## Iterační metody

**Gradientní** →

$$\vec{x}_{k+1} = x_k - \alpha \nabla f(\vec{x_k})$$

Velikost kroku

$f(\varphi) = (\quad)$

$f'(\varphi)$ → řešíme pro $\alpha$ (záporné)

**Newton**

$$\vec{x}_{k+1} = \vec{x_k} - \underbrace{f''(\vec{x_k})^{-1}}_{\text{teoreticky obecně } \alpha} \nabla f(\vec{x_k})$$

**Gauss Newton**

$$x_{k+1} = x_k - g'(x_k)^{+} g(x_k)$$

$$x_{k+1} = x_k - (g'(x_k)^T g'(x_k))^{-1} g'(x_k)^T g(x_k)$$

## Taylorův polynom

$$T_0(x_0) = f(x_0)$$

$$T_1(x_0) = T_0(x_0) + f'(x_0)(x - x_0)$$

$$T_2(x_0) = T_1(x_0) + \tfrac{1}{2} f''(x)(x - x_0)^2$$

**Regulární bod:**
$g(x)$ je v $x$ diferencovatelná a $g'(x)$ má LN řádky

---

$X = \{(x,y,z) \mid x = t, y = 2t, z = 1 - t, t \in \mathbb{R}\}$

$X = \text{span}\{(1,2,-1)\} + (0,0,1)$

(a) najděte $A, b$, aby platilo $X = \{u \mid Au = b\}$

je to $\text{null}\{(1\ 2\ -1)\} \Rightarrow A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix}$

(b) Vzdálenost $\begin{pmatrix}1\\1\\1\end{pmatrix}$ od $X$ $\quad b = A\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$

$$P_x = \frac{aa^T}{a^Ta} = \frac{1}{6}\begin{bmatrix}1 & 2 & -1\\ 2 & 4 & -2\\ -1 & -2 & 1\end{bmatrix} \Rightarrow P_x^{\perp} = I - P$$

Posuň $a$ o $b$ $\Rightarrow \begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix}$ vzdálenost: $\left\|\begin{pmatrix}1/2\\0\\1/2\end{pmatrix}\right\| = 1/\sqrt{2}$

rejekce $(a-b) = P_x^{\perp} \cdot \begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}1/2\\0\\1/2\end{pmatrix}$

$$\sqrt{(Av-b)^T (AA^T)^{-1} (Av-b)}$$

---

## SVD: Matici A napište ve tvaru SVD

$A = \begin{bmatrix}-3 & 1\\ 6 & -2\\ 6 & -2\end{bmatrix}$ $\quad A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$ $\quad A = USV^T$

$A^TA = \begin{bmatrix}81 & -27\\ -27 & 9\end{bmatrix}$ eigen → $\begin{vmatrix}81-\lambda & -27\\ -27 & 9-\lambda\end{vmatrix}$ $\quad \lambda^2 - 90\lambda = 0$

→ $A^TA - \lambda I$ $\quad \lambda(\lambda - 90) = 0$

$\lambda = \begin{bmatrix}0 & 0\\ 0 & 90\end{bmatrix}$ $\quad 0 \quad 90$

↓ (z vlastních čísel, jenž je nenulové má rank = 1)

$\sigma_1 = \sqrt{90} = 3\sqrt{10}$

$A^TA - 90I = \begin{bmatrix}-9 & -27\\ -27 & -81\end{bmatrix} \sim \begin{bmatrix}1 & 3\\ 3 & 9\end{bmatrix}$

řešíme soustavu

$\begin{bmatrix}1 & 3\\ 3 & 9\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$ $\quad$ řešení $\text{span}\begin{pmatrix}3\\-1\end{pmatrix}$

potom $\vec{v_1} = 1/\sqrt{10}\begin{pmatrix}3\\-1\end{pmatrix}$

levý singulární vektor

$u_1 = \frac{1}{\sigma_1} A\vec{v_1} = \frac{1}{3\sqrt{10}} \frac{1}{\sqrt{10}}\begin{bmatrix}-3 & 1\\ 6 & -2\\ 6 & -2\end{bmatrix}\begin{pmatrix}3\\-1\end{pmatrix}$

$= 1/30\begin{pmatrix}-10\\ 20\\ 20\end{pmatrix}$

$A = V\Lambda V^{-1}$ (spektr. $Av = \lambda v$)

$u_1 = \frac{1}{\sigma_1} A v_1^T$

---

$A = \begin{bmatrix}1 & 2 & 3\\ 2 & 4 & 6\\ 1 & 0 & 1\end{bmatrix}$ $\quad x = \begin{pmatrix}1\\1\\1\end{pmatrix}$ $\quad$ proj na a

