11. Skalární součin

Skalární součin je jeden ze základních matematických pojmů. Je to pokročilejší způsob, jak měřit vzdálenosti a úhly ve vektorových prostorech. Budeme se jej ilustrovat na nejjednodušším případu, měření úhlů a vzdáleností v rovině. V aritmetickém reálném vektorovém prostoru $\mathcal{R}^2$ dimenze 2 si zvolíme dva nenulové vektory $\mathbf{x} = (x_1, x_2)$ a $\mathbf{y} = (y_1, y_2)$. Pro jednoduchost je zvolíme oba z prvního kvadrantu. Buď α úhel, který svírá spojnice bodu (x_1, x_2) s počátkem souřadnic s kladným směrem vodorovné osy, a β úhel, který svírá spojnice bodu (y_1, y_2) s počátkem souřadnic s kladným směrem vodorovné osy. Dále předpokládejme, že $0 < \beta < \alpha < 90°$ (nakreslete si obrázek). Pomocí trigonometrických vzorců dostáváme

$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \frac{x_1y_1 + x_2y_2}{||\mathbf{x}|| \cdot ||\mathbf{y}||},$$

kde $||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2}$ označuje délku vektoru $||\mathbf{x}||$.

Výraz $x_1y_1 + x_2y_2$ nazýváme skalární součin vektorů $\mathbf{x} = (x_1, x_2)$ a $\mathbf{y} = (y_1, y_2)$ a označujeme ho $(\mathbf{x}, \mathbf{y})$. Ze skalárního součinu $(\mathbf{x}, \mathbf{y})$ pak můžeme vypočítat kosinus úhlu η mezi vektory $\mathbf{x}$ a $\mathbf{y}$ pomocí vztahu

$$\cos\eta = \frac{(\mathbf{x}, \mathbf{y})}{||\mathbf{x}|| \cdot ||\mathbf{y}||},$$

známe-li délky vektorů $\mathbf{x}$ a $\mathbf{y}$. Ty ale také spočítáme pomocí skalárního součinu, neboť $||\mathbf{x}|| = \sqrt{(\mathbf{x}, \mathbf{x})}$ a $||\mathbf{y}|| = \sqrt{(\mathbf{y}, \mathbf{y})}$.

Definice 11.28. *Skalární součin* na reálném vektorovém prostoru $\mathcal{V}$ (tj. vektorovém prostoru nad reálnými čísly $\mathbf{R}$) je funkce, která každé uspořádané dvojici vektorů $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ přiřazuje číslo $(\mathbf{x}, \mathbf{y}) \in \mathbf{R}$ takové, že platí podmínky

a) $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$ pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{V}$,

b) $(a_1\mathbf{x}_1 + a_2\mathbf{x}_2, \mathbf{y}) = a_1(\mathbf{x}_1, \mathbf{y}) + a_2(\mathbf{x}_2, \mathbf{y})$ pro každé vektory $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in \mathcal{V}$ a skaláry $a_1, a_2 \in \mathbf{R}$,

c) $(\mathbf{x}, \mathbf{x}) \geq 0$ pro každé $\mathbf{x} \in \mathcal{V}$, přičemž $(\mathbf{x}, \mathbf{x}) = 0$ právě když $\mathbf{x} = \mathbf{0}$.

V prostoru $\mathcal{V}$ se skalárním součinem nazýváme číslo $\sqrt{(\mathbf{x}, \mathbf{x})} = ||\mathbf{x}||$ *délka* (*norma*) vektoru $\mathbf{x}$. Reálný prostor $\mathcal{V}$ se skalárním součinem se také nazývá *euklidovský prostor.*

Je-li $\mathcal{V}$ vektorový prostor nad komplexními čísly, pak *skalárním součinem* na prostoru $\mathcal{V}$ nazýváme funkci, která přiřazuje každé uspořádané dvojici vektorů $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ komplexní číslo $(\mathbf{x}, \mathbf{y})$ takové, že platí

a) $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$ pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{V}$,

b) $(a_1\mathbf{x}_1 + a_2\mathbf{x}_2, \mathbf{y}) = a_1(\mathbf{x}_1, \mathbf{y}) + a_2(\mathbf{x}_2, \mathbf{y})$ pro každé vektory $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in \mathcal{V}$ a skaláry $a_1, a_2 \in \mathbf{C}$,

c) $(\mathbf{x}, \mathbf{x}) \geq 0$ pro každé $\mathbf{x} \in \mathcal{V}$, přičemž $(\mathbf{x}, \mathbf{x}) = 0$ právě když $\mathbf{x} = \mathbf{0}$.

