10.3.2015 – 3. cvičení

Ortogonální projekce a Cauchy-Schwarzova nerovnost

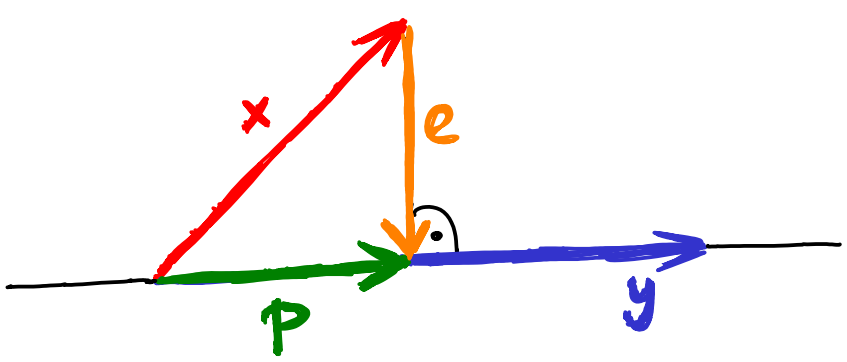
V tomto cvičení objevíme geometrickou interpretaci pro skalární součin.

Ortogonální projekce na přímku:

Mějme dva vektory x a y.

Chceme nalézt kolmou projekci

p vektoru x na přímku $\langle y \rangle$.



Má platit $p = c \cdot y$. Chceme zvolit c tak, aby platilo

$p = c \cdot y$ … nějaká konstanta

$e = p - x \perp y$.

↙ dosadíme za p.

$$0 = \langle p - x \mid y \rangle = \langle cy - x \mid y \rangle = c\langle y|y\rangle - \langle x|y\rangle, \text{ neboli } c = \frac{\langle x|y\rangle}{\langle y|y\rangle}.$$

pouze normovací faktor

Věta: Kolmá projekce na přímku $\langle y \rangle$ je lineární zobrazení

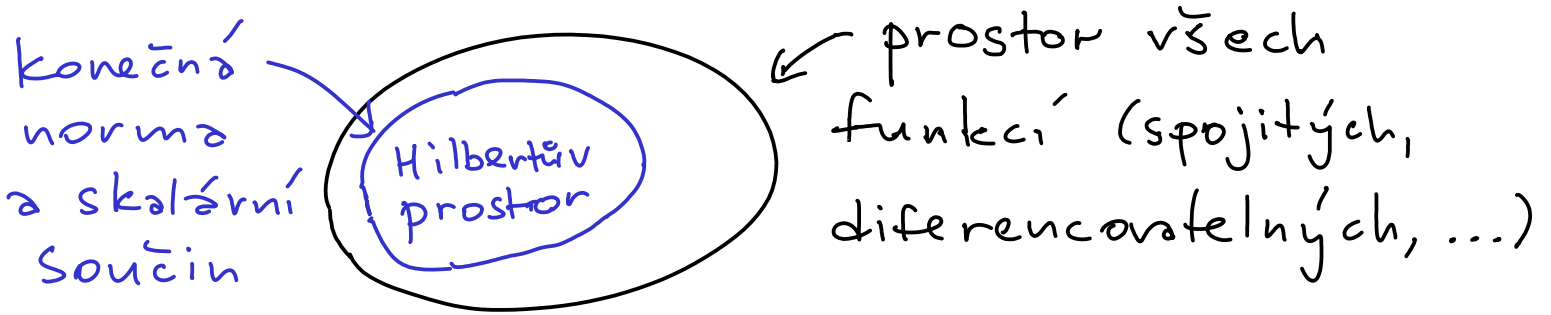
$$x \mapsto \frac{\langle x|y\rangle}{\langle y|y\rangle} y.$$

□

Cauchy-Schwarzova nerovnost:

Pro libovolné dva vektory x a y platí $|\langle x|y\rangle| \leq \|x\| \cdot \|y\|$.

Motivace třeba ve funkcionální analýze, ve vektorových prostorech nekonečné dimenze. Zde mohou mít funkce nekonečnou normu, proto se omezíme na Hilbertovy podprostory pouze funkcí konečné normy. Podle nerovnosti je i konečný skalární součin.



18 Důkaz nerovnosti:

Délka rozdílu projekce e je nezáporná, tedy i

$$0 \leq \|e\|^2 = \langle e \mid e \rangle = \langle p - x \mid p - x \rangle = \left\langle \frac{\langle x|y\rangle}{\langle y|y\rangle} y - x \;\middle|\; \frac{\langle x|y\rangle}{\langle y|y\rangle} y - x \right\rangle =$$

(symetrie, linearita)

stačí elementární úpravy

$$= \left\langle \frac{\langle x|y\rangle}{\langle y|y\rangle} y \;\middle|\; \frac{\langle x|y\rangle}{\langle y|y\rangle} y \right\rangle - 2 \left\langle \frac{\langle x|y\rangle}{\langle y|y\rangle} y \;\middle|\; x \right\rangle + \langle x|x\rangle =$$