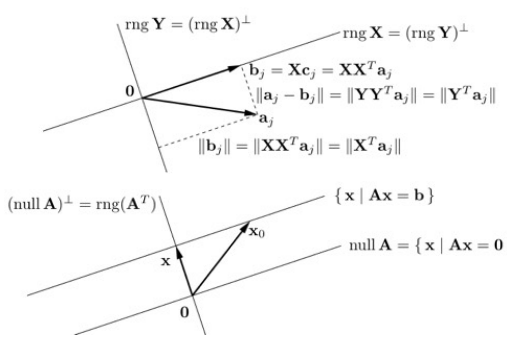
a) $\text{span}(1,1,2)$ $\quad \frac{aa^T}{a^Ta} \cdot x$

b) $\text{rng}\, A$ $\quad x = (c)$

c) $\text{null}\, A = \text{span}\{(\quad)\}$

d) $\text{rng}\, A^T$ $\quad x = (c)$

e) $\text{null}\, A^T$ $\quad \text{span}\{(2,-9,0)\}$




Projektor $j$ $UU^T$, kde $U$ je matice ortonorm. báze, normalizeč ortogon.

---

data $(x_i, y_i) \in \mathbb{R}^2$ $\quad \min$

$p(x) = ax + b$, LSQ $p(x_i) - y_i$

$A = \begin{bmatrix}x_i & 1\end{bmatrix}$ $\quad u = \begin{bmatrix}a\\ b\end{bmatrix}$ $\quad \min \|Au - y\|^2$

$A = \begin{bmatrix}0 & 1\\ 1 & 1\\ 2 & 1\end{bmatrix}$ $\quad y = \begin{pmatrix}0\\2\\2\end{pmatrix}$ $\longrightarrow$ $A^TAu = A^Ty$

$\begin{bmatrix}5 & 3\\ 3 & 3\end{bmatrix} u = \begin{bmatrix}6\\ 4\end{bmatrix}$ $\Rightarrow (a,b) = (1, 1/3)$

optimální hodnota je

$\left\|\begin{bmatrix}0 & 1\\ 1 & 1\\ 2 & 1\end{bmatrix}\begin{bmatrix}1\\ 1/3\end{bmatrix} - \begin{bmatrix}0\\2\\2\end{bmatrix}\right\|^2 = 2/3$

Linprog

**Příklad 15.1.** Následující dvojice lineárních programů je navzájem duální:

$$\begin{array}{lrcl} \min & 2x_1 - 3x_3 + x_4 & & \\ \text{za podm.} & 2x_1 - x_2 + x_3 + 2x_4 &=& 6 \\ & -x_1 + 2x_2 - 3x_3 &\le& 5 \\ & x_1 - x_2 - x_3 - 3x_4 &\ge& 0 \\ & x_1 &\ge& 0 \\ & x_2 &\in& \mathbb{R} \\ & x_3 &\ge& 0 \\ & x_4 &\le& 0 \end{array} \qquad \begin{array}{lrcl} \max & 6y_1 + 5y_2 & & \\ \text{za podm.} & y_1 &\in& \mathbb{R} \\ & y_2 &\le& 0 \\ & y_3 &\ge& 0 \\ & 2y_1 - y_2 + y_3 &\le& 2 \\ & -y_1 + 2y_2 - y_3 &=& 0 \\ & y_1 - 3y_2 - y_3 &\le& -3 \\ & 2y_1 - 3y_3 &\ge& 1 \end{array} \quad \blacklozenge$$

Pro speciální tvary LP se dvojice duálních úloh přehledněji napíše v maticové formě. Např. pro $I_0 = I_- = J_0 = J_- = \emptyset$ obdržíme

$$\begin{array}{lrcl} \min & \mathbf{c}^T\mathbf{x} & & \\ \text{za podm.} & \mathbf{A}\mathbf{x} &\ge& \mathbf{b} \\ & \mathbf{x} &\ge& \mathbf{0} \end{array} \qquad \begin{array}{lrcl} \max & \mathbf{b}^T\mathbf{y} & & \\ \text{za podm.} & \mathbf{y} &\ge& \mathbf{0} \\ & \mathbf{A}^T\mathbf{y} &\le& \mathbf{c} \end{array} \qquad (15.2)$$

4. Máme data $(x_1, y_1), \ldots, (x_n, y_n)$ a snažíme se zjistit jejich závislost pomocí $y \approx f(x; \mathbf{w})$, kde $f(x; \mathbf{w}) = xe^{w_1} + x^2 w_2$. Ve výsledcích nesmí být obecné derivace (například $g'$), ale musí být spočteny. Případné inverze matic se počítat nemusí.
   (a) (2 b) Formulujte hledání neznámých parametrů jako úlohu nelineárních nejmenších čtverců.
   (b) (5 b) Napište iterace gradientní a Newtonovy metody.
   (c) (1 b) Napište, za jaké podmínky selže Newtonova metoda v první iterací.
   (d) (2 b) Napište konkrétní data (tedy specifikujte $n$ a poté $(x_1, y_1), \ldots, (x_n, y_n)$) a počáteční iteraci, pro které Newtonova metoda selže v první iteraci.

