Lineární algebra pro pokročilé

Martin Černý, Nikola Kalábová

kam.mff.cuni.cz/$\sim$cerny
cerny@kam.mff.cuni.cz

March 6, 2023

Ortogonální báze a Gram-Schmidtova ortogonalita

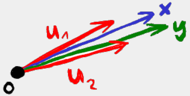
Připomenutí:

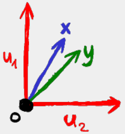
- $\mathbf{v}_1, \ldots, \mathbf{v}_n \neq \mathbf{o}$ $\mathbf{v}_i \perp \mathbf{v}_j$ $\forall i \neq j$ ... ortogonální systém vektorů
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ jsou lineárně nezávislé
- Steinitz: Libovolná množina nenulových lineární nezávislých vektorů lze rozšířit na bázi prostoru.
- **Dnes:** Libovolná množina nenulových ortogonálních vektorů lze rozšířit na ortogonální bázi.

Idea:

- *Lineárně nezávislé vektory ukazují odlišným směrem*
- *Ortogonální vektory ukazují maximálním možným odlišným směrem*

1. **Výhodné numerické vlastnostni**

- 

versus


- vlevo: souřadnice $\mathbf{x}$ a $\mathbf{y}$ vůči $\mathbf{u}_1, \mathbf{u}_2$ jsou velice odlišné, špatné pro zaokrouhlování
- vpravo: změna v souřadnicích je úměrná $\|\mathbf{x} - \mathbf{y}\|$

1. Výhodné numerické vlastnostni ortogonální bázi

Příklad: ***Krylovův podprostor*** $\mathcal{K}_n(A, \mathbf{v}) = \langle \mathbf{v}, A\mathbf{x}, A^2\mathbf{v}, \ldots, A^{n-1}\mathbf{v} \rangle$

- obsahuje hodně informací o matici A
- lze využít např. k řešení $A\mathbf{x} = \mathbf{o}$
- úhly mezi $\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$ moc malé
- $\implies$ kumulace chyb a ztráta informace při zaokrouhlování
- Řeší se konstrukcí ortogonální báze $\mathcal{K}_n(A, \mathbf{v})$
  - *Arnoldiho algoritmus*

Další motivace ortogonálních bází:

2. Snadný výpočet souřadnic pomocí ortogonálních projekcí
    - pro obecné báze řešení soustavy $A\alpha = \mathbf{x}$

Další motivace ortogonálních bází:

2. Snadný výpočet souřadnic pomocí ortogonálních projekcí
   - pro obecné báze řešení soustavy $A\alpha = \mathbf{x}$

Souřadnice ortogonálních bází

Pro ortogonální bázi $\mathbf{v}_1, \ldots, \mathbf{v}_n$ jsou souřadnice vektoru $\mathbf{x} = \sum \alpha_i \mathbf{v}_i$ rovny

$$\alpha_i = \frac{\langle \mathbf{x}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle}.$$

Důkaz:

- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ ... ortogonální báze
- $\mathbf{x} = \sum \alpha_i \mathbf{v}_i$
- $\langle \mathbf{x}, \mathbf{v}_i \rangle = \langle \sum \alpha_k \mathbf{v}_k, \mathbf{v}_i \rangle = \alpha_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle$
  - z ortogonality ($\langle \mathbf{v}_k, \mathbf{v}_i \rangle = 0$) a linearity $\langle \bullet, \bullet \rangle$

2. Snadný výpočet souřadnic pomocí ortogonálních projekcí

Příklad: ***Fourierova řada*** $f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$

- $B = \{\cos(nx), \sin(nx), 1 : n \in \mathbb{N}\}$
  - $\cos(nx), \sin(nx), 1$ jsou *hladké* funkce $f\colon [0, 2\pi] \to \mathbb{R}$
  - jsou **ortogonální**
    - hladké funkce $f\colon [0, 2\pi] \to \mathbb{R}$ tvoří VP
    - standardní skalární součin:
    - $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)dx.$
    - $\int \sim \sum$
- $a_n = \frac{\langle f(x), \cos(nx)\rangle}{\langle \cos(nx), \cos(nx)\rangle}$
- $b_n = \frac{\langle f(x), \sin(nx)\rangle}{\langle \sin(nx), \sin(nx)\rangle}$
- ortogonální projekce na přímku vektorů z B
- $f(x) \in \langle B\rangle \implies f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)$
- Jedna z úloh analýzy: které funkce náleží do $\langle B\rangle$
- lze uvažovat i na obecnějších prostorech

O ortogonálních bázích

Libovolnou množinu nenulových ortogonálních vektorů lze rozšířit na ortogonální bázi.

