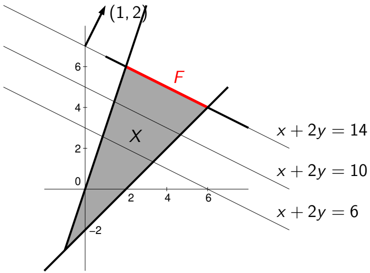
$$\min\{\, f(\mathbf{x}) \mid \mathbf{g}(\mathbf{x}) \le \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0},\ \mathbf{x} \in \mathbb{R}^n \,\}$$

je LP, právě když funkce $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$ je lineární
a zobrazení $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}$ a $\mathbf{h}(\mathbf{x}) = \mathbf{d} - \mathbf{C}\mathbf{x}$ jsou afinní.

Příklad

$$\begin{array}{rl} \max & x + y \\ \text{za podm.} & x + 2y \le 14 \\ & 3x - y \ge 0 \\ & x - y \le 2 \end{array} \qquad\qquad \begin{array}{rl} \max & x + 2y \\ \text{za podm.} & x + 2y \le 14 \\ & 3x - y \ge 0 \\ & x - y \le 2 \end{array}$$





Tvrzení: Pro každou úlohu LP mohou nastat jen tři případy:

- úloha má (alespoň jedno) optimální řešení,
- úloha je **nepřípustná** (množina přípustných řešení je prázdná),
- úloha je **neomezená** (kritérium lze bez porušení omezení libovolně zlepšovat).

Ekvivalentní úpravy

Snadná pozorování:

- maximalizaci $\mathbf{c}^T\mathbf{x}$ lze nahradit minimalizací $-\mathbf{c}^T\mathbf{x}$,
- rovnici $\mathbf{a}^T\mathbf{x} = b$ lze nahradit dvěma nerovnicemi $\mathbf{a}^T\mathbf{x} \leq b, \;\; \mathbf{a}^T\mathbf{x} \geq b$,
- nerovnici $\mathbf{a}^T\mathbf{x} \leq b$ lze nahradit nerovnicí $-\mathbf{a}^T\mathbf{x} \geq -b$.

Každou úlohu LP lze tedy vždy převést na tvar

$$\begin{aligned} \text{min} \quad & c_1x_1 + \cdots + c_nx_n \\ \text{za podmínek} \quad & a_{i1}x_1 + \cdots + a_{in}x_n \geq b_i, \quad i = 1, \ldots, m \end{aligned}$$

neboli

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b},\ \mathbf{x} \in \mathbb{R}^n \,\}$$

Ekvivalentní úprava na rovnicový tvar

Každou úlohu LP lze také převést do **rovnicového tvaru**

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$$

- Nerovnice $\mathbf{a}^T\mathbf{x} \geq b$ nahradíme $\mathbf{a}^T\mathbf{x} - u = b$ kde $u \geq 0$ je **slacková proměnná**,
- Proměnné $x \in \mathbb{R}$ nahradíme $x = x^+ - x^-$ kde $x^+ \geq 0$ a $x^- \geq 0$.

Příklad: V úloze

$$\begin{array}{rrcrcr}
\max & x & + & y & & \\
\text{za podm.} & x & + & 2y & \leq & 14 \\
 & 3x & - & y & \geq & 0 \\
 & x & - & y & \leq & 2
\end{array}$$

přidáme slacky a z max uděláme min:

$$\begin{array}{rrcrcrcr}
\min & -x & - & y & & & & \\
\text{za podmínek} & x & + & 2y & + & u & = & 14 \\
 & 3x & - & y & - & v & = & 0 \\
 & x & - & y & + & w & = & 2 \\
 & & & & & u, v, w & \geq & 0
\end{array}$$

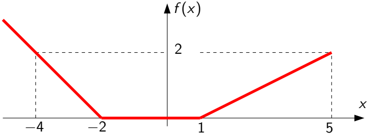
Pak rozdělíme neomezené proměnné:

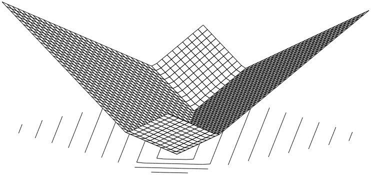
$$\begin{array}{rrcrcrcrcrcr}
\min & -x_+ & + & x_- & - & y_+ & + & y_- & & & & \\
\text{za podmínek} & x_+ & - & x_- & + & 2y_+ & - & 2y_- & + & u & = & 14 \\
 & 3x_+ & - & 3x_- & - & y_+ & + & y_- & - & v & = & 0 \\
 & x_+ & - & x_- & - & y_+ & + & y_- & + & w & = & 2 \\
 & & & & & & x_+, x_-, y_+, y_-, u, v, w & & & & \geq & 0
\end{array}$$