**Řešení:**

(a) Jedná se o minimalizaci $\frac{1}{2}\sum_{i=1}^n (x_i e^{w_1} + x_i^2 w_2 - y_i)^2$.

(b) Pro tyto metody potřebujeme spočítat derivaci a Hessián. Derivace má tvar

$$\begin{bmatrix} \sum_{i=1}^n (x_i e^{w_1} + x_i^2 w_2 - y_i) x_i e^{w_1} \\ \sum_{i=1}^n (x_i e^{w_1} + x_i^2 w_2 - y_i) x_i^2 \end{bmatrix}$$

a Hessián se rovná

$$\begin{bmatrix} \sum_{i=1}^n (2x_i e^{w_1} + x_i^2 w_2 - y_i) x_i e^{w_1} & e^{w_1} \sum_{i=1}^n x_i^3 \\ e^{w_1} \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{bmatrix}.$$

Při updatech je nutno myslet na to, že se dělají pro $\mathbf{w}$ a ne pro $x$.

(c) Newtonova metoda selže, pokud je Hessián singulární.

(d) Což nastane například pro $n = 1$ a $x_1 = 0$.

Regulární bod: fce $f$ je v nom grajik a Jacobiho matice v nom má LN řádky.

5. Máme data $\mathbf{a}_1, \ldots, \mathbf{a}_{500} \in \mathbb{R}^3$. Hledáme rovinu procházející počátkem, která minimalizuje součet čtverců vzdáleností k těm bodům. Uvažujeme matici $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_{500}] \in \mathbb{R}^{3\times 500}$ a víme, že matice $\mathbf{A}\mathbf{A}^T$ má vlastní čísla $\sqrt{6}, 1, 3$ a jim odpovídající jednotkové vlastní vektory $\frac{1}{\sqrt{2}}(0,1,1)$, $\frac{1}{\sqrt{6}}(2,1,-1)$, $\frac{1}{\sqrt{3}}(1,-1,1)$.
   (a) (3 b) Formulujte tuto úlohu jako optimalizační problém.
   (b) (6 b) Nalezněte hledanou rovinu a popište ji pomocí báze.
   (c) (1 b) Jaká bude optimální hodnota úlohy (chyba proložení)?

**Řešení:**

(a) Prokládáme body lineárním podprostorem kodimenze 1. Tedy úlohu můžeme vyjádřit jako minimalizaci $\sum_{i=1}^{500} |\mathbf{x}^T\mathbf{a}_i|^2 = \|\mathbf{x}^T\mathbf{A}\|^2 = \mathbf{x}^T\mathbf{A}\mathbf{A}^T\mathbf{x}$, kde $\mathbf{x} \in \mathbb{R}^3$ splňuje $\|\mathbf{x}\| = 1$. Zde $\mathbf{x}$ vyjadřuje směr přímky, která je ortogonálním doplňkem hledané roviny.

(b) Spektrální rozklad $\mathbf{A}\mathbf{A}^T = \mathbf{V}\Lambda\mathbf{V}^T$, kde $\lambda_1 = 1 \le \lambda_2 = \sqrt{6} \le \lambda_3 = 3$ a odpovídající sloupce $\mathbf{V}$ jsou podle toho seřazeny. Řešením $\mathbf{x}$ je vlastní vektor $\mathbf{v}_1 = \frac{1}{\sqrt{6}}(2,1,-1)$ odpovídající nejmenšímu vlastnímu číslu. Hledaná rovina je popsána ortonormální bází $(\mathbf{v}_2, \mathbf{v}_3) = (\frac{1}{\sqrt{2}}(0,1,1), \frac{1}{\sqrt{3}}(1,-1,1))$.

(c) Chyba je optimální hodnota $\mathbf{v}_1^T\mathbf{A}\mathbf{A}^T\mathbf{v}_1 = \lambda_1 = 1$.