Důkaz:

- $\mathbf{v}_1, \ldots, \mathbf{v}_k$ ortogonální množina vektorů
- $V_k = \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle$
- chceme $\mathbf{v}_{k+1}$ z doplňku V_k
    - a ortogonální
- $\implies$ **z ortogonálního doplňku!**
- ${V_k}^{\perp}$ ... ortogonální doplněk
- Jak ho snadno zkonstruovat?
- $\mathcal{R}(A) = Ker(A)^{\perp}$

O ortogonálních bázích

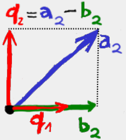
Libovolnou množinu nenulových ortogonálních vektorů lze rozšířit na ortogonální bázi.

Důkaz:

- $\mathbf{v}_1, \ldots, \mathbf{v}_k$ ortogonální množina vektorů
- A tvořena řádky $\mathbf{v}_1{}^T, \ldots, \mathbf{v}_k{}^T$
- $\mathbf{o} \neq \mathbf{v}_{k+1} \in Ker(A)$
- $\implies$ systém o 1 větší
- iterujeme $n-k$-krát

Věta o ortogonálních bázích

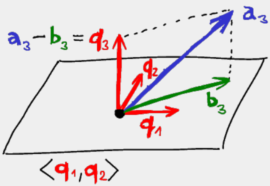
- Jaká je nevýhoda předchozího důkazu?
- Byť konstruktivní, tak výpočetně náročný
  - v každém kroku řešíme soustavu lineárních rovnic
- ukážeme si alternativní konstrukci ortogonální báze
  - kromě praktického výpočtu i jako důkaz věty

- Vstup: $\mathbf{a}_1, \ldots, \mathbf{a}_n$ ... lineárně nezávislá množina
- Výstup: $\mathbf{q}_1, \ldots, \mathbf{q}_n$ ... ortogonální množina (stejného obalu)
- Idea: vektory ortogonalizuje krok za krokem
  1. $\mathbf{q}_1 = \mathbf{a}_1$
  2. $\mathbf{q}_2 = \mathbf{a}_2 - \frac{\langle \mathbf{a}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{q}_1, \mathbf{q}_1 \rangle}\mathbf{q}_1 = \mathbf{a}_2 - \mathbf{b}_2$
  - 

  - odečteme kolmou projekci na přímku $\mathbf{q}_1$

Gram-Schmidtova ortogonalizace

- Vstup: $\mathbf{a}_1, \ldots, \mathbf{a}_n$ ... lineárně nezávislá množina
- Výstup: $\mathbf{q}_1, \ldots, \mathbf{q}_n$ ... ortogonální množina (stejného obalu)
- vektory ortogonalizuje krok za krokem
  1. $\mathbf{q}_1 = \mathbf{a}_1$
  2. $\mathbf{q}_2 = \mathbf{a}_2 - \frac{\langle \mathbf{a}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{q}_1, \mathbf{q}_1 \rangle} \mathbf{q}_1 = \mathbf{a}_2 - \mathbf{b}_2$
  3. $\mathbf{q}_i = \mathbf{a}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i, \mathbf{q}_j \rangle}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle} \mathbf{q}_j = \mathbf{a}_i - \mathbf{b}_i$

  ▶ 

  ▶ odečteme kolmou projekci na přímky $\mathbf{q}_1, \ldots, \mathbf{q}_{i-1}$

Ale funguje to?

- odčítáme postupně kolmou projekci na jednotlivé přímky
- **ALE:** odčítáme kolmou projekci do podprostoru?

Funguje to

- odčítáme postupně kolmou projekci na jednotlivé přímky
- **ALE:** odčítáme kolmou projekci do podprostoru? *ANO*

Správnost G-S

Gram-Schmidtova ortogonalizace je korektní, tj. vrací ortogonální bázi generující stejný podprostor.

