Kapitola 9

Skalární součin

Skalární součin je nástroj, jak měřit velikost vektorů a úhly mezi vektory v reálných a komplexních vektorových prostorech.

Definice 9.1 *Je-li* $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathcal{R}^{n\times 1}$ *reálný vektor dimenze* n*, pak definujeme* euklidovskou normu (délku) *vektoru* $\mathbf{x}$ *jako reálné číslo*

$$\|\mathbf{x}\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} = \sqrt{\mathbf{x}^T\mathbf{x}}.$$

Pokud je $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathcal{C}^{n\times 1}$ *komplexní vektor dimenze* n*, pak jeho* euklidovská norma (délka) *je reálné číslo*

$$\|\mathbf{x}\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} = \sqrt{\mathbf{x}^*\mathbf{x}}.$$

Cvičení 9.1 *Dokažte, že pro euklidovskou normu každého reálného (komplexního) vektoru* $\mathbf{x}$ *platí*

- $\|\mathbf{x}\| \geq 0$,
- $\|\mathbf{x}\| = 0$ *právě když* $\mathbf{x} = \mathbf{0}$,
- $\|a\mathbf{x}\| = |a| \cdot \|\mathbf{x}\|$ *pro libovolný skalár* $a \in \mathbf{R}(\mathbf{C})$.

Pro každý reálný (komplexní) vektor $\mathbf{x} \neq \mathbf{0}$ definujeme *jednotkový vektor stejného směru* jako vektor $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|$. Název jednotkový vyplývá ze skutečnosti, že

$$\|\mathbf{u}\| = \left\|\frac{\mathbf{x}}{\|\mathbf{x}\|}\right\| = \frac{1}{\|\mathbf{x}\|}\|\mathbf{x}\| = 1.$$

Říkáme také, že jsme vektor $\mathbf{x}$ *normalizovali*, nebo že vektor $\mathbf{u}$ je *normalizovaný* vektor $\mathbf{x}$.

Definice normy vektoru nám umožňuje také definovat *vzdálenost dvou vektorů* $\mathbf{x}, \mathbf{y} \in \mathcal{R}^{n\times 1}(\mathcal{C}^{n\times 1})$ jako $\|\mathbf{x} - \mathbf{y}\|$.

Definice 9.2 Standardní skalární součin *libovolných dvou reálných vektorů* $\mathbf{x} = (x_1, \ldots, x_n)^T$ *a* $\mathbf{y} = (y_1, \ldots, y_n)^T$ *definujeme jako reálné číslo*

$$\mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Jsou-li $\mathbf{x} = (x_1, \ldots, x_n)^T$ *a* $\mathbf{y} = (y_1, \ldots, y_n)^T$ *komplexní vektory, pak jejich* standardní skalární součin *je komplexní číslo*

$$\mathbf{x}^*\mathbf{y} = \sum_{i=1}^{n} \overline{x}_i y_i.$$

V této definici označuje $\overline{x}$ číslo komplexně sdružené k číslu x. Všimněte si rovněž, že skalární součin dvou reálných vektorů je speciálním případem skalárního součinu dvou komplexních vektorů.

Vztah mezi normou a skalárním součinem udává *Cauchyova-Schwarzova-Bunjakovského* nerovnost, zkráceně *CSB-nerovnost.*

Věta 9.3 *Pro libovolné dva vektory* $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$ *platí*

$$|\mathbf{x}^*\mathbf{y}| \le \|\mathbf{x}\| \cdot \|\mathbf{y}\|,$$

přičemž rovnost nastává právě když

$$\mathbf{y} = \frac{\mathbf{x}^*\mathbf{y}}{\mathbf{x}^*\mathbf{x}} \cdot \mathbf{x}.$$

Důkaz. Nerovnost zřejmě platí, pokud je $\mathbf{x} = \mathbf{0}$. V případě $\mathbf{x} \neq \mathbf{0}$ položíme

$$a = \frac{\mathbf{x}^*\mathbf{y}}{\mathbf{x}^*\mathbf{x}} = \frac{\mathbf{x}^*\mathbf{y}}{\|\mathbf{x}\|^2}.$$

Potom platí

$$\mathbf{x}^*(a\mathbf{x} - \mathbf{y}) = a\mathbf{x}^*\mathbf{x} - \mathbf{y}^*\mathbf{x} = \frac{\mathbf{x}^*\mathbf{y}}{\|\mathbf{x}\|^2} \cdot \|\mathbf{x}\|^2 - \mathbf{x}^*\mathbf{y} = 0.$$