Zde $\overline{c}$ označuje číslo komplexně sdružené k c. Vzhledem k podmínce a) je první část podmínky c) splněna. Komplexní vektorový prostor $\mathcal{V}$ se skalárním součinem se nazývá *unitární* prostor.

V euklidovském prostoru $\mathcal{V}$ se skalárním součinem $(\cdot\,,\cdot)$ bychom rádi definovali úhel φ mezi libovolnými dvěma nenulovými vektory $\mathbf{x}$ a $\mathbf{y}$. Aby byla naše definice úhlu v souladu s geometrickými úvahami, kterými jsme motivovali skalární součin, mělo by platit

$$\cos\varphi = \frac{(\mathbf{x}, \mathbf{y})}{||\mathbf{x}|| \cdot ||\mathbf{y}||}.$$

K tomu ale potřebujeme vědět, že platí

$$-1 \leq \frac{(\mathbf{x}, \mathbf{y})}{||\mathbf{x}|| \cdot ||\mathbf{y}||} \leq 1$$

pro libovolné dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{V}$. To je obsahem následující věty.

Věta 11.51. Cauchy-Schwarzova-Bunjakovského nerovnost. Je-li $\mathcal{V}$ euklidovský prostor, pak pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ platí

$$|(\mathbf{x}, \mathbf{y})| \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||.$$

Důkaz. Buď t libovolné reálné číslo. Potom platí

$$0 \leq (\mathbf{x} - t\mathbf{y}, \mathbf{x} - t\mathbf{y}) = (\mathbf{x}, \mathbf{x}) - (\mathbf{x}, t\mathbf{y}) - (t\mathbf{y}, \mathbf{x}) + (t\mathbf{y}, t\mathbf{y}) = t^2(\mathbf{y}, \mathbf{y}) - 2t(\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{x}).$$

Poslední výraz je polynom druhého stupně v reálné proměnné t. Má-li platit $0 \leq t^2(\mathbf{y}, \mathbf{y}) - 2t(\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{x})$ pro každé reálné číslo t, nemůže být diskriminant tohoto polynomu kladný. Platí proto

$$4(\mathbf{x}, \mathbf{y})^2 - 4(\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y}) \leq 0,$$

neboli $|(\mathbf{x}, \mathbf{y})| \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||$. $\square$

Důsledek 11.52. Pro libovolné dva vektory $\mathbf{x}, \mathbf{y}$ v euklidovském prostoru $\mathcal{V}$ platí $||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$. Speciálně platí, že jsou-li vektory $\mathbf{x}, \mathbf{y}$ ortogonální, tj. platí $(\mathbf{x}, \mathbf{y}) = 0$, pak $||\mathbf{x} + \mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2$ (Pythagorova věta).

Důkaz. Platí

$$||\mathbf{x} + \mathbf{y}||^2 = (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = (\mathbf{x}, \mathbf{x}) + 2(\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{y}) = ||\mathbf{x}||^2 + 2(\mathbf{x}, \mathbf{y}) + ||\mathbf{y}||^2.$$

Odtud ihned plyne Pythagorova věta. Použitím Věty 11.51 dostaneme dále

$$||\mathbf{x} + \mathbf{y}||^2 = ||\mathbf{x}||^2 + 2(\mathbf{x}, \mathbf{y}) + ||\mathbf{y}||^2 \leq ||\mathbf{x}||^2 + 2||\mathbf{x}|| \cdot ||\mathbf{y}|| + ||\mathbf{y}||^2 = (||\mathbf{x}|| + ||\mathbf{y}||)^2.$$

□

Definice 11.29. Nenulové vektory $\mathbf{e}_1, \ldots, \mathbf{e}_n$ tvoří *ortogonální bázi* v n-dimenzionálním euklidovském vektoru $\mathcal{V}$, platí-li

$$(\mathbf{e}_i, \mathbf{e}_j) = 0, \text{ kdykoliv } i \neq j.$$

Je-li navíc $||\mathbf{e}_i|| = 1$ pro všechna $i = 1, \ldots, n$, nazýváme $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *ortonormální báze* ve $\mathcal{V}$.