Důkaz:

- $\langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle = \langle \mathbf{q}_1, \ldots, \mathbf{q}_n \rangle$
  - z konstrukce (pouze lineární kombinace)
- $\mathbf{q}_1, \ldots, \mathbf{q}_{i-1}$ jsou ortogonální $\implies \langle \mathbf{q}_i, \mathbf{q}_t \rangle$ pro $t < i$
  - $\langle \mathbf{q}_i, \mathbf{q}_t \rangle = \langle \mathbf{a}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i, \mathbf{q}_j \rangle}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle} \mathbf{q}_j, \mathbf{q}_t \rangle = \langle \mathbf{a}_i, \mathbf{q}_t \rangle - \sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i, \mathbf{q}_j \rangle}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle} \langle \mathbf{q}_j, \mathbf{q}_t \rangle =$
  - $= \langle \mathbf{a}_i, \mathbf{q}_t \rangle - \langle \mathbf{a}_i, \mathbf{q}_t \rangle = 0$

- vektory $\mathbf{q}_1, \ldots, \mathbf{q}_n$ je možné průběžné normovat $\mathbf{q}_i := \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|}$
  - vyhneme se dělení $\langle \mathbf{q}_i, \mathbf{q}_i \rangle$ v sumě
- G-S ortogonalizace:
  1. $\mathbf{q}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$
  2. $\mathbf{q}_i = \mathbf{a}_i - \sum_{j=1}^{k-1} \langle \mathbf{a}_i, \mathbf{q}_j \rangle \mathbf{q}_j$
  3. $\mathbf{q}_i = \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|}$

O ORTOGONÁLNÍCH BÁZÍCH 2

- vektory $\mathbf{q}_1, \ldots, \mathbf{q}_n$ je možné průběžné normovat $\mathbf{q}_i := \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|}$
  - vyhneme se dělení $\langle \mathbf{q}_i, \mathbf{q}_i \rangle$ v sumě
- G-S ortogonalizace:
  1. $\mathbf{q}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$
  2. $\mathbf{q}_i = \mathbf{a}_i - \sum_{j=1}^{k-1} \langle \mathbf{a}_i, \mathbf{q}_j \rangle \mathbf{q}_j$
  3. $\mathbf{q}_i = \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|}$

O ortogonálních bázích

Libovolnou množinu nenulových ortogonálních vektorů lze rozšířit na ortogonální bázi.

Důkaz:

- $\mathbf{q}_1, \ldots, \mathbf{q}_k$ (ortogonální množina)

1. Steinitz: $\mathbf{q}_1, \ldots, \mathbf{q}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ (báze prostoru)
2. G-S: $\mathbf{q}_1, \ldots, \mathbf{q}_n$ (ortogonální báze prostoru)

Ortogonální matice

Matice Q je ***ortogonální***, pokud má orto*normální* sloupce.

- ortogonální sloupce + sloupce normy 1

Ortogonální matice

Matice Q je ***ortogonální***, pokud má ortonormální sloupce.

- ortogonální sloupce + sloupce normy 1

Alternativní charakterizace ortogonálních matic

Pro ortogonální matice platí $Q^T Q = I_n$, tedy Q^T je levá inverze Q^{-1}. Speciálně pro čtvercové matice platí $Q^T = Q^{-1}$.

Důkaz:

- $(Q^T Q)_{ij} = \langle \mathbf{q}_i, \mathbf{q}_j \rangle$
  - $i = j \implies \langle \mathbf{q}_i, \mathbf{q}_i \rangle = \|\mathbf{q}_i\|^2 = 1$
  - $i \neq j \implies \langle \mathbf{q}_i, \mathbf{q}_j \rangle = 0$ … $\mathbf{q}_i \perp \mathbf{q}_j$

Zpátky na začátek...