Největší stopa $\max\{x^TAx \mid x^Tx=1\}$ pro $k \in n$ pak: $\max\{+r(X^TAX) \mid X^TX=I\}$
opt. hodnota $\lambda_{max}$, opt. řešení $\vec{v}_{max}$ (spektrální rozklad), $A$ symetrická!! opt. hodnota: $k$ největších $\lambda_i$
opt. řešení: $k$ vlastních vektorů příslušných $\lambda_i$
rozbor na $x_1^TAx_1 + x_2^TAx_2 + \ldots$

PCA: co nejlepší proložení bodů dim $m$ prostorem o dim $k$
→ ortonorm báze v nomi $X^{m\times k}$, $X^TX=I$
$(\text{mg}X)^\perp = \text{mg}V$
$b_j = Xc = XX^Ta_j \to \|X^Ta_j\|$
$\text{mg}X$, $a_j$, $a_j - b_j = VV^Ta_j$
→ $\max\{\|X^TA\|^2 \mid X^TX=I\}$ chyba: $\mathcal{E}$ m-k
$\sim \min\{+r(Y^TAA^TY) \mid Y^TY=I\}$ nejmenších v. čísel $AA^T$
$\sim \max\{+r(X^TAA^TX) \mid X^TX=I\}$ relativní: $\frac{\mathcal{E}\ m-k}{\mathcal{E}\ \text{všech}}$
největší stopa: spektr. rozklad $AA^T$
$AA^T = V\Lambda V^T$, $V = [X\ Y]$, $X$ hledaný, $Y$ doplněk

Nejbližší matice nižší hodnosti: $\min\{\|A-B\|^2 \mid B^{m\times n}, \text{rank}B \le k\}$, vektor PCA s $A$, $B = XX^TA$
Přeurčená homogenní soustava: $Ax=0$, $\min\{\|Ax\|^2 \mid x^Tx=1\}$, $x^TA^TAx$, nejmenší $\lambda$
např. proložení kuželosečkou

SVD: $A^{m\times n} = U^{m\times p}S^{p\times p}V^{T\ p\times n} = s_1u_1v_1^T + s_2u_2v_2^T + \ldots$
$U, V$ ortonorm sloupce, $S$ diag, $p = \min\{m,n\}$
$s_1 \ge s_2 \ge \ldots \ge 0$
plné: zbytek $s_i$ jsou 0, matice plný rozšíř
rank-minimální: zahodím všechny $s_i = 0$
$A = USV^T$: $A^TA = VS^2V^T$, $AA^T = US^2U^T$
Nejbližší matice z SVO (Eckart-Young): $B = US_kV^T$, $S_k$ vynuluje $p-k$ nejmenších $s_i$

- → pokud $\mathbf{x}, \mathbf{y}$ jsou přípustná řešení primární a duální úlohy a $\mathbf{c}^\intercal\mathbf{x} = \mathbf{b}^\intercal\mathbf{y}$, pak jsou to optimální hodnoty obou úloh.
- *PODM. KOMPL.*: $\mathbf{c}^\intercal\mathbf{x} = \mathbf{b}^\intercal\mathbf{y}$ iff na každém řádku alespoň jedna rovnost (aktivní podmínka)
- *SILNÁ DUALITA:* Primár má opt. řeš. iff duál má opt. řeš. $\mathbf{y}$ máli Primár má opt. řeš. $\mathbf{x}$ a duál má opt. řeš. $\mathbf{y}$, pak $\mathbf{c}^\intercal\mathbf{x} = \mathbf{b}^\intercal\mathbf{y}$
- 

| primární/duální | má optimum | neomezená | nepřípustná |
|---|---|---|---|
| má optimum | ano | ne | ne |
| neomezená | ne | ne | ano |
| nepřípustná | ne | ano | ano |

- *STÍNOVÉ CENY:* $f : \mathbb{R}^m \to \mathbb{R}$, $f(\mathbf{b}) = \min\{\mathbf{c}^\intercal\mathbf{x} | \mathbf{A}\mathbf{x} \ge \mathbf{b}, \mathbf{x} \ge \mathbf{0}\} = \max\{\mathbf{b}^\intercal\mathbf{y} | \mathbf{A}^\intercal\mathbf{y} \le \mathbf{c}, \mathbf{y} \ge \mathbf{0}\}$ (přičemž primár i duál mají opt. řeš.) Pokud má duální úloha pro dané $\mathbf{b}$ jediné opt. řeš. $\mathbf{y}^*$, pak je fce $f$ v bodě $\mathbf{b}$ diferencovatelná a $\frac{\partial f(\mathbf{b})}{\partial b_i} = \mathbf{y}_i^*$.