Věta 11.53. Vektory v libovolné ortogonální bázi $\mathbf{e}_1, \ldots, \mathbf{e}_n$ jsou lineárně nezávislé.

Důkaz. Kdyby byly vektory $\mathbf{e}_1, \ldots, \mathbf{e}_n$ lineárně závislé, existovalo by $i \geq 1$ takové, že vektor $\mathbf{e}_{i+1}$ by byl lineárně závislý na vektorech $\mathbf{e}_1, \ldots, \mathbf{e}_i$. Platilo by tedy $\mathbf{e}_{i+1} = a_1\mathbf{e}_1 + \cdots a_i\mathbf{e}_i$. Pak ale

$$(\mathbf{e}_{i+1}, \mathbf{e}_{i+1}) = (\mathbf{e}_{i+1}, a_1\mathbf{e}_1 + \cdots + a_i\mathbf{e}_i) = a_1(\mathbf{e}_{i+1}, \mathbf{e}_1) + \cdots + a_i(\mathbf{e}_{i+1}, \mathbf{e}_i) = 0.$$

Z vlastnosti c) skalárního součinu odtud plyne $\mathbf{e}_{i+1} = \mathbf{0}$, což je ve sporu s tím, že prvky ortogonální báze jsou nenulové. □

Dále si ukážeme, že v každém euklidovském prostoru existuje ortonormální báze. K důkazu použijeme metodu, které se říká **Gramm-Schmidtova ortogonalizace.** Uděláme to takhle: buď $\mathbf{x}_1, \ldots, \mathbf{x}_n$ lineárně nezávislé vektory v euklidovském prostoru $\mathcal{V}$. Sestrojíme vektory $\mathbf{e}_1, \ldots, \mathbf{e}_n$ takové, že generují prostor $\mathcal{V}$, a jsou navíc navzájem ortogonální.

Položíme $\mathbf{e}_1 = \mathbf{x}_1$. Vektor $\mathbf{e}_2$ budeme hledat ve tvaru $\mathbf{e}_2 = \mathbf{x}_2 + a\mathbf{e}_1$. Chceme, aby platilo $(\mathbf{e}_2, \mathbf{e}_1) = 0$, tj. $0 = (\mathbf{x}_2 + a\mathbf{e}_1, \mathbf{e}_1) = (\mathbf{x}_2, \mathbf{e}_1) + a(\mathbf{e}_1, \mathbf{e}_1)$. Odtud vypočteme

$$a = -\frac{(\mathbf{x}_2, \mathbf{e}_1)}{(\mathbf{e}_1, \mathbf{e}_1)}$$

a tedy

$$\mathbf{e}_2 = \mathbf{x}_2 - \frac{(\mathbf{x}_2, \mathbf{e}_1)}{(\mathbf{e}_1, \mathbf{e}_1)}\,\mathbf{e}_1.$$

Předpokládejme, že pro nějaké $k < n$ už máme sestrojené vektory $\mathbf{e}_1, \ldots, \mathbf{e}_k$, které jsou po dvou ortogonální a generují stejný podprostor $\mathcal{V}$ jako vektory $\mathbf{x}_1, \ldots, \mathbf{x}_k$. Vektor $\mathbf{e}_{k+1}$ budeme hledat ve tvaru

$$\mathbf{e}_{k+1} = \mathbf{x}_{k+1} + a_1\mathbf{e}_1 + \cdots + a_k\mathbf{e}_k.$$

Skaláry a_i najdeme tak, že poslední rovnost postupně skalárně násobíme vektory $\mathbf{e}_i$, $i = 1, \ldots, k$. Po úpravě dostáváme

$$(\mathbf{e}_{k+1}, \mathbf{e}_i) = (\mathbf{x}_{k+1}, \mathbf{e}_i) + a_i(\mathbf{e}_i, \mathbf{e}_i),$$

čili

$$a_i = -\frac{(\mathbf{x}_{k+1}, \mathbf{e}_i)}{(\mathbf{e}_i, \mathbf{e}_i)}$$

pro $i = 1, \ldots, k$.