(opět linearita)

$$= \frac{\langle x|y\rangle^2}{\langle y|y\rangle^2} \langle y|y\rangle - 2 \frac{\langle x|y\rangle^2}{\langle y|y\rangle} + \langle x|x\rangle,$$

neboli $\langle x|y\rangle^2 \leq \langle x|x\rangle \langle y|y\rangle$. ↗ Cauchy–Schwarzova nerovnost plyne odmocněním. □

Dokažme nyní, že skalární součin indukuje normu $\sqrt{\langle x|x\rangle}$, konkrétně že platí trojúhelníková nerovnost:

$$\|x + y\| \leq \|x\| + \|y\|$$

↓ vše kladné, umocníme druhou mocninou

$$\langle x+y \mid x+y \rangle \leq \langle x|x\rangle + 2\|x\|\cdot\|y\| + \langle y|y\rangle$$

$$\langle x|x\rangle + 2\langle x|y\rangle + \langle y|y\rangle \leq \langle x|x\rangle + 2\|x\|\cdot\|y\| + \langle y|y\rangle$$

$$\langle x|y\rangle \leq \|x\|\cdot\|y\|$$ → Cauchy-Schwarzova nerovnost

Tedy trojúhelníková nerovnost je ekvivalentní s Cauchy-Schwarzovou nerovností.

Ortogonální projekce na obecný podprostor:

Popišme nejprve matici P kolmé projekce na přímku:

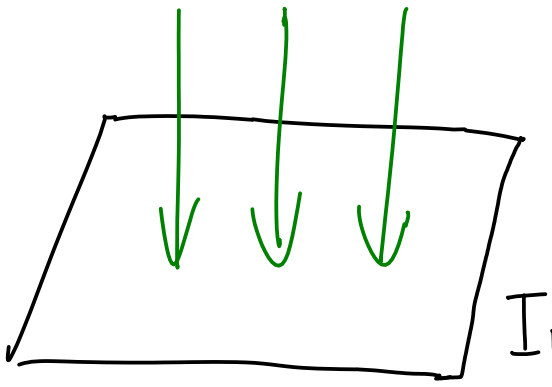
$x \mapsto \frac{\langle x|y\rangle}{\langle y|y\rangle} y$, ve standardním skalárním součinu:

(prohodíme pořadí členů)

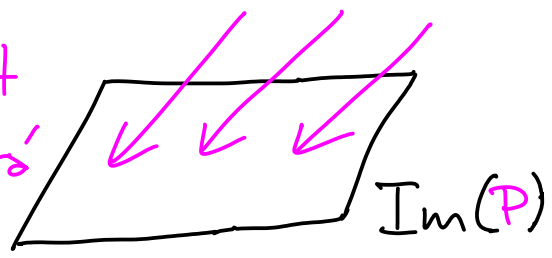
$$x \mapsto \frac{x^T y}{y^T y} y = y \frac{y^T x}{y^T y}, \quad \text{neboli} \quad P = \frac{y\,y^T}{y^T y}.$$

Odvodit přesný předpis pro obecný podprostor je složitější, zatím se spokojíme pouze s vlastnostmi P.

Vlastnosti, co musí P splňovat:



projekce na podprostor



nemusí být obecně kolmá

① $\mathrm{Im}(P)$ je fixován:

$\forall \mathbf{y} \in \mathrm{Im}(P)$: $P\mathbf{y} = \mathbf{y}$, tedy

$\forall \mathbf{x} \in \mathbb{R}^n$: $P(P\mathbf{x}) = P\mathbf{x}$.

Protože to platí pro každý vektor $\mathbf{x}$, je $P^2 = P$.

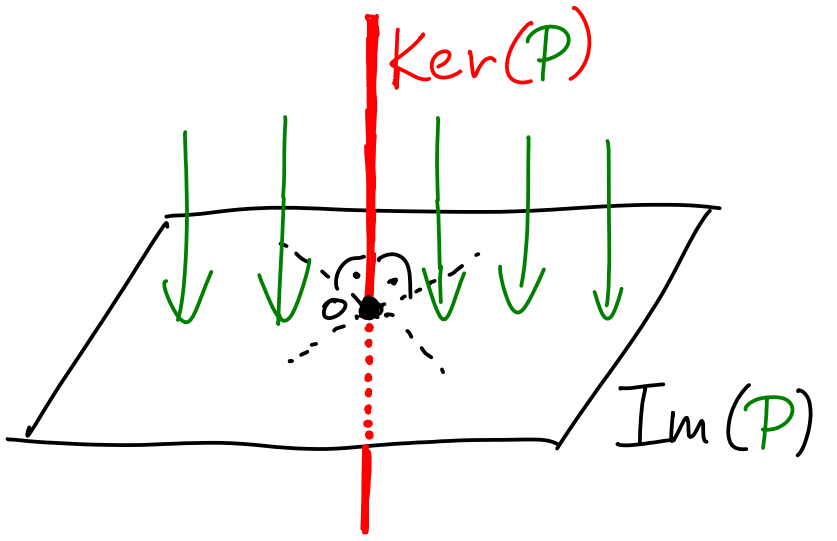
Věta: Matice P je projekce na podprostor $\mathrm{Im}(P)$, právě když $P^2 = P$. □

② Projekce je kolmá: $\forall \mathbf{x} \in \mathbb{R}^n \quad \mathbf{x} - P\mathbf{x} \perp \mathrm{Im}(P)$.

