Optimalizace

9. Úvod do lineárního programování

Tomáš Werner, Tomáš Kroupa

26. listopadu 2024

FEL ČVUT

Soustava lineárních nerovnic

- Násobení nerovnice mínus jedničkou: $a \leq b \iff -a \geq -b$

- **Lineární nerovnice** s n proměnnými má jeden z těchto tvarů:

$$a_1x_1 + \cdots + a_nx_n \geq b \qquad \text{neboli} \quad \mathbf{a}^T\mathbf{x} \geq b$$
$$a_1x_1 + \cdots + a_nx_n \leq b \qquad \text{neboli} \quad \mathbf{a}^T\mathbf{x} \leq b$$

  Nerovnice $\mathbf{a}^T\mathbf{x} \leq b$ je ekvivalentní nerovnici $-\mathbf{a}^T\mathbf{x} \geq -b$.

- Soustava m lineárních nerovnic:

$$\begin{array}{ccccc} a_{11}x_1 & + \cdots + & a_{1n}x_n & \geq & b_1 \\ \vdots & & \vdots & & \\ a_{m1}x_1 & + \cdots + & a_{mn}x_n & \geq & b_m \end{array} \qquad \text{neboli} \quad \mathbf{A}\mathbf{x} \geq \mathbf{b}$$

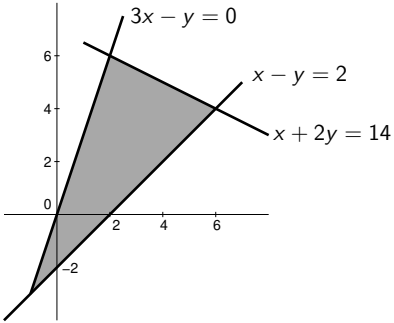
Příklad pro jednu proměnnou ($n = 1$):

$$2x \leq 1$$
$$x \geq -3$$

Snadné: množina řešení je průnik intervalů $(-\infty, \frac{1}{2}] \cap [-3, +\infty) = [-3, \frac{1}{2}]$

Příklad pro dvě proměnné ($n = 2$):

$$\begin{aligned} x + 2y &\leq 14 \\ 3x - \ \ y &\geq \ \ 0 \\ x - \ \ y &\leq \ \ 2 \end{aligned}$$



Množina řešení je průnik polorovin.

Úloha lineárního programování (LP)

Lineární program (LP) (neboli úloha **lineární optimalizace**) je minimalizace nebo maximalizace lineární funkce za podmínek lineárních nerovnic a rovnic.

- **Lineární funkce** je $c_1x_1 + \cdots + c_nx_n$ neboli $\mathbf{c}^T\mathbf{x}$

Obecná úloha spojité optimalizace

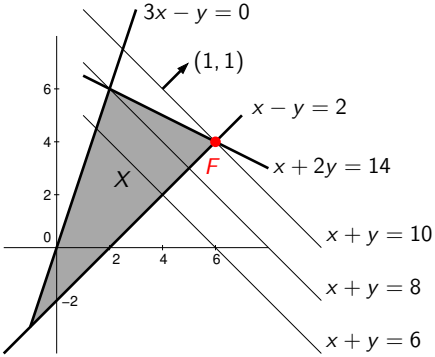
$$\min\{\, f(\mathbf{x}) \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0},\ \mathbf{x} \in \mathbb{R}^n \,\}$$

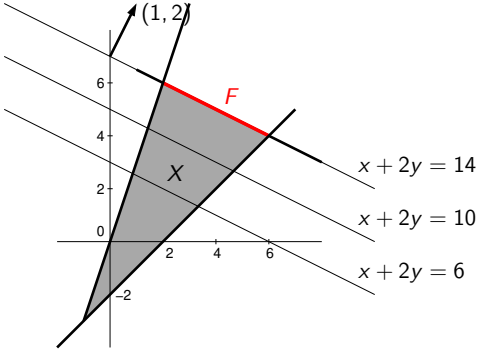
je LP, právě když funkce f je lineární a zobrazení $\mathbf{g}$ a $\mathbf{h}$ jsou afinní.