Vzhledem k tomu, že vektory $\mathbf{x}_1, \ldots, \mathbf{x}_k$ a $\mathbf{e}_1, \ldots, \mathbf{e}_k$ generují stejný podprostor, a vektor $\mathbf{x}_{k+1}$ je lineárně nezávislý na vektorech $\mathbf{x}_1, \ldots, \mathbf{x}_k$, je vektor $\mathbf{x}_{k+1}$ také lineárně nezávislý na vektorech $\mathbf{e}_1, \ldots, \mathbf{e}_k$. Proto je $\mathbf{e}_{k+1} \neq 0$. Z první rovnosti předchozího odstavce pak plyne, že vektor $\mathbf{x}_{k+1}$ je lineárně závislý na vektorech $\mathbf{e}_1, \ldots, \mathbf{e}_{k+1}$ a proto vektory $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ generují stejný podprostor $\mathcal{V}$ jako vektory $\mathbf{e}_1, \ldots, \mathbf{e}_{k+1}$.

Dále platí

$$(\mathbf{e}_{k+1}, \mathbf{e}_i) = (\mathbf{x}_{k+1}, \mathbf{e}_i) + a_1(\mathbf{e}_1, \mathbf{e}_i) + \cdots + a_k(\mathbf{e}_k, \mathbf{e}_i) = (\mathbf{x}_{k+1}, \mathbf{e}_i) + a_i(\mathbf{e}_i, \mathbf{e}_i) = 0$$

pro $i = 1, \ldots, k$. Tím je dokázáno indukcí podle dimenze prostoru $\mathcal{V}$ první tvrzení následující věty.

K důkazu druhé části stačí nahradit každý vektor $\mathbf{e}_i$ vektorem

$$\mathbf{e}'_i = \frac{\mathbf{e}_i}{(\mathbf{e}_i, \mathbf{e}_i)}.$$

Věta 11.54. V každém euklidovském prostoru $\mathcal{V}$ existuje ortogonální báze. V každém euklidovském prostoru $\mathcal{V}$ existuje ortonormální báze. □

Věta 11.55. Buď $\mathbf{e}_1, \ldots, \mathbf{e}_n$ ortonormální báze v euklidovském prostoru $\mathcal{V}$, $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ libovolné dva vektory a $\mathbf{x} = \sum_1^n a_i\mathbf{e}_i$, $\mathbf{y} = \sum_1^n b_i\mathbf{e}_i$ jejich vyjádření pomocí báze $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Potom

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n a_i b_i.$$

Důkaz. Stačí využít rovností $(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$ pro $i, j = 1, \ldots, n$. Pak dostaneme

$$(\mathbf{x}, \mathbf{y}) = (\sum_1^n a_i\mathbf{e}_i, \sum_1^n b_i\mathbf{e}_i) = \sum_{i,j=1}^n a_i b_j(\mathbf{e}_i, \mathbf{e}_j) = \sum_{i,j=1}^n a_i b_j \delta_{ij} = \sum_{i=1}^n a_i b_i.$$

□

Věta 11.56. Buď $\mathbf{e}_1, \ldots, \mathbf{e}_n$ ortonormální báze v euklidově prostoru $\mathcal{V}$ a $\mathbf{x} \in \mathcal{V}$ libovolný vektor. Je-li

$$\mathbf{x} = \sum_{i=1}^n a_i\mathbf{e}_i$$

vyjádření vektoru $\mathbf{x}$ pomocí báze $\mathbf{e}_1, \ldots, \mathbf{e}_n$, pak $a_i = (\mathbf{x}, \mathbf{e}_i)$ pro $i = 1, \ldots, n$.

Důkaz. Buď $\mathbf{e}_1, \ldots, \mathbf{e}_n$ ortonormální báze v prostoru $\mathcal{V}$ a

$$\mathbf{x} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n$$

vyjádření vektoru $\mathbf{x}$ pomocí této báze. Vynásobíme-li toto vyjádření skalárně vektorem $\mathbf{e}_i$, dostaneme $(\mathbf{x}, \mathbf{e}_i) = a_i$ pro $i = 1, \ldots, n$. $\square$

Dále si ukážeme, že všechny euklidovské prostory téže dimenze n jsou izomorfní.

Definice 11.30. Buď $\mathcal{U}$ a $\mathcal{V}$ dva euklidovské prostory. Izomorfizmus $F : \mathcal{U} \to \mathcal{V}$ lineárních prostorů $\mathcal{U}$, $\mathcal{V}$ se nazývá *izomorfizmus euklidových prostorů* $\mathcal{U}$ a $\mathcal{V}$, platí-li

$$(\mathbf{x}, \mathbf{y})_{\mathcal{U}} = (F(\mathbf{x}), F(\mathbf{y}))_{\mathcal{V}}$$

pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{U}$.

