17.3.2015 – 4. cvičení

# Ortogonální báze a Gram-Schmidtova ortogonalizace

Připomeňme, že nenulové ortogonální vektory jsou lineárně nezávislé, tedy pojem ortogonální báze má smysl: $\mathbf{v}_1, \dots, \mathbf{v}_n \neq \mathbf{0} : \mathbf{v}_i \perp \mathbf{v}_j \quad \forall i \neq j.$

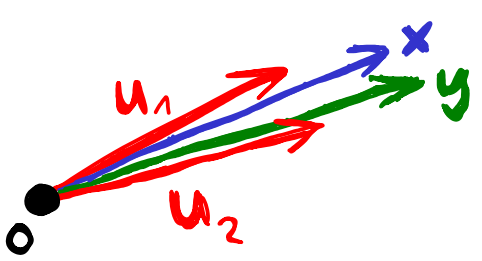
Chceme dokázat, že jich je hodně a mají bohatou strukturu:

**Věta:** Libovolnou množinu nenulových ortogonálních vektorů lze rozšířit na ortogonální bázi.

**Motivace: Proč jsou ortogonální báze důležité?**

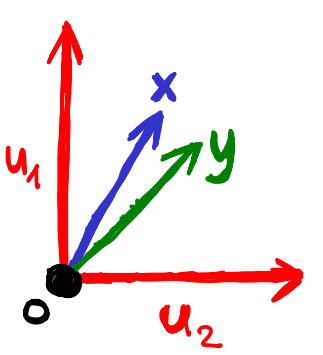
→ Ortogonální vektory ukazují maximálním možným odlišným směrem.

① Výhodné numerické vlastnosti




Souřadnice **x** a **y** vůči $\mathbf{u}_1, \mathbf{u}_2$ jsou velice odlišné, špatné pro zaokrouhlování.

versus




Změna v souřadnicích je úměrná $\|\mathbf{x}-\mathbf{y}\|$.

Krylovův podprostor $\mathcal{K}_n(A,\mathbf{v}) = \langle \mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^{n-1}\mathbf{v} \rangle$ obsahuje hodně informací o matici, lze využít například k řešení $A\mathbf{x}=\mathbf{b}$. Problém je, že úhly mezi vektory jsou hodně malé, což při zaokrouhlování vede ke kumulaci chyb a ztrátě informace.

Řešení: konstrukce ortogonální báze $\mathcal{K}_n(A,\mathbf{v})$ (tzv. Arnoldiho algoritmus)

22 (2.) Snadný výpočet souřadnic pomocí ortogonálních projekcí, zatímco pro obecné báze je to řešení soustavy $A\alpha = x$.

**Lemma:** Pro ortogonální bázi $v_1, \dots, v_n$ je $x = \sum \alpha_i v_i$, kde $\alpha_i = \dfrac{\langle x | v_i \rangle}{\langle v_i | v_i \rangle}$. ← kolmá projekce na jednotlivé přímky

**Důkaz:** $\langle x | v_i \rangle = \langle \sum \alpha_i v_i | v_i \rangle = \alpha_i \langle v_i | v_i \rangle$. (ortogonalita báze) □

Mnohem snazší než pro obecné báze.

Základní myšlenka Fourierovy řady:

Hladké funkce $f : [0; 2\pi] \to \mathbb{R}$ tvoří vektorový prostor se standardním skalárním součinem

$$\langle f | g \rangle = \int_0^{2\pi} f(x)\, g(x)\, dx.$$

← $\int$ je jako $\Sigma$ i, rozmyslete si, že splňuje vlastnosti (i) až (iii).

Množina funkcí $B = \{\cos(nx), \sin(nx), 1 : n \in \mathbb{N}\}$ je ortogonální, což lze ověřit výpočtem integrálů.

Fourierova řada $f(x) = a_0 + \sum_1^\infty a_n \cos(nx) + \sum_1^\infty b_n \sin(nx)$,

kde $a_n = \dfrac{\langle f(x) | \cos(nx) \rangle}{\langle \cos(nx) | \cos(nx) \rangle}$, $b_n = \dfrac{\langle f(x) | \sin(nx) \rangle}{\langle \sin(nx) | \sin(nx) \rangle}$.

Za určitých podmínek konverguje k $f(x)$, pokud $f(x) \in \langle B \rangle$. Je to aplikace lemmatu, souřadnice jsou ortogonální projekce na přímky vektorů $B$.

V analýze se řeší, které funkce leží v $\langle B \rangle$ a které vlastnosti jsou zachovány řadami funkcí. Lze uvažovat i na obecnějších prostorech.