Minimalizace maxima afiních funkcí

Nechť $f\colon \mathbb{R}^n \to \mathbb{R}$ je maximum afinních funkcí:

$$f(\mathbf{x}) = \max_{i=1}^{k}(\mathbf{c}_i^T\mathbf{x} + d_i)$$





$f(x) = \max\{-x-2,\ 0,\ \frac{1}{2}(x-1)\}$ $\qquad$ $f(x,y) = \frac{1}{2}\max\{x+y-1,\ -x+y,\ x-y,\ -x-y-1\}$

Transformace na úlohu LP ($X \subseteq \mathbb{R}^n$ je množina přípustných řešení):

$$\begin{aligned}
\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X \,\} &= \min\{\, y \mid (\mathbf{x}, y) \in X \times \mathbb{R},\ f(\mathbf{x}) \le y \,\} \\
&= \min\{\, y \mid (\mathbf{x}, y) \in X \times \mathbb{R},\ \max_i(\mathbf{c}_i^T\mathbf{x} + d_i) \le y \,\} \\
&= \min\{\, y \mid (\mathbf{x}, y) \in X \times \mathbb{R},\ \mathbf{c}_i^T\mathbf{x} + d_i \le y\ \forall i \,\}
\end{aligned}$$

neboť

$$\max_i a_i \le b \iff a_i \le b\ \ \forall i$$

Příklad

Úloha

$$\begin{aligned} \min \quad & \max\{\, 3x + 4y + 1,\ 2x - 3y \,\} \\ \text{za podm.} \quad & x + 2y \le 14 \\ & 3x - \ \ y \ge 0 \\ & x - \ \ y \le 2 \end{aligned}$$

je ekvivalentní lineárnímu programu

$$\begin{array}{rl} \min & z \\ \text{za podm.} & \begin{array}{rcl} 3x + 4y - z & \le & -1 \\ 2x - 3y - z & \le & 0 \\ x + 2y & \le & 14 \\ 3x - \ \ y & \ge & 0 \\ x - \ \ y & \le & 2 \end{array} \end{array}$$

Příklady použití LP

Optimální výrobní program

Maximalizuj celkový zisk $\mathbf{c}^T\mathbf{x}$ za podmínek $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ a $\mathbf{x} \geq \mathbf{0}$, kde

- a_{ij} = množství suroviny druhu i potřebné na výrobu výrobku druhu j
- b_i = množství suroviny druhu i, které máme k dispozici
- c_j = zisk z vyrobení jednoho výrobku druhu j
- x_j = počet vyrobených výrobků druhu j

Příklad: Pán prodává lupínky a hranolky, které vyrábí z brambor a voleje.

$$\begin{aligned} \text{max} \quad & 120l + 76h \\ \text{za podmínek} \quad & 2l + 1.5h \leq 100 \\ & 0.4l + 0.2h \leq 16 \\ & l, h \geq 0 \end{aligned}$$

Optimální řešení: $l = 20\,\text{kg}$ lupínků a $h = 40\,\text{kg}$ hranolků.

Směšovací (výživová) úloha (diet problem)

Minimalizuj celkovou cenu $\mathbf{c}^T\mathbf{x}$ surovin za podmínek $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ a $\mathbf{x} \geq \mathbf{0}$, kde

- a_{ij} = množství látky druhu i obsažené v jednotkovém množství suroviny druhu j
- b_i = nejmenší požadované množství látky druhu i v konečném produktu
- c_j = jednotková cena suroviny druhu j
- x_j = množství suroviny druhu j

Příklad: Optimální složení zeleninového salátu

	Mrkev	*Bílé zelí*	*Okurka*	**Požadavek**
Vitamín A [mg/kg]	35	0.5	0.28	0.5 mg
Vitamín C [mg/kg]	60	300	80	15 mg
Vláknina [g/kg]	30	20	10	4 g
Cena [Kč/kg]	26	22	60	

$$
\begin{array}{rrrrrrrl}
\text{min} & 26x_1 & + & 22x_2 & + & 60x_3 & & \\
\text{za podmínek} & 35x_1 & + & 0.5x_2 & + & 0.28x_3 & \geq & 0.5 \\
 & 60x_1 & + & 300x_2 & + & 80x_3 & \geq & 15 \\
 & 30x_1 & + & 20x_2 & + & 10x_3 & \geq & 4 \\
 & & & & & x_1, x_2, x_3 & \geq & 0
\end{array}
$$

Optimální řešení: $x_1 \doteq 0.12$, $x_2 \doteq 0.03$, $x_3 = 0$ za cenu 3.59 Kč.

Při požadavku $x_3 \geq 0.1$ (okurka!) je řešení $x_1 \doteq 0.097$, $x_2 \doteq 0.004$, $x_3 = 0.1$ za 8.62 Kč.