Věta 11.57. Každé dva euklidovské prostory téže dimenze n jsou izomorfní.

Důkaz. Dokážeme, že každý n-dimenzionální euklidův prostor je izomorfní jistému *standardnímu* euklidovu prostoru. Standardní euklidův prostor je aritmetický n-dimenzionální vektorový prostor $\mathcal{R}^n$ vybavený obvyklým skalárním součinem: jsou-li $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ a $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ dva vektory prostoru $\mathcal{R}^n$, pak jejich skalární součin je definován

$$(\mathbf{u}, \mathbf{v})_{\mathcal{R}^n} = \sum_{i=1}^{n} u_i v_i.$$

Buď nyní $\mathcal{V}$ libovolný euklidův prostor dimenze n. Podle Věty 11.54 v něm existuje ortonormální báze $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Vezměme nyní vektory $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ a jejich vyjádření $\mathbf{x} = \sum_1^n a_i\mathbf{e}_i$ a $\mathbf{y} = \sum_1^n b_i\mathbf{e}_i$ pomocí této báze. Podle Věty 11.55 platí

$$(\mathbf{x}, \mathbf{y})_{\mathcal{V}} = \sum_{i=1}^{n} a_i b_i.$$

Zobrazení $F : \mathcal{V} \to \mathcal{R}^n$ definované předpisem

$$F(\mathbf{x}) = (a_1, a_2, \ldots, a_n),$$

kde $\mathbf{x} = \sum_1^n a_i\mathbf{e}_i$ je vyjádření $\mathbf{x}$ pomocí báze $\mathbf{e}_1, \ldots, \mathbf{e}_n$, je izomorfizmus lineárních prostorů $\mathcal{V}$ a $\mathcal{R}^n$ podle Věty 6.24. Potom $F(\mathbf{x}) = (a_1, a_2, \ldots, a_n)$ a $F(\mathbf{y}) = (b_1, b_2, \ldots, b_n)$ a tedy

$$(F(\mathbf{x}), F(\mathbf{y}))_{\mathcal{R}^n} = \sum_{i=1}^{n} a_i b_i = (\mathbf{x}, \mathbf{y})_{\mathcal{V}}.$$

Tím je dokázáno, že každý n-dimenzionální euklidův prostor $\mathcal{V}$ je izomorfní standardnímu n-dimenzionálnímu euklidovu prostoru $\mathcal{R}^n$. Odtud plyne, že každé dva n-dimenzionální euklidovy prostory jsou izomorfní. $\square$

Definice 11.31. Buď $\mathcal{V}$ euklidův prostor a $\mathcal{P}$ jeho podprostor. Vektor $\mathbf{x} \in \mathcal{V}$ je *ortogonální* k $\mathcal{P}$, platí-li $(\mathbf{x}, \mathbf{z}) = 0$ pro každý prvek $\mathbf{z} \in \mathcal{P}$. Množina všech vektorů prostoru $\mathcal{V}$ ortogonálních k $\mathcal{P}$ se nazývá *ortogonální doplněk* podprostoru $\mathcal{P}$ v prostoru $\mathcal{V}$ a označuje se $\mathcal{P}^{\perp}$.

Lemma 11.58. Pro každý podprostor $\mathcal{P}$ euklidova prostoru $\mathcal{V}$ platí
a) $\mathcal{P}^{\perp}$ je podprostor $\mathcal{V}$,
b) $(\mathcal{P}^{\perp})^{\perp}$.

Důkaz. a) Jsou-li $\mathbf{x}, \mathbf{y}$ dva libovolné prvky $\mathcal{P}^{\perp}$, a, b libovolná dvě reálná čísla a $\mathbf{z} \in \mathcal{P}$, pak

$$(a\mathbf{x} + b\mathbf{y}, \mathbf{z}) = a(\mathbf{x}, \mathbf{z}) + b(\mathbf{y}, \mathbf{z}) = 0,$$

tedy také $a\mathbf{x} + b\mathbf{y} \in \mathcal{P}^{\perp}$.

b) Jsou-li $\mathbf{x} \in \mathcal{P}^{\perp}$ a $\mathbf{z} \in \mathcal{P}$ libovolné dva prvky, pak $(\mathbf{z}, \mathbf{x}) = (\mathbf{x}, \mathbf{z}) = 0$, tedy $\mathbf{z} \in (\mathcal{P}^{\perp})^{\perp}$. Proto $\mathcal{P} \subseteq (\mathcal{P}^{\perp})^{\perp}$.