Speciálně $\mathrm{Ker}(P) \perp \mathrm{Im}(P)$, kvůli dimenzím je $\mathrm{Ker}(P) = \mathrm{Im}(P)^\perp$.

Pro každou matici platí, že $\mathrm{Ker}(P) = \mathcal{R}(P)^\perp$, tedy

$$\mathrm{Im}(P) = \mathcal{R}(P).$$



Chceme dokázat, že $P = P^T$.

Definice transpozice přes duální zobrazení:

Platí, že pro dané lineární zobrazení $f: \mathbb{R}^n \to \mathbb{R}^m$ existuje právě jedno lineární zobrazení $f^*: \mathbb{R}^m \to \mathbb{R}^n$, že platí:

duální zobrazení

$$\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m \quad \underbrace{\langle f(\mathbf{x}) \mid \mathbf{y} \rangle}_{\mathbb{R}^m} = \underbrace{\langle \mathbf{x} \mid f^*(\mathbf{y}) \rangle}_{\mathbb{R}^n}.$$

Dokažme alespoň, že v případě standardního skalárního součinu a maticové reprezentace A, kde f^* je reprezentováno A^T:

20 **Důkaz:** Nechť A je matice, chceme zkonstruovat B tak, aby platilo $\langle A\mathbf{x}|\mathbf{y}\rangle = \langle \mathbf{x}|B\mathbf{y}\rangle$. Zjevně $B=A^T$ toto splňuje. Dokažme, že žádná jiná matice to nesplňuje, stačí pro kanonickou bázi $\mathbf{e}_1,\dots,\mathbf{e}_n$:

$$a_{ji} = (A\mathbf{e}_i)_j = \sum_{k=1}^{n} (A\mathbf{e}_i)_k \cdot (\mathbf{e}_j)_k = \sum_{k=1}^{n} (\mathbf{e}_i)_k \cdot (B\mathbf{e}_j)_k = b_{ij}.$$

nenulové pouze pro $k=j$ — nenulové pouze pro $k=i$

Tedy $B=A^T$. □

Vraťme se k ortogonální projekci P. Dokažme, že $\langle P\mathbf{x}|\mathbf{y}\rangle = \langle \mathbf{x}|P\mathbf{y}\rangle$. Protože $\mathrm{Ker}(P) = \mathrm{Im}(P)^\perp$, můžeme vyjádřit:

$$\mathbf{x} = \mathbf{x}_K + \mathbf{x}_I, \quad \mathbf{y} = \mathbf{y}_K + \mathbf{y}_I, \quad \text{kde} \quad \mathbf{x}_K, \mathbf{y}_K \in \mathrm{Ker}(P), \quad \mathbf{x}_I, \mathbf{y}_I \in \mathrm{Im}(P).$$

Protože $P\mathbf{x} = \mathbf{x}_I$ a $P\mathbf{y} = \mathbf{y}_I$, dostáváme:

$$\langle P\mathbf{x}|\mathbf{y}\rangle = \langle \mathbf{x}_I | \mathbf{y}_K + \mathbf{y}_I\rangle = \underbrace{\langle \mathbf{x}_I|\mathbf{y}_K\rangle}_{=0 \text{ z ortogonality}} + \langle \mathbf{x}_I|\mathbf{y}_I\rangle = \langle \mathbf{x}_I|\mathbf{y}_I\rangle.$$

$$\langle \mathbf{x}|P\mathbf{y}\rangle = \langle \mathbf{x}_K + \mathbf{x}_I | \mathbf{y}_I\rangle = \langle \mathbf{x}_K|\mathbf{y}_I\rangle + \langle \mathbf{x}_I|\mathbf{y}_I\rangle = \langle \mathbf{x}_I|\mathbf{y}_I\rangle.$$

Tedy $P^T = P$. Naopak z této vlastnosti vyplývá, že $\mathrm{Ker}(P) = \mathrm{Im}(P)^\perp$, tedy projekce je ortogonální.

Věta: Projekce P je ortogonální, právě když $P=P^T$.

Brzo ukážeme, že když sloupce A tvoří bázi $\mathrm{Im}(P)$, potom $P = A(A^TA)^{-1}A^T$, speciální případ je

projekce na přímku: $P = \dfrac{\mathbf{y}\mathbf{y}^T}{\mathbf{y}^T\mathbf{y}}$ ($\mathbf{y}\mathbf{y}^T \leftarrow AA^T$, $\mathbf{y}^T\mathbf{y} \leftarrow (A^TA)^{-1}$).

Jiný způsob:

P zobrazuje $\mathbf{x}$ na nejbližší bod $\mathrm{Im}(P)$. Lze využít analýzu a derivace. Zkuste pro projekci na přímku.