Historie: Stigler diet

Dopravní úloha

Máme m skladů a n spotřebitelů.

- a_i = množství ve skladě i
- b_j = množství zboží požadované spotřebitelem j
- c_{ij} = cena dopravy jednotky zboží ze skladu i ke spotřebiteli j
- x_{ij} = množství zboží vezené ze skladu i ke spotřebiteli j

Chceme co nejlevněji rozvézt zboží ze skladů ke spotřebitelům.

$$\begin{aligned}
\text{min} \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \\
\text{za podm.} \quad & \sum_{j=1}^{n} x_{ij} = a_i, \qquad i = 1, \ldots, m \\
& \sum_{i=1}^{m} x_{ij} = b_j, \qquad j = 1, \ldots, n \\
& x_{ij} \geq 0, \qquad i = 1, \ldots, m,\ j = 1, \ldots, n
\end{aligned}$$

Použití LP na přibližné řešení přeurčených lin. soustav

Vektorové normy

Definice

Funkce $\|\cdot\|: \mathbb{R}^n \to \mathbb{R}$ je (vektorová) **norma**, pokud platí:

- $\|\mathbf{x}\| = 0 \implies \mathbf{x} = \mathbf{0}$
- $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ pro každé $\alpha \in \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

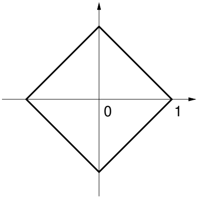
Z axiomů plyne $\|\mathbf{0}\| = 0$ a $\|\mathbf{x}\| \geq 0$ pro každé $\mathbf{x}$.

p-norma

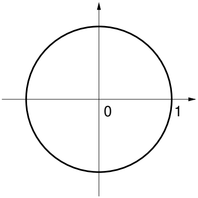
$$\|\mathbf{x}\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p} \qquad (p \geq 1)$$

- $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|$ ('manhattanská' norma)
- $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$ (eukleidovská norma)
- $\|\mathbf{x}\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$ (max-norma, Čebyševova norma)

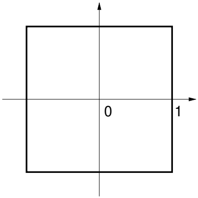
Jednotkové sféry:



$\|\mathbf{x}\|_1 = 1$



$\|\mathbf{x}\|_2 = 1$



$\|\mathbf{x}\|_\infty = 1$

Příklad normy, která není p-norma: $\|\mathbf{A}\mathbf{x}\|$ kde $\|\cdot\|$ je norma a $\mathbf{A}$ má l.n. sloupce

Přibližné řešení přeurčené lineární soustavy v ∞-normě a 1-normě

Minimalizuj funkci $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_p$

- Pro $p \to \infty$ je $f(\mathbf{x}) = \max_{i=1}^{m} |\mathbf{a}_i^T \mathbf{x} - b_i|$. Úlohu lze převést na LP

$$\begin{aligned} \min \quad & z \\ \text{za podm.} \quad & -z \leq \mathbf{a}_i^T \mathbf{x} - b_i \leq z, \quad i = 1, \ldots, m \end{aligned}$$

- Pro $p = 2$ je $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$. Úloha je lin. úloha nejmenších čtverců.

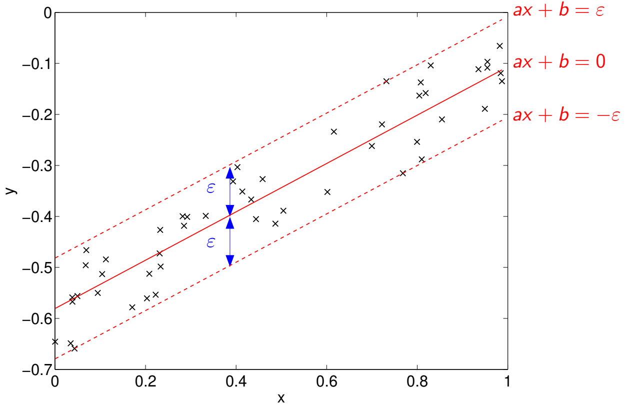
- Pro $p = 1$ je $f(\mathbf{x}) = \sum_{i=1}^{m} |\mathbf{a}_i^T \mathbf{x} - b_i|$. Úlohu lze převést na LP

$$\begin{aligned} \min \quad & z_1 + \cdots + z_m \\ \text{za podm.} \quad & -z_i \leq \mathbf{a}_i^T \mathbf{x} - b_i \leq z_i, \quad i = 1, \ldots, m \end{aligned}$$

Použití řešení v ∞-normě: Prokládání bodů přímkou

Dány body $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^2$.
Hledáme $\boldsymbol{\theta} = (a, b) \in \mathbb{R}^2$, které minimalizuje

$$\varepsilon(\boldsymbol{\theta}) = \max_{i=1}^{m} |ax_i + b - y_i| = \|\mathbf{A}\boldsymbol{\theta} - \mathbf{y}\|_\infty$$