Abychom dokázali opačnou inkluzi, zvolme nějakou bázi $\mathbf{x}_1, \ldots, \mathbf{x}_m$ v podprostoru $\mathcal{P}$ a doplňme ji do báze $\mathbf{x}_1, \ldots, \mathbf{x}_n$ celého prostoru $\mathcal{V}$. Pomocí Gram-Schmidtovy ortogonalizace dostaneme ortonormální bázi $\mathbf{e}_1, \ldots, \mathbf{e}_n$ prostoru $\mathcal{V}$, ve které prvních m vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_m$ tvoří ortonormální bázi $\mathcal{P}$. Potom $\mathbf{e}_i \in \mathcal{P}^{\perp}$ pro každé $i = m+1, \ldots, n$. Vyjádříme libovolný vektor $\mathbf{z} \in \mathcal{V}$ pomocí báze $\mathbf{e}_1, \ldots, \mathbf{e}_n$: $\mathbf{z} = \sum_1^n a_i \mathbf{e}_i$. Jestliže nyní $\mathbf{z} \notin \mathcal{P}$, pak $a_i \neq 0$ pro nějaký index $i \geq m + 1$. Pro tento index i pak platí $(\mathbf{z}, \mathbf{e}_i) = a_i \neq 0$ a tedy $\mathbf{z} \notin (\mathcal{P}^{\perp})^{\perp}$. Tím je dokázána opačná inkluze $(\mathcal{P}^{\perp})^{\perp} \subseteq \mathcal{P}$. $\square$

Věta 11.59. Buď $\mathcal{V}$ euklidův prostor dimenze n a $\mathcal{P}$ jeho podprostor dimenze $m < n$. Buď $\mathbf{x} \in \mathcal{P}$ libovolný prvek, který neleží v $\mathcal{P}$. Potom existuje právě jeden vektor $\mathbf{x}_0 \in \mathcal{P}$ takový, že

$$||\mathbf{x} - \mathbf{x}_0|| < ||\mathbf{x} - \mathbf{y}||$$

pro libovolný prvek $\mathbf{y} \in \mathcal{P}$ různý od $\mathbf{x}_0$. Pro tento prvek $\mathbf{x}_0$ platí, že $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}^{\perp}$.

Vektor $\mathbf{x}_0$ se nazývá *pata kolmice* spuštěné z vektoru $\mathbf{x}$ na podprostor $\mathcal{P}$ nebo také *ortogonální projekce* vektoru $\mathbf{x}$ na podprostor $\mathcal{P}$.

Důkaz. Je-li $\mathbf{x}_0 \in \mathcal{P}$ takový vektor, že $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}^{\perp}$ a $\mathbf{y} \neq \mathbf{x}_0$ libovolný prvek $\mathcal{P}$, pak $\mathbf{x}_0, \mathbf{y} \in \mathcal{P}$ a tedy $\mathbf{x} - \mathbf{x}_0$ je ortogonální k $\mathbf{x}_0 - \mathbf{y}$. Podle Důsledku 11.52 (Pythagorova věta) dostáváme

$$||\mathbf{x} - \mathbf{y}||^2 = ||\mathbf{x} - \mathbf{x}_0||^2 + ||\mathbf{x}_0 - \mathbf{y}||^2 > ||\mathbf{x} - \mathbf{x}_0||^2.$$

K dokončení důkazu Věty 11.59 stačí ukázat, že existuje vektor $\mathbf{x}_0 \in \mathcal{P}$ takový, že $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}^{\perp}$. K tomu zvolíme nějakou bázi $\mathbf{e}_1, \ldots, \mathbf{e}_m$ podprostoru $\mathcal{P}$. Potom můžeme vyjádřit hledaný vektor $\mathbf{x}_0 \in \mathcal{P}$ ve tvaru