## První důkaz věty – fundamentální podprostory

Nechť $\mathbf{v}_1, \dots, \mathbf{v}_k$ je ortogonální množina, dejme je do řádků matice $A$. Tedy

$$\mathcal{R}(A) = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle, \quad \dim \operatorname{Ker}(A) = n - k.$$

Protože $\operatorname{Ker}(A) = \mathcal{R}(A)^\perp$, můžeme zvolit libovolný vektor $\mathbf{0} \neq \mathbf{v}_{k+1} \in \operatorname{Ker}(A)$, čímž zvětšíme ortogonální množinu o jedna. Opakováním vytvoříme bázi. □

## Gram-Schmidtova ortogonalizace:

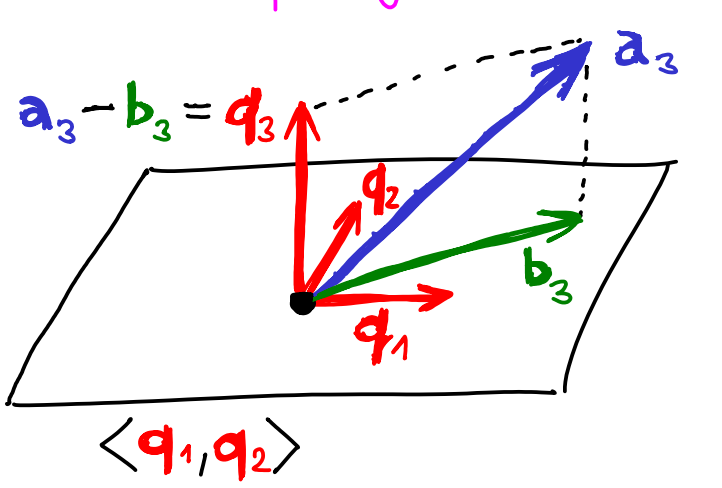
Z lineárně nezávislé množiny vektorů $\mathbf{a}_1, \dots, \mathbf{a}_k$ vyrobí ortogonální množinu $\mathbf{q}_1, \dots, \mathbf{q}_k$ stejného obalu.

Vektory ortogonalizujeme krok za krokem:

$$\mathbf{q}_1 = \mathbf{a}_1,$$

$$\mathbf{q}_i = \mathbf{a}_i - \underbrace{\sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i | \mathbf{q}_j \rangle}{\langle \mathbf{q}_j | \mathbf{q}_j \rangle} \mathbf{q}_j}_{= \mathbf{b}_i}.$$

odečítáme kolmé projekce na $\mathbf{q}_1, \dots, \mathbf{q}_{i-1}$




### Důkaz správnosti:

Je zjevné, že $\langle \mathbf{q}_1, \dots, \mathbf{q}_k \rangle = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$. Nevíme však, že $\mathbf{q}_1, \dots, \mathbf{q}_k$ jsou ortogonální, budeme dokazovat indukcí:

$$\langle \mathbf{q}_i | \mathbf{q}_t \rangle = \left\langle \mathbf{a}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i | \mathbf{q}_j \rangle}{\langle \mathbf{q}_j | \mathbf{q}_j \rangle} \mathbf{q}_j \,\middle|\, \mathbf{q}_t \right\rangle \overset{\text{linearita}}{=} \langle \mathbf{a}_i | \mathbf{q}_t \rangle -$$

$\forall t < i$ ← nenulové pouze pro $j = t$

$$\sum_{j=1}^{i-1} \frac{\langle \mathbf{a}_i | \mathbf{q}_j \rangle}{\langle \mathbf{q}_j | \mathbf{q}_j \rangle} \langle \mathbf{q}_j | \mathbf{q}_t \rangle = \langle \mathbf{a}_i | \mathbf{q}_t \rangle - \langle \mathbf{a}_i | \mathbf{q}_t \rangle = 0.$$

□

Vlastně platí, že $\mathbf{b}_i$ je kolmá projekce $\mathbf{a}_i$ na podprostor $\langle \mathbf{q}_1, \dots, \mathbf{q}_{i-1} \rangle$.

24 Je možné vektory $q_1, \dots, q_k$ průběžně normovat a vyhnout se dělení:

$$q_i := \frac{q_i}{\|q_i\|}, \text{ potom } \langle q_j | q_j \rangle = 1.$$

Druhý důkaz věty – přes Gram-Schmidtovu ortogonalizaci

Množinu $v_1, \dots, v_k$ libovolně doplníme na bázi a aplikujeme ortogonalizaci. □

Ortogonální matice: Matice Q je ortogonální, pokud má ortonormální sloupce.

↖ ortogonalita + norma rovna jedné

Platí, že $Q^TQ = I_n$, tedy $Q^T$ je levá inverze $Q$.
Pro čtvercové matice Q je $Q^T$ jejich inverze.