Motivace ortogonálních bází:

- Pro obecnou bázi jsou souřadnice vektoru rovny $A\alpha = \mathbf{x}$
- Pro ortogonální bázi jsou souřadnice vektoru rovny $Q\alpha = \mathbf{x}$
  - $\alpha = Q^T Q\alpha = Q^T\mathbf{x}$
  - odpovídá skalárním součinům:
  - $$\begin{bmatrix} - & q_1 & - \\ - & q_2 & - \\ & \vdots & \\ - & q_n & - \end{bmatrix} \cdot \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} \langle x | q_1 \rangle \\ \langle x | q_2 \rangle \\ \vdots \\ \langle x | q_n \rangle \end{bmatrix}$$

Vlastnosti ortogonálních matic

- ortogonální matice tvoří grupu $\mathbf{O}_n$
  - součin dvou ortogonálních matic $Q_1 Q_2$ je opět ortogonální matice

  1. asociativita $(Q_1 Q_2) Q_3 = Q_1 (Q_2 Q_3)$
  2. neutrální prvek $I_n : Q I_n = I_n Q = Q$
  3. inverzní prvek $Q^T : Q Q^T = Q^T Q = I_n$
- geometricky: *rotace* a *zrcadlení*
- podgrupa $\mathbf{SO}_n$
  - pouze rotace
  - poloviční velikost

Vlastnosti ortogonálních matic

- *rotace* a *zrcadlení*
- $\implies$ zachovávají velikosti vektorů a úhly
- $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$
  - $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{x})^T Q\mathbf{y} = \mathbf{x}^T Q^T Q\mathbf{y} = \mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$
- $\|Q\mathbf{x}\| = \|\mathbf{x}\|$
  - pro indukované normy:
  - $\|Q\mathbf{x}\| = \sqrt{\langle Q\mathbf{x}, O\mathbf{x}\rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle} = \|\mathbf{x}\|$

Maticový zápis Gram-Schmidtovy ortogonalizace

- $\mathbf{a}_1, \ldots, \mathbf{a}_n \overset{G-S}{\Longrightarrow} \mathbf{q}_1, \ldots, \mathbf{q}_n$

$$A = \begin{bmatrix} | & & | \\ a_1, & \ldots, & a_k \\ | & & | \end{bmatrix} \quad Q = \begin{bmatrix} | & & | \\ q_1, & \ldots, & q_k \\ | & & | \end{bmatrix}$$

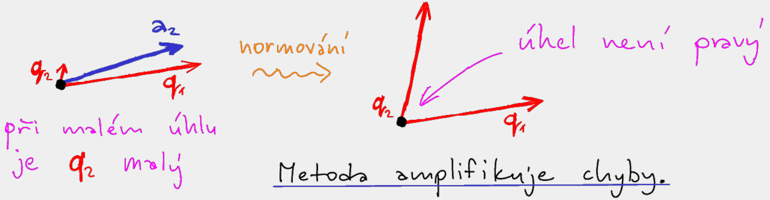
-
- $AR_1R_2 \ldots R_\ell = Q$
- R_t jsou dvou typů:
  1. odčítání $\mathbf{b}_i = \sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i, \mathbf{q}_j \rangle}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle} \mathbf{q}_j$
  2. normování $\mathbf{q}_i = \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|}$

Maticový zápis Gram-Schmidtovy ortogonalizace

- R_t jsou dvou typů:
  1. odčítání $\mathbf{b}_i = \sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i, \mathbf{q}_j \rangle}{\langle \mathbf{q}_j, \mathbf{q}_j \rangle} \mathbf{q}_j$
     - $$\begin{pmatrix} 1 & & & -\langle a_i | q_1 \rangle & & \\ & 1 & & -\langle a_i | q_2 \rangle & & \\ & & \ddots & \vdots & & \\ & & & 1 \; -\langle a_i | q_{i-1} \rangle & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$
  2. normování $\mathbf{q}_i = \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|}$
     - $$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \frac{1}{\|q_i\|} & & \\ & & & & 1 & \\ & & & & & \ddots \; 1 \end{pmatrix}$$

Jedná se o slavný QR rozklad $A = QR$

- Q je ortogonální
- R je horní trojúhelníková
  - $AR_1R_2 \dots R_\ell = Q$
  - $R = (R_1R_2 \dots R_\ell)^{-1}$ je horní trojúhelníková
    - víme z minulého semestru
- špatné numerické vlastnosti:



- ortogonální úpravy: $Q_\ell \dots Q_1A = R$
  - R stejné numerické vlastnosti jako A
  1. Householderovy reflexe
  2. Givensovy rotace