$$\mathbf{x}_0 = c_1\mathbf{e}_1 + \cdots + c_m\mathbf{e}_m,$$

kde $c_1, \ldots, c_m$ jsou nějaká reálná čísla. Má-li být $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}^{\perp}$, musí platit $(\mathbf{x} - \mathbf{x}_0, \mathbf{e}_i) = 0$ pro $i = 1, \ldots, m$, neboli $(\mathbf{x}_0, \mathbf{e}_i) = (\mathbf{x}, \mathbf{e}_i)$ pro $i = 1, \ldots, m$. Dosadíme-li za $\mathbf{x}_0$ jeho vyjádření pomocí báze $\mathbf{e}_1, \ldots, \mathbf{e}_m$, dostaneme dále

$$(c_1\mathbf{e}_1 + \cdots + c_m\mathbf{e}_m, \mathbf{e}_i) = c_1(\mathbf{e}_1, \mathbf{e}_i) + \cdots + c_m(\mathbf{e}_m, \mathbf{e}_i) = (\mathbf{x}, \mathbf{e}_i)$$

pro $i = 1, \ldots, m$. Neznámé koeficienty $c_1, \ldots, c_m$ tak musí splňovat soustavu m lineárních rovnic o m neznámých.

Předpokládejme dále, že $\mathbf{e}_1, \ldots, \mathbf{e}_m$ je ortonormální báze prostoru $\mathcal{P}$. Potom i-tá rovnice naší soustavy se rovná

$$(\mathbf{x}, \mathbf{e}_i) = c_i,$$

takže naše soustava má v tomto případě právě jedno řešení $c_1, \ldots, c_m$. Protože v každém euklidově prostoru existuje ortonormální báze, je tím dokázána existence vektoru $\mathbf{x}_0 \in \mathcal{P}$, pro který platí $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}^\perp$. Tím je věta dokázána. $\square$

Vraťme se ještě k soustavě

$$c_1(\mathbf{e}_1, \mathbf{e}_i) + c_2(\mathbf{e}_2, \mathbf{e}_i) + \cdots + c_m(\mathbf{e}_m, \mathbf{e}_i) = (\mathbf{x}, \mathbf{e}_i),$$

$i = 1, \ldots, m$, kde $\mathbf{e}_1, \ldots, \mathbf{e}_m$ je libovolná (ne nutně ortonormální) báze podprostoru $\mathcal{P}$. Jsou-li reálná čísla $c_1, \ldots, c_m$ řešením této soustavy, pak vektor $\mathbf{x}_0 = \sum_{i=1}^m c_i \mathbf{e}_i$ má tu vlastnost, že $\mathbf{x} - \mathbf{x}_0 \in \mathcal{P}^\perp$. Podle předchozí věty takový vektor existuje právě jeden. Naše soustava má tedy rovněž právě jedno řešení. Odtud plyne, že její matice

$$\begin{pmatrix} (\mathbf{e}_1, \mathbf{e}_1) & (\mathbf{e}_2, \mathbf{e}_1) & \ldots & (\mathbf{e}_m, \mathbf{e}_1) \\ (\mathbf{e}_1, \mathbf{e}_2) & (\mathbf{e}_2, \mathbf{e}_2) & \ldots & (\mathbf{e}_m, \mathbf{e}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{e}_1, \mathbf{e}_m) & (\mathbf{e}_2, \mathbf{e}_m) & \ldots & (\mathbf{e}_m, \mathbf{e}_m) \end{pmatrix}$$

je regulární.

Poslední matici můžeme definovat pro libovolný systém m vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_m$ (i v případě, že netvoří bázi). Tato matice se nazývá *Grammova matice* systému vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_m$.

Metoda nejmenších čtverců. Předpokládejme, že y a $x_1, \ldots, x_m$ jsou nějaké proměnné veličiny, které mohou nabývat reálných hodnot. Dále předpokládáme, že veličina y je lineárně závislá na veličinách $x_1, \ldots, x_m$, tj. platí

$$y = c_1 x_1 + c_2 x_2 + \cdots + c_m x_m,$$

kde $c_1, \ldots, c_m$ jsou nějaká reálná čísla. Tato čísla neznáme, můžeme ale provádět měření veličin $y, x_1, \ldots, x_m$. Provedeme n nezávislých měření a nechť při k-tém měření má proměnná y hodnotu y_k a proměnné $x_1, \ldots, x_n$ mají hodnoty $x_{1k}, \ldots, x_{mk}$. Koeficienty $c_1, \ldots, c_m$ se pak můžeme pokusit vypočítat ze soustavy

$$c_1 x_{1k} + c_2 x_{2k} + \cdots + c_m x_{mk} = y_k,$$

kde $k = 1, \ldots, n$. Budeme dále předpokládat, že počet měření n je větší než m. Naše soustava má tedy více rovnic než neznámých. Měření hodnot proměnných veličin $y, x_1, \ldots, x_m$ jsou navíc zatížena chybami, a tak je pravděpodobné, že naše soustava nebude mít žádné řešení.