↳ sloupce jsou ortogonální báze

K ② motivaci: Vyjádření souřadnic vůči obecné bázi vede na $A\alpha = x$, pro ortogonální bázi na $Q\alpha = x$, kde $\alpha = Q^TQ\alpha = Q^Tx$, což odpovídá skalárním součinům u kolmých projekcí.

↑ alternativně $\alpha = A^{-1}x = Q^Tx$, řešení inverzní maticí

$$\begin{bmatrix} — & q_1 & — \\ — & q_2 & — \\ & \vdots & \\ — & q_n & — \end{bmatrix} \cdot \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} \langle x | q_1 \rangle \\ \langle x | q_2 \rangle \\ \vdots \\ \langle x | q_n \rangle \end{bmatrix}$$

Vlastnosti ortogonálních matic

→ Součin dvou ortogonálních matic $Q_1Q_2$ je ortogonální.

Ortogonální matice tvoří grupu $O_n$, geometricky jsou to rotace a zrcadlení. Podgrupa poloviční velikosti $SO_n$ je tvořena jen rotacemi.

↑ speciální ortogonální grupa

→ Zachovávají standardní skalární součin:

$$\langle Qx | Qy \rangle = (Qx)^TQy = x^TQ^TQy = x^Ty = \langle x | y \rangle.$$

Zachovávají úhly, délky: $\|Qx\| = \|x\|$. ← výhodné numericky, neamplifikují chyby

# Maticový zápis Gram-Schmidtovy ortogonalizace:

Matici $A = \left[\begin{array}{c|c|c} & & \\ a_1, \dots, & a_k & \\ & & \end{array}\right]$ transformujeme na $Q = \left[\begin{array}{c|c|c} & & \\ q_1, \dots, & q_k & \\ & & \end{array}\right]$,

která je ortogonální: $AR_1R_2 \cdots R_e = Q.$

$R_t$ jsou dvou typů:

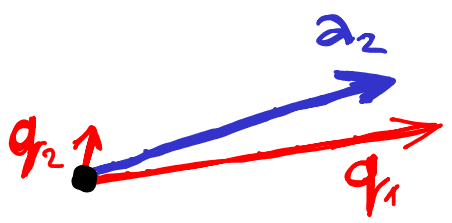
$$\begin{pmatrix} 1 & & & -\langle a_i|q_1\rangle & & \\ & \ddots & & -\langle a_i|q_2\rangle & & \\ & & 1 & \vdots & & \\ & & & -\langle a_i|q_{i-1}\rangle & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \frac{1}{\|q_i\|} & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & 1 \end{pmatrix}$$

odečítá $b_i$ — normuje $q_i$

## Slavný QR rozklad:

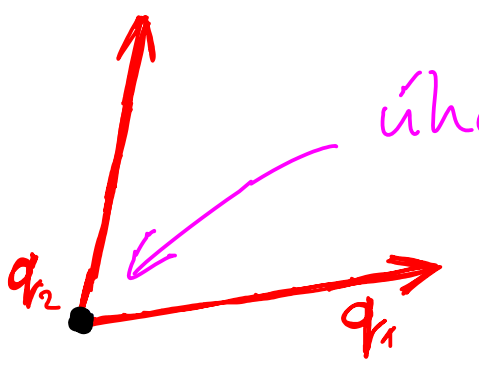
$A = QR$ ← $Q$ ortogonální, $R$ horní trojúhelníková