Proto

$$\begin{aligned} 0 &\leq \|a\mathbf{x}-\mathbf{y}\|^2 = (a\mathbf{x}-\mathbf{y})^*(a\mathbf{x}-\mathbf{y}) = \overline{a}\mathbf{x}^*(a\mathbf{x}-\mathbf{y}) - \mathbf{y}^*(a\mathbf{x}-\mathbf{y}) = \\ &= -\mathbf{y}^*(a\mathbf{x}-\mathbf{y}) = \mathbf{y}^*\mathbf{y} - a\mathbf{y}^*\mathbf{x} = \frac{\|\mathbf{y}\|^2 \cdot \|\mathbf{x}\|^2 - (\mathbf{x}^*\mathbf{y})(\mathbf{y}^*\mathbf{x})}{\|\mathbf{x}\|^2}. \end{aligned}$$

Vzhledem k tomu, že $\mathbf{y}^*\mathbf{x} = \overline{\mathbf{x}^*\mathbf{y}}$, platí $(\mathbf{x}^*\mathbf{y})(\mathbf{y}^*\mathbf{x}) = |\mathbf{x}^*\mathbf{y}|^2$, takže

$$0 \leq \frac{\|\mathbf{y}\|^2 \cdot \|\mathbf{x}\|^2 - \|\mathbf{x}^*\mathbf{y}\|^2}{\|\mathbf{x}\|^2}$$

a odtud vyplývá $0 \leq \|\mathbf{y}\|^2 \cdot \|\mathbf{x}\|^2 - \|\mathbf{x}^*\mathbf{y}\|^2$, tj. $|\mathbf{x}^*\mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$. Rovnost nastává právě když $0 = \|a\mathbf{x}-\mathbf{y}\|^2$, tj. právě když $\mathbf{y} = a\mathbf{x}$. $\square$

CSB-nerovnost umožňuje dokázat následující *trojúhelníkovou nerovnost* pro reálné a komplexní vektory libovolné dimenze.

Tvrzení 9.4 *Pro libovolné dva vektory* $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$ *platí*

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

Důkaz. Pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$ platí

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y})^*(\mathbf{x}+\mathbf{y}) = \mathbf{x}^*\mathbf{x} + \mathbf{x}^*\mathbf{y} + \mathbf{y}^*\mathbf{x} + \mathbf{y}^*\mathbf{y} = \\ &= \|\mathbf{x}\|^2 + \mathbf{x}^*\mathbf{y} + \mathbf{y}^*\mathbf{x} + \|\mathbf{y}\|^2. \end{aligned}$$

Protože $\mathbf{y}^*\mathbf{x} = \overline{\mathbf{x}^*\mathbf{y}}$, dostáváme s použitím CSB-nerovnosti

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

Platí proto

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

$\square$

CSB-nerovnost také dovoluje definovat úhly mezi každými dvěma nenulovými reálnými vektory $\mathbf{x} = (x_1, \ldots, x_n)^T$ a $\mathbf{y} = (y_1, \ldots, y_n)^T$ libovolné dimenze n jako úhel $\alpha \in [0, \pi]$, pro který platí

$$\cos\alpha = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}.$$

Protože funkce $\cos\alpha$ je vzájemně jednoznačná na intervalu $[0, \pi]$, určuje poslední rovnost úhel α jednoznačně. Všimněte si také, že v případě $n = 2$

je tato definice v souladu s kosinovou větou pro trojúhelník, jehož vrcholy jsou počátek souřadnic a koncové body vektorů $\mathbf{x}$ a $\mathbf{y}$ úhel α je úhel sevřený vektory $\mathbf{x}$ a $\mathbf{y}$. V tom případě totiž kosinová věta říká, že

$$\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos\alpha.$$

Protože

$$\begin{aligned} \|\mathbf{x}-\mathbf{y}\|^2 &= (\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y}) = \mathbf{x}^T\mathbf{x} + \mathbf{y}^T\mathbf{y} - \mathbf{x}^T\mathbf{y} - \mathbf{y}^T\mathbf{x} = \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x}^T\mathbf{y}, \end{aligned}$$

dostaneme po dosazení do kosinové věty

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos\alpha,$$

tj.

$$\frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} = \cos\alpha.$$

Euklidovská norma není jedinou používanou mírou velikosti vektorů. Následující definice ukazuje, že pro každé reálné číslo $p \geq 1$ lze definovat normu vektorů z prostoru $\mathcal{C}^{n\times 1}$.