Příklad

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

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





Řeštelnost lineárních programů

Tvrzení

Pro každou úlohu LP mohou nastat jen tři případy:

- úloha má (alespoň jedno) optimální řešení (tedy účelová funkce je na množině přípustných řešení zdola/shora omezená),
- úloha je **nepřípustná** (množina přípustných řešení je prázdná),
- úloha je **neomezená** (účelová funkce je na množině přípustných řešení zdola/shora neomezená, tedy na ní nemá minimum/maximum).

Důkaz (zatím) vynecháme.

Pro obecné (nelineární) úlohy spojité optimalizace je ještě čtvrtá možnost:

- účelová funkce je na množině přípustných řešení zdola/shora omezená, ale minimum/maximum na ní nemá.

Tvrzení tedy říká, že tato možnost pro lineární programy nemůže nastat.

Ekvivalentní úpravy lineárních programů

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

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

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 vždy 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

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

přidáme slacky a z max uděláme min:

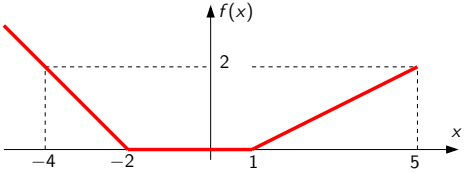
$$\begin{array}{rrcl} \min & -x - y & & \\ \text{za podmínek} & x + 2y + u & = & 14 \\ & 3x - y - v & = & 0 \\ & x - y + w & = & 2 \\ & u, v, w & \geq & 0 \end{array}$$

Pak rozdělíme neomezené proměnné:

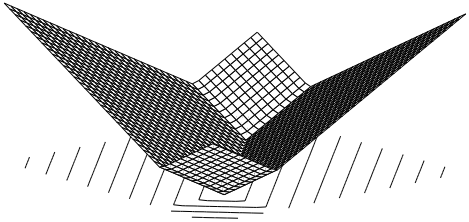
$$\begin{array}{rrcl} \min & -x_+ + x_- - y_+ + y_- & & \\ \text{za podmínek} & x_+ - x_- + 2y_+ - 2y_- + u & = & 14 \\ & 3x_+ - 3x_- - y_+ + y_- - v & = & 0 \\ & x_+ - x_- - y_+ + y_- + w & = & 2 \\ & x_+, x_-, y_+, y_-, u, v, w & \geq & 0 \end{array}$$

Minimalizace maxima afiních funkcí

Nechť $f\colon \mathbb{R}^n \to \mathbb{R}$ je maximum afinních funkcí: $\quad f(\mathbf{x}) = \max\limits_{i=1}^{k}(\mathbf{c}_i^T\mathbf{x} + d_i)$



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



$f(x_1, x_2) = \frac{1}{2}\max\{x_1 + x_2 - 1,\ -x_1 + x_2,\ x_1 - x_2,\ -x_1 - x_2 - 1\}$

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

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

Příklad

Úloha

$$\begin{aligned} \min \quad & \max\{\, 3x_1 + 4x_2 + 1,\ 2x_1 - 3x_2 \,\} \\ \text{za podm.} \quad & x_1 + 2x_2 \le 14 \\ & 3x_1 - x_2 \ge 0 \\ & x_1 - x_2 \le 2 \end{aligned}$$

je ekvivalentní lineárnímu programu

$$\begin{aligned} \min \quad & y \\ \text{za podm.} \quad & 3x_1 + 4x_2 + 1 \le y \\ & 2x_1 - 3x_2 \le y \\ & x_1 + 2x_2 \le 14 \\ & 3x_1 - x_2 \ge 0 \\ & x_1 - x_2 \le 2 \end{aligned}$$

Příklad

Úloha

$$\begin{aligned} \min \quad & |x_1 + x_2| \\ \text{za podm.} \quad & x_1 + 2x_2 \leq 14 \\ & 3x_1 - x_2 \geq 0 \\ & x_1 - x_2 \leq 2 \end{aligned}$$

Všimneme si, že $|x_1 + x_2| = \max\{-x_1 - x_2, x_1 + x_2\}$.