## Má špatné numerické vlastnosti:




při malém úhlu je $q_2$ malý

normování

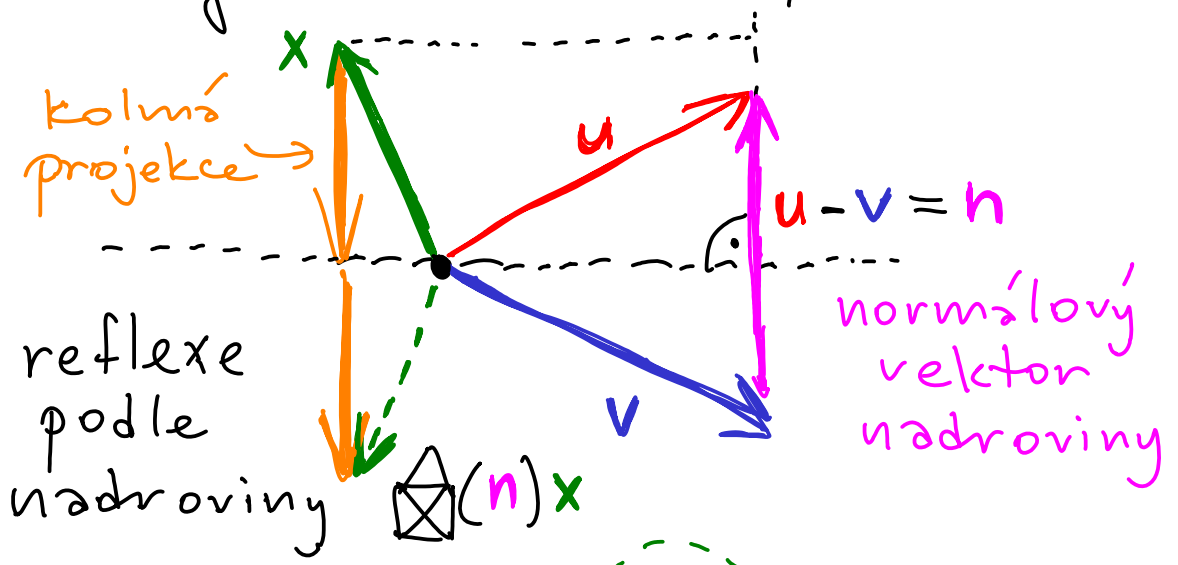



Metoda amplifikuje chyby.

**Řešení:** ortogonální úpravy $Q_e \cdots Q_1 A = R$, potom $R$ má stejné numerické vlastnosti jako $A$.

## 1. Householderovy reflexe $\boxtimes(n)$

Ortogonální matice, co zobrazuje $u \mapsto \|u\| e_1 = v$.




$\boxtimes(n)x$ získáme odečtením dvojnásobku kolmé projekce $x$ na $n$.

Tedy $\boxtimes(n) = I - 2\frac{nn^T}{n^Tn}$. ← matice kolmé projekce na přímku

26

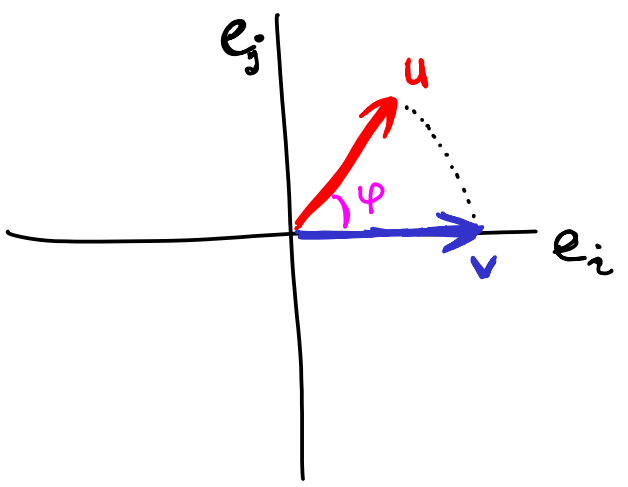
$$\boxtimes(\mathbf{a}_1 - \|\mathbf{a}_1\|\mathbf{e}_1)\cdot A = \begin{bmatrix} \|\mathbf{a}_1\| & \\ 0 & \\ \vdots & ? \\ 0 & \end{bmatrix}, \qquad \begin{pmatrix} 1 & \\ & \boxtimes \end{pmatrix} \boxtimes \cdot A = \begin{bmatrix} * & * & \\ & * & \\ 0 & & ? \end{bmatrix},$$

a tak dál. Získáme $A = QR$ po $n-1$ krocích.

## ② Givensovy rotace:

Vyrobí pouze jednu 0, musíme násobit $O(n^2)$-krát, vhodné pro matice s hodně nulami.

Uvažme rovinu určenou $\mathbf{e}_i$ a $\mathbf{e}_j$, v ní rotujeme:




$$\begin{pmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_j \\ \vdots \\ u_n \end{pmatrix} \mapsto v = \begin{pmatrix} u_1 \\ \vdots \\ \sqrt{u_i^2 + u_j^2} \\ \vdots \\ 0 \\ \vdots \\ u_n \end{pmatrix}$$

← pouze dvě souřadnice změněny

$$R^{\varphi}_{i,j} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\varphi & & -\sin\varphi & & \\ & & & \ddots & & & \\ & & \sin\varphi & & \cos\varphi & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

(sloupce $i$, $j$; řádky $i$, $j$)

$$\varphi = \arccos\left(\frac{u_i}{\sqrt{u_i^2 + u_j^2}}\right).$$

vhodné úhly

$$R^{\varphi_{1,n}}_{1,n} \cdots R^{\varphi_{1,3}}_{1,3} R^{\varphi_{1,2}}_{1,2} A = R^{\varphi_{1,n}}_{1,n} \cdots R^{\varphi_{1,3}}_{1,3} \begin{bmatrix} * & \\ 0 & ? \\ \text{původní část } A & \end{bmatrix} = \cdots = \begin{bmatrix} * & \\ 0 & \\ \vdots & ? \\ 0 & \end{bmatrix}.$$

a dodělám zbytek podobně jako v ①.