Definice 9.5 *Pro $p \geq 1$ definujeme p-*normu *vektorů $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathcal{C}^{n\times 1}$ jako reálné číslo*

$$\|\mathbf{x}\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}.$$

Euklidovská norma $\|\mathbf{x}\|$ vektoru $\mathbf{x}$ se tak rovná $\|\mathbf{x}\|_2$. Lze dokázat následující vlastnosti p-norem:

- $\|\mathbf{x}\|_p \geq 0$ a $\|\mathbf{x}\|_p = 0$ právě když $\mathbf{x} = \mathbf{0}$.
- $\|a\mathbf{x}\|_p = |a| \cdot \|\mathbf{x}\|_p$ pro každá skalár $a \in \mathbf{C}$ a každý vektor $\mathbf{x}$,
- $\|\mathbf{x}+\mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$ pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$.

CBS-nerovnost pro p-normy má tvar následující *Hölderovy nerovnosti.* Jsou-li $p, q > 1$ reálná čísla, pro která platí $1/p + 1/q = 1$, pak pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$ platí

$$|\mathbf{x}^*\mathbf{y}| \leq \|\mathbf{x}\|_p \cdot \|\mathbf{y}\|_q.$$

Viděli jsme, že takto definované p-normy splňují podmínky, které jsou požadovány od obecných vektorových norem ve smyslu následující definice.

Definice 9.6 Obecná vektorová norma *na reálném nebo komplexním vektorovém prostoru $\mathcal{V}$ je zobrazení, které každému vektoru $\mathbf{x} \in \mathcal{V}$ přiřazuje reálné číslo $\mathbf{x} \in \mathbf{R}$, a které splňuje následující podmínky*

- $\|\mathbf{x}\| \geq 0$ *a* $\|\mathbf{x}\| = 0$ *právě když* $\mathbf{x} = \mathbf{0}$.
- $\|a\mathbf{x}\| = |a| \cdot \|\mathbf{x}\|$ *pro každá skalár a a každý vektor* $\mathbf{x} \in \mathcal{V}$,
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ *pro každé dva vektory* $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

Prostory se skalárním součinem

Standardní skalární součin dvou vektorů v aritmetickém vektorovém prostoru nad reálnými nebo komplexními čísly závisí na souřadnicích těchto vektorů vzhledem ke standardní bázi. Následující definice ukazuje, jak lze definovat skalární součin na reálném nebo komplexním vektorovém prostoru obecně bez předchozí volby báze tohoto prostoru.

Definice 9.7 Skalární součin *na reálném (komplexním) vektorovém prostoru $\mathcal{V}$ je funkce, která každé uspořádané dvojici $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ přiřazuje reálné (komplexní) číslo $< \mathbf{x}|\mathbf{y} >$ a splňuje následující podmínky*

- $< \mathbf{x}|\mathbf{x} > \geq 0$ *pro každý vektor* $\mathbf{x} \in \mathcal{V}$,
- $< \mathbf{x}|\mathbf{x} > = 0$ *právě když* $\mathbf{x} = \mathbf{0}$,
- $< \mathbf{x}|a\mathbf{y} > = a < \mathbf{x}|\mathbf{y} >$ *pro libovolný skalár a a vektory* $\mathbf{x}, \mathbf{y} \in \mathcal{V}$,
- $< \mathbf{x}|\mathbf{y} + \mathbf{z} > = < \mathbf{x}|\mathbf{y} > + < \mathbf{x}|\mathbf{z} >$ *pro každé vektory* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$,
- $< \mathbf{x}|\mathbf{y} > = \overline{< \mathbf{y}|\mathbf{x} >}$ *pro* $\mathbf{x}, \mathbf{y} \in \mathcal{V}$, *(v případě reálného prostoru $\mathcal{V}$ to znamená, že $< \mathbf{x}|\mathbf{y} > = < \mathbf{y}|\mathbf{x} >$).*

Všimněte si, že z poslední podmínky plyne $< \mathbf{x}|\mathbf{x} > \in \mathbf{R}$ bez ohledu na to, je-li $\mathcal{V}$ reálný nebo komplexní vektorový prostor.