Pokusíme se najít koeficienty $c_1, \ldots, c_m$ tak, aby si levé a pravé strany naší soustavy byly "co nejblíže". "Blízkost" můžeme měřit různými způsoby. Jedna z často používaných měr blízkosti je *kvadratická odchylka* levých stran naší soustavy od pravých stran. Kvadratická odchylka se definuje jako

$$\sum_{k=1}^{n}(c_1x_{1k} + c_2x_{2k} + \cdots + c_mx_{mk} - y_k)^2.$$

Budeme hledat koeficienty $c_1, \ldots, c_m$ tak, aby kvadratická odchylka levých stran soustavy od pravých byla co nejmenší. Vezmeme si standardní n-dimenzionální euklidův prostor $\mathcal{R}^n$ a v něm vektory $\mathbf{e}_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ pro $i = 1, \ldots, m$, a vektor $\mathbf{f} = (y_1, y_2, \ldots, y_n)$.

Levé strany naší soustavy jsou souřadnice vektoru $c_1\mathbf{e}_1 + \cdots c_m\mathbf{e}_m$ a pravé strany soustavy jsou souřadnice vektoru $\mathbf{f}$. Kvadratická odchylka levých stran soustavy od pravých se potom rovná druhé mocnině euklidovy vzdálenosti vektorů $c_1\mathbf{e}_1 + \cdots c_m\mathbf{e}_m$ a $\mathbf{f}$. Označme $\mathcal{P}$ podprostor $\mathcal{R}^n$ generovaný vektory $\mathbf{e}_1, \ldots, \mathbf{e}_m$. Podle Věty 11.59 existuje právě jeden vektor $\mathbf{f}_0 \in \mathcal{P}$, který má od $\mathbf{f}$ mezi všemi vektory podprostoru $\mathcal{P}$ nejmenší vzdálenost. Je to pata kolmice spuštěné z vektoru $\mathbf{f}$ na podprostor $\mathcal{P}$. Tuto patu kolmice můžeme vypočítat pomocí postupu z důkazu Věty 11.59 a poznámek následujících, pokud jsou vektory $\mathbf{e}_1, \ldots, \mathbf{e}_m$ lineárně nezávislé. Koeficienty $c_1, \ldots, c_m$ jsou potom určené jednoznačně. Jsou řešením soustavy

$$c_1(\mathbf{e}_1, \mathbf{e}_i) + \cdots + c_m(\mathbf{e}_m, \mathbf{e}_i) = (\mathbf{f}, \mathbf{e}_i),$$

$i = 1, \ldots, m$, kde

$$(\mathbf{f}, \mathbf{e}_i) = \sum_{k=1}^{n} x_{ik}y_k, \quad (\mathbf{e}_j, \mathbf{e}_i) = \sum_{k=1}^{n} x_{jk}x_{ik}.$$

Pokud nejsou vektory $\mathbf{e}_1, \ldots, \mathbf{e}_m$ lineárně nezávislé, má jednoznačně určená ortogonální projekce vektoru $\mathbf{f}$ na podprostor $\mathcal{P}$ různá vyjádření jako lineární kombinace vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_m$. Proto také koeficienty $c_1, \ldots, c_m$ nejsou v tomto případě určené jednoznačně. Pro výpočet koeficientů $c_1, \ldots, c_m$ napřed najdeme nějakou lineárně nezávislou podmnožinu množiny vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_m$ a potom postupujeme podobně jako v předchozím případě (lineárně nezávislých vektorů $\mathbf{e}_1, \ldots, \mathbf{e}_m$).

Příklad. Nechť jsou $(x_1, y_1), \ldots, (x_n, y_n)$ body v euklidově rovině. Najděte přibližné řešení soustavy rovnic

$$x_kc = y_k, \quad k = 1, \ldots, n,$$

metodou nejmenších čtverců. Jaký je geometrický význam konstanty c?