Dále tedy jako předtím.

Příklady použití LP

Optimální výrobní program

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

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

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

Obecná formulace: 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

Směšovací (výživová) úloha (diet problem)

Příklad: Nejlevnější dovolené složení zeleninového salátu

	mrkev	*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	

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

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

Obecná formulace: Minimalizuj celkovou cenu $\mathbf{c}^T\mathbf{x}$ surovin za podmínek $\mathbf{Ax} \geq \mathbf{b}$ 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

Dopravní úloha

Máme m skladů a n spotřebitelů. Rozvez co nejlevněji zboží ze skladů ke spotřebitelům.

- 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

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

Použití LP na přibližné řešení přeurčených lineárních soustav

Vektorové normy

Definice

Funkce $\|\cdot\| \colon \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

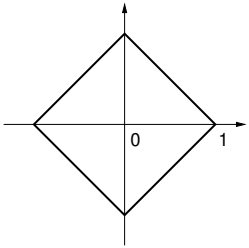
Z axiomů plyne $\|\mathbf{0}\| = 0$ a $\|\mathbf{x}\| \geq 0$ pro každé $\mathbf{x}$.

p-norma

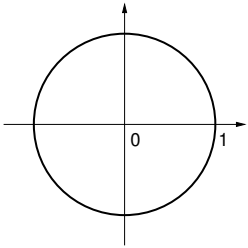
$$\|\mathbf{x}\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p} \qquad (p \geq 1)$$

- $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|$ ('manhattanská' norma)
- $\|\mathbf{x}\|_2 = (x_1^2 + \cdots + x_n^2)^{1/2}$ (eukleidovská norma)
- $\|\mathbf{x}\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$ (max-norma, Čebyševova norma)

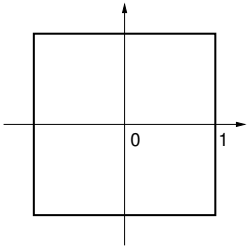
Jednotkové sféry:



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



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



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

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

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

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

- $p = 2$: $\quad f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 = \sqrt{\sum_{i=1}^m (\mathbf{a}_i^T \mathbf{x} - b_i)^2}$ $\qquad$ (to dobře známe)

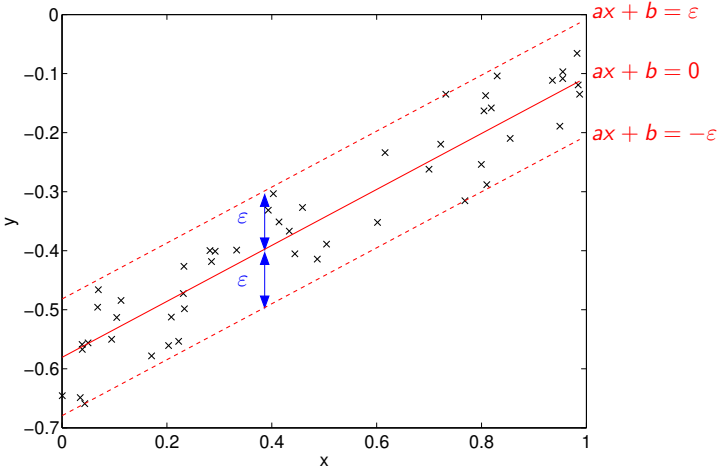
- $p = \infty$: $\quad f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_\infty = \max_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|$
  Úlohu lze převést na LP:
  $$\min\ y \quad \text{za podmínek} \quad -y \le \mathbf{a}_i^T \mathbf{x} - b_i \le y \ \ \forall i = 1, \ldots, m$$
  neboli
  $$\min\{\, y \mid \mathbf{x} \in \mathbb{R}^n,\ y \in \mathbb{R},\ -y\mathbf{1} \le \mathbf{A}\mathbf{x} - \mathbf{b} \le y\mathbf{1} \,\}$$