Cvičení 9.2 *Dokaže, že v každém prostoru $\mathcal{V}$ se skalárním součinem platí*

- $< \mathbf{x} + \mathbf{y}|\mathbf{z} > = < \mathbf{x}|\mathbf{z} > + < \mathbf{y}|\mathbf{z} >$ *pro libovolné tři vektory* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$,
- $< a\mathbf{x}|\mathbf{y} > = \overline{a} < \mathbf{x}|\mathbf{y} >$ *pro každý skalár a a vektory* $\mathbf{x}, \mathbf{y} \in \mathcal{V}$,
- $< \mathbf{0}|\mathbf{x} > = < \mathbf{x}|\mathbf{0} > = 0$ *pro každý vektor* $\mathbf{x} \in \mathcal{V}$.

Příklad 9.8 *Aritmetické prostory $\mathcal{R}^{n\times 1}$ a $\mathcal{C}^{n\times 1}$ se standardním skalární součinem jsou příkladem obecných prostorů se skalárním součinem, pokud definujeme $<\mathbf{x}|\mathbf{y}>=\mathbf{x}^T\mathbf{y}$, případně $<\mathbf{x}|\mathbf{y}>=\mathbf{x}^*\mathbf{y}$.*

Je-li $\mathbf{A}$ regulární komplexní matice řádu n, potom předpis $<\mathbf{x}|\mathbf{y}>=\mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{y}$ definuje skalární součin na aritmetickém prostoru $\mathcal{C}^{n\times 1}$. Říká se mu eliptický skalární součin.

Na prostoru $\mathcal{R}^{m\times n}$ reálných matic tvaru $m\times n$ můžeme definovat skalární součin předpisem

$$<\mathbf{A}|\mathbf{B}>=tr(\mathbf{A}^T\mathbf{B}),$$

kde $tr(\mathbf{C})$ označuje stopu *čtvercové matice $\mathbf{C}=(c_{ij})$ řádu n, která je definována jako součet prvků na hlavní diagonále matice $\mathbf{C}$*

$$tr(\mathbf{C})=\sum_{i=1}^{n} a_{ii}.$$

Podobně lze definovat skalární součin na prostoru komplexních matic $\mathcal{C}^{m\times n}$ předpisem

$$<\mathbf{A}|\mathbf{B}>=tr(\mathbf{A}^*\mathbf{B}).$$

Podobně jako standardní skalární součin na aritmetickém vektorovém prostoru $\mathcal{C}^{n\times 1}$ definuje euklidovskou normu, lze definovat rovněž normu každého vektoru $\mathbf{x}\in\mathcal{V}$ v libovolném prostoru se skalárním součinem $\mathcal{V}$ jako

$$\|\mathbf{x}\|=\sqrt{<\mathbf{x}|\mathbf{x}>}.$$

Abychom dokázali, že takto definovaná norma splňuje podmínky obecné definice normy podle Definice 9.6, potřebujeme dokázat, že v libovolném prostoru se skalárním součinem platí variantna CSB=nerovnosti. Předtím ještě jedno cviceni.

Cvičení 9.3 *Jaká je norma vektoru $\mathbf{x}\in\mathcal{C}^{n\times 1}$ v prostoru s eliptickým skalární součinem určeným regulární maticí $\mathbf{A}$ řádu n?*

Jaká je norma matice $\mathbf{A}=(a_{ij})\in\mathcal{R}^{m\times n}$ určená skalárním součinem $<\mathbf{A}|\mathbf{B}>=tr(\mathbf{A}^T\mathbf{B})$?

Jaká je norma matice $\mathbf{A}=(a_{ij})\in\mathcal{C}^{m\times n}$ určená skalárním součinem $<\mathbf{A}|\mathbf{B}>=tr(\mathbf{A}^*\mathbf{B})$?

V následující věte je dokázána *obecná CSB-nerovnost.*

Věta 9.9 *Je-li $\mathcal{V}$ prostor se skalárním součinem a $\|\mathbf{x}\|$ je norma na $\mathcal{V}$ definovaná tímto skalárním součinem, pak*

$$|<\mathbf{x}|\mathbf{y}>| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

pro každé dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{V}$. Rovnost nastává právě když $\mathbf{y} = a\mathbf{x}$ pro

$$a = \frac{<\mathbf{x}|\mathbf{y}>}{\|\mathbf{x}\|^2}.$$

Důkaz. Můžeme postupovat stejně jako v případě důkazu Věty 9.3 s využitím Cvičení 9.2. Případ $\mathbf{x} = \mathbf{0}$ je zřejmý a pro $\mathbf{x} \neq \mathbf{0}$ položíme

$$a = \frac{<\mathbf{x}|\mathbf{y}>}{\|\mathbf{x}\|^2}$$

a spočítáme, že

$$<\mathbf{x}|a\mathbf{x}-\mathbf{y}>= a<\mathbf{x}|\mathbf{x}> - <\mathbf{x}|\mathbf{y}>=<\mathbf{x}|\mathbf{y}> - <\mathbf{x}|\mathbf{y}>= 0.$$