Použití řešení v ∞-normě: Aproximace funkce polynomem

Aproximuj funkci $\sin x$ na intervalu $[0, \frac{\pi}{2}]$ polynomem

$$f(x, \boldsymbol{\theta}) = \theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3.$$

Hledáme $\boldsymbol{\theta} \in \mathbb{R}^4$ které minimalizuje maximální chybu aproximace

$$\varepsilon(\boldsymbol{\theta}) = \max_{x \in [0, \frac{\pi}{2}]} |f(x, \boldsymbol{\theta}) - \sin x|.$$

To neumíme vyřešit (neumíme ani vypočítat hodnoty funkce ε).

Ale když nahradíme interval $[0, \frac{\pi}{2}]$ konečnou množinou $\{ \frac{\pi i}{2m} \mid i = 1, \ldots, m \}$, vede to na minimalizace $\|\mathbf{A}\boldsymbol{\theta} - \mathbf{b}\|_\infty$ kde

$$a_{ij} = \left(\frac{\pi i}{2m}\right)^{j-1}, \quad b_i = \sin \frac{\pi i}{2m}.$$

Použití řešení v 1-normě: Robustní odhad skaláru z m měření

Hledáme odhad $\hat{x} \in \mathbb{R}$ veličiny z m nepřesných měření $(x_1, \ldots, x_m) = \mathbf{x} \in \mathbb{R}^m$.
Hledejme $\hat{x}$ jako argument minima funkce

$$\|\mathbf{x} - \mathbf{1}\hat{x}\|_p = \|(x_1 - \hat{x}, \ldots, x_m - \hat{x})\|_p$$

- Pro $p \to \infty$ minimalizujeme $\max_{i=1}^m |x_i - \hat{x}|$.
  Optimální řešení je $\hat{x} = \frac{1}{2}(\min_i x_i + \max_i x_i)$.

- Pro $p = 2$ minimalizujeme $\sqrt{\sum_i (x_i - \hat{x})^2}$.
  Optimální řešení je aritmetický průměr $\hat{x} = \frac{1}{m} \sum_{i=1}^m x_i$.

- Pro $p = 1$ minimalizujeme $\sum_{i=1}^m |x_i - \hat{x}|$.
  Optimální řešení $\hat{x}$ je **medián** z čísel $x_1, \ldots, x_m$.

Příklad: $\mathbf{x} = (1.02, 1.04, 0.99, 2.03)$. Měření x_4 je **vychýlený bod (outlier)**.

p	∞	2	1
$\hat{x}$	1.51	1.27	1.03

- Pro $p \in \{2, \infty\}$ outlier výrazně ovlivnil (učinil nepoužitelným) odhad $\hat{x}$.
- Pro $p = 1$ je odhad téměř neovlivněn.

Robustní estimátor

Estimátor je zde funkce $(x_1, \ldots, x_m) \mapsto \hat{x}$.

Estimátor je **robustní**, jestliže ho přítomost outlierů příliš neovlivní.

Definice (Bod zlomu (breakdown point) estimátoru)

Bod zlomu (breakdown point) estimátoru je nejmenší počet měření, jejichž změnou lze docílit libovolně velkou změnu odhadu.

Formálně: je to nejmenší k takové, že pro každé $x_1, \ldots, x_{m-k}$ a každé $\delta > 0$ existují $x_{m-k+1}, \ldots, x_m$ tak, že $|\hat{x}| > \delta$.

- Pro $p \in \{2, \infty\}$ má náš estimátor bod zlomu 1 (pro $m \to \infty$ tedy 0%). Tedy není robustní.
- Pro $p = 1$ má náš estimátor (medián) bod zlomu $m/2$ (tedy 50%). Tedy je robustní.

Lze zobecnit na případ, kdy odhad $\hat{\mathbf{x}}$ počítáme jako opt. argument úlohy

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_p.$$

Řídké řešení nedourčené lin. soustavy

- Označme $\|\mathbf{x}\|_0$ počet nenulových složek vektoru $\mathbf{x} \in \mathbb{R}^n$ (pozor, není norma!).
- Vektor $\mathbf{x}$ je **řídký**, jestliže $\|\mathbf{x}\|_0 \ll n$.
- Najdi řídké řešení nedourčené lin. soustavy:

$$\min\{\, \|\mathbf{x}\|_0 \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} = \mathbf{b} \,\}$$

  To je velmi obtížná (NP-těžká) úloha.
- Dobrá aproximace (tzv. **relaxace**) úlohy je

$$\min\{\, \|\mathbf{x}\|_1 \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} = \mathbf{b} \,\}$$