## Kapitola 15 (Konvexní funkce)

- Funkce $f : \mathbb{R}^n \to \mathbb{R}$ na množině $X \subseteq \mathbb{R}^n$ se nazývá konvexní, iff: $x, y \in X, 0 \le \alpha \le 1 \to f((1-\alpha)x + \alpha y) \le (1-\alpha)f(x) + \alpha f(y)$
- konkávní je funkce na množině iff $-f$ je konvexní na množině
- Jensenova nerovnost zobecňuje podmínku konvexity fce: $x_i \in X, \alpha_i \ge 0, \sum \alpha_i = 1 \to f(\alpha_1 x_1 + .. + \alpha_k x_k) \le \alpha_1 f(x_1) + .. + \alpha_k f(x_k)$
- Př: $f(x) = \max(x_i) \to f((1-\alpha)x + \alpha y) = \max((1-\alpha)x_i + \alpha y_i) \le \max((1-\alpha)x_i) + \max(\alpha y_i) = (1-\alpha)f(x) + \alpha f(y)$..fce je konvexní
- Příklady konvexních fcí: $f(x) = e^{ax}$ $a \in \mathbb{R}$ na $\mathbb{R}$, $f(x) = x^a$ $a \ge 1 \vee a \le 0$ na $\mathbb{R}_{++}$, $f(x) = |x|^a$ $a \ge 1$ na $\mathbb{R}$, $f(x) = \mathbf{a}^\intercal\mathbf{x} + b$ je zároveň konv. i konk, $f(x) = \mathbf{x}^\intercal\mathbf{A}\mathbf{x}$ konv. pro $\mathbf{A}$ PSD, každá norma je konv., $x\log x$ je konvexní na $\mathbb{R}_{++}$
- Příklady konkávních fcí: $f(x) = \log(x)$ na $\mathbb{R}_{++}$, $\mathbf{x}^\intercal\mathbf{A}\mathbf{x}$ konk. pro $\mathbf{A}$ NSD, $f(x) = x^a$ $0 \le a \le 1$ na $\mathbb{R}_{++}$
- Epigraf fce je množina $\{(x, y) \in \mathbb{R}^{n+1} | f(x) \le y\}$..množina nad grafem fce
- Subkontura výšky $y$ je $\{x \in \mathbb{R}^n | f(x) \le y\}$
- Funkce je konvexní iff její epigraf je konvexní. Obousměrná implikace.
- Každá subkontura konv. fce je kovexní množina. Obrácená implikace neplatí.
- Ať je $f : \mathbb{R}^n \to \mathbb{R}$ diferencovatelná. Pak je $f$ konvexní na $\mathbb{R}^n$, iff pro $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ platí: $f(\mathbf{y}) \ge f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x})$..tzn $T_1$ fce $f$ je v každém bodě $x \in X$ všude menší nebo roven fci $f$
- Ať je $f : \mathbb{R}^n \to \mathbb{R}$ dvakrát diferencovatelná. Pak je $f$ konvexní na $\mathbb{R}^n$, iff v každém $x \in \mathbb{R}^n$ je $f''(x)$ PSD.
- Ať $g_1..g_k : \mathbb{R}^n \to \mathbb{R}$ jsou konvexní, pak pro $\alpha_i \ge 0$ je $f = \alpha_1 g_1 + .. + \alpha_k g_k$ také konvexní. Může se stát, že i takováto kombinace nekonvex. fcí je nakonec konvexní fce ($f(x) = x^3 - x^3$).
- Skládání konvex. fcí nemusí být konvex. fce. Např.
- Ať fce $g : \mathbb{R}^m \to \mathbb{R}$ je konvex., $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, pak fce $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{b})$ je konvexní.
- Ať $I$ je libovolná množina, $g_i : \mathbb{R}^n \to \mathbb{R}$ jsou konv. fce. Pak: $f(x) = \max_{i\in I}(g_i(x))$ je konv fce, předpokládáme že pro každé $x$ maximum existuje. Dané tím, že epigraf $f$ je průnikem epigrafů $g_i$.

## Kapitola 16 (Konvexní optimalizace)

- Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$ (tj. je to konvexní optim. úloha). Pak každé lok. min. fce $f$ na $X$ je zároveň globální.
- Příklady konvexních úloh: lineární programování, kvadratické programování, kvadratické programování s kvadratickými omezením, programování na kuželu druhého řádu, semidefinitní programování.
- Příklad nepřípustné úlohy: $\max\{b \mid b = z + s + 100p, b \le 40, b \ge 20\}$, kde $z$...znalosti, $s$...štěstí, $p$...tento přehled, $b$...body ze zkoušky
- Konvexní relaxace: když mám konv. fci na složité (nekonv.) množině, vezmu její konv. podmnožinu a získám alespoň horní odhad minima.