Proto

$$\begin{aligned} 0 &\leq \|a\mathbf{x}-\mathbf{y}\|^2 =< a\mathbf{x}-\mathbf{y}|a\mathbf{x}-\mathbf{y}>= \overline{a}<\mathbf{x}|a\mathbf{x}-\mathbf{y}> - <\mathbf{y}|a\mathbf{x}-\mathbf{y}>= \\ &= -<\mathbf{y}|a\mathbf{x}-\mathbf{y}>=<\mathbf{y}|\mathbf{y}> -a<\mathbf{y}|\mathbf{x}>= \\ &= \frac{\|\mathbf{y}\|^2 \cdot \|\mathbf{x}\|^2 - <\mathbf{x}|\mathbf{y}><\mathbf{y}|\mathbf{x}>}{\|\mathbf{x}\|^2}. \end{aligned}$$

Z rovnosti $<\mathbf{x}|\mathbf{y}>= \overline{<\mathbf{y}|\mathbf{x}>}$ vyplývá $<\mathbf{x}|\mathbf{y}><\mathbf{y}|\mathbf{x}>= |<\mathbf{x}|\mathbf{y}>|^2$, proto

$$0 \leq \|\mathbf{y}\|^2 \cdot \|\mathbf{x}\|^2 - <\mathbf{x}|\mathbf{y}><\mathbf{y}|\mathbf{x}>,$$

tj. $|<\mathbf{x}|\mathbf{y}>| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$. Rovnost nastává právě když nastává rovnost v prvním výpočtu důkazu, tj. právě když $a\mathbf{x}-\mathbf{y} = \mathbf{0}$, tj. $\mathbf{y} = a\mathbf{x}$. □

Důsledek 9.10 *Je-li $\mathcal{V}$ prostor se skalárním součinem, pak předpis*

$$\|\mathbf{x}\| = \sqrt{<\mathbf{x}|\mathbf{x}>}$$

definuje normu na prostoru $\mathcal{V}$, tj. splnňuje podmínky Definice 9.6.

Důkaz. Z definice skalárního součinu plyne, že $<\mathbf{x}|\mathbf{x}>\in \mathbf{R}$ je vždy nezáporné číslo, proto rovněž $\|\mathbf{x}\| \geq 0$ pro každý vektor $\mathbf{x} \in \mathcal{V}$. Rovněž $\|\mathbf{x}\| = 0$ právě když $<\mathbf{x}|\mathbf{x}>= 0$ právě když $\mathbf{x} = \mathbf{0}$.

Je-li a skalár a $\mathbf{x} \in \mathcal{V}$, pak

$$\|a\mathbf{x}\|^2 =< a\mathbf{x}|a\mathbf{x} >= a\overline{a} < \mathbf{x}|\mathbf{x} >= |a|^2 \cdot \|\mathbf{x}\|^2.$$

Pouze důkaz trojúhelníkové nerovnosti vyžaduje použití CSB-nerovnosti.

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= < \mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y} >= \\
&= < \mathbf{x}|\mathbf{x} > + < \mathbf{x}|\mathbf{y} > + < \mathbf{y}|\mathbf{x} > + < \mathbf{y}|\mathbf{y} >= \\
&= \|\mathbf{x}\|^2 + 2\mathrm{Re} < \mathbf{x}|\mathbf{y} > + \|\mathbf{y}\|^2 \le \|\mathbf{x}\|^2 + 2| < \mathbf{x}|\mathbf{y} > | + \|\mathbf{y}\|^2 \le \\
&\le \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.
\end{aligned}$$

□

Ortogonalita

Definice 9.11 *Dva vektory* $\mathbf{x},\mathbf{y} \in \mathcal{V}$ *v prostoru se skalárním součinem se nazývají* kolmé (ortogonální)*, pokud* $< \mathbf{x}|\mathbf{y} >= 0$*. Kolmost vektorů* $\mathbf{x},\mathbf{y}$ *zapisujeme* $\mathbf{x}\perp\mathbf{y}$*.*