- $p = 1$: $\quad f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1 = \sum_{i=1}^m |\mathbf{a}_i^T \mathbf{x} - b_i|$
  Úlohu lze převést na LP
  $$\min\ \textstyle\sum_i y_i \quad \text{za podmínek} \quad -y_i \le \mathbf{a}_i^T \mathbf{x} - b_i \le y_i \ \ \forall i = 1, \ldots, m$$
  neboli
  $$\min\{\, \mathbf{1}^T \mathbf{y} \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{y} \in \mathbb{R}^m,\ -\mathbf{y} \le \mathbf{A}\mathbf{x} - \mathbf{b} \le \mathbf{y} \,\}$$

Příklad řešení v ∞-normě: Prokládání bodů přímkou

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

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



Příklad 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

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

Funkci $\varepsilon(\boldsymbol{\theta})$ neumíme minimalizovat.

Diskretizace: nahradíme interval $[0, \frac{\pi}{2}]$ konečnou množinou $\{ \frac{\pi i}{2m} \mid i = 1, \ldots, m \}$.

$$\varepsilon(\boldsymbol{\theta}) \approx \max_{i=1,\ldots,m} \left| f(\tfrac{\pi i}{2m}, \boldsymbol{\theta}) - \sin \tfrac{\pi i}{2m} \right| = \|\mathbf{A}\boldsymbol{\theta} - \mathbf{b}\|_\infty$$

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

Příklad 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

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

- $p = \infty$: $\quad f(\hat{x}) = \max_{i=1}^m |x_i - \hat{x}|, \quad \hat{x} = \frac{1}{2}(\min_i x_i + \max_i x_i)$
- $p = 2$: $\quad f(\hat{x}) = \left(\sum_i (x_i - \hat{x})^2\right)^{1/2}, \quad \hat{x} = \frac{1}{m}\sum_{i=1}^m x_i.$
- $p = 1$: $\quad f(\hat{x}) = \sum_{i=1}^m |x_i - \hat{x}|, \quad \hat{x} = \underset{i=1,\ldots,m}{\text{median}}\, x_i$

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 je nejmenší počet měření, jejichž změnou lze docílit libovolně velkou změnu odhadu.

Formálně: je to nejmenší k takové, že pro každé $x_{k+1}, \ldots, x_m$ a každé $\delta > 0$ existují $x_1, \ldots, x_k$ 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.

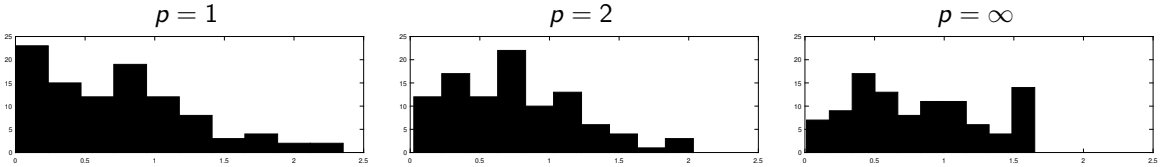
Lze zobecnit na jiné estimátory (např. prokládáme body přímkou).

Histogramy residuí pro různá p

Prvky $\mathbf{A}, \mathbf{b}$ generovány náhodně i.i.d. z Gaussova rozdělení $\mathcal{N}(0, 1)$.

Minimalizujeme $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_p$ pro různá p.

Zobrazeny histogramy residuí $r_i = |\mathbf{a}^T\mathbf{x} - b_i|$ v optimálním řešení $\mathbf{x}$.



Řídké řešení nedourčené lineární soustavy

- Označme $\|\mathbf{x}\|_0$ počet nenulových složek vektoru $\mathbf{x} \in \mathbb{R}^n$ (misnomer, 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

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