Stejně jako standardní skalární součin umožňuje definovat úhel mezi dvěma vektory aritmetických vektorových prostorů $\mathcal{R}^{n\times 1}$ a $\mathcal{C}^{n\times 1}$ libovolné dimenze, můžeme také definovat úhel mezi každými dvěma vektory $\mathbf{x},\mathbf{y} \in \mathcal{V}$ v libovolném prostoru $\mathcal{V}$ se skalárním součinem jako jednoznačně určený úhel $\alpha \in [0,\pi]$, pro který platí

$$\cos\alpha = \frac{< \mathbf{x}|\mathbf{y} >}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}.$$

Z obecné CSB-nerovnosti plyne, že

$$\left| \frac{< \mathbf{x}|\mathbf{y} >}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \right| \le 1,$$

úhel α je tedy dobře definován. Tato definice úhlu mezi dvěma vektory je v souladu s definicí kolmosti, neboť platí $\mathbf{x}\perp\mathbf{y}$ právě když $< \mathbf{x}|\mathbf{y} >= 0$, což je právě když $\cos\alpha = 0$ pro úhel α mezi vektory $\mathbf{x},\mathbf{y}$, neboli $\alpha = \pi/2$.

Definice 9.12 *Posloupnost* $\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$ *vektorů prostoru* $\mathcal{V}$ *se skalárním součinem se nazývá* ortonormální*, jestliže platí*

$$< \mathbf{u}_i|\mathbf{u}_j >= \delta_{ij}$$

pro libovolné dva indexy $i,j = 1,2,\ldots,n$*.*

Tvrzení 9.13 *Každá ortonormální posloupnost* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ *vektorů prostoru* $\mathcal{V}$ *se skalárním součinem je lineárně nezávislá.*

Důkaz. Předpokládejme, že $a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0}$ pro nějaké skaláry $a_1, \ldots, a_n$. Potom pro každé $i = 1, 2, \ldots, n$ platí

$$\begin{aligned} 0 &= < \mathbf{u}_i | \sum_{j=1}^{n} > a_j\mathbf{u}_j = \sum_{j=1}^{n} < \mathbf{u}_i | a_j\mathbf{u}_j >= \\ &= \sum_{j=1}^{n} a_j < \mathbf{u}_i | \mathbf{u}_j >= \sum_{j=1}^{n} a_j\delta_{ij} = a_i. \end{aligned}$$

Posloupnost $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ je proto lineárně nezávislá. □

Známe-li nějakou ortonormální bázi $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ prostoru $\mathcal{V}$ se skalárním součinem, můžeme snadno spočítat souřadnice libovolného vektoru $\mathbf{x} \in \mathcal{V}$ vzhledem k této ortonormální bázi a rovněž snadno spočítáme skalární součin $< \mathbf{x}|\mathbf{y} >$ dvou vektorů $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

Tvrzení 9.14 *Je-li* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ *ortonormální báze prostoru* $\mathcal{V}$ *se skalárním součinem, pak pro každý vektor* $\mathbf{x} \in \mathcal{V}$ *platí*

$$\mathbf{x} =< \mathbf{u}_1|\mathbf{x} > \mathbf{u}_1 + < \mathbf{u}_2|\mathbf{x} > \mathbf{u}_2 + \cdots + < \mathbf{u}_n|\mathbf{x} > \mathbf{u}_n,$$

tj. i-tá souřadnice vektoru $\mathbf{x}$ *vzhledem k bázi* $\mathbf{u}_1, \ldots, \mathbf{u}_n$ *se rovná skalárnímu součinu* $< \mathbf{u}_1|\mathbf{x} >$ *pro každé* $i = 1, 2, \ldots, n$.

Pokud $\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots a_n\mathbf{u}_n$ *a* $\mathbf{y} = b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots b_n\mathbf{u}_n$*, pak*

$$< \mathbf{x}|\mathbf{y} >= \sum_{i=1}^{n} \overline{a}_i b_i.$$

Důkaz. Pokud

$$\mathbf{x} = \sum_{j=1}^{n} a_j\mathbf{u}_j,$$

pak pro každé $i = 1, 2, \ldots, n$ platí

$$< \mathbf{u}_i|\mathbf{x} >=< \mathbf{u}_i| \sum_{j=1}^{n} a_j\mathbf{u}_j >= \sum_{j=1}^{n} a_j < \mathbf{u}_i|\mathbf{u}_j >= \sum_{j=1}^{n} a_j\delta_{ij} = a_i.$$

Dále platí

$$< \mathbf{x}|\mathbf{y} > \;=\; < \sum_{i=1}^{n} a_i\mathbf{u}_i | \sum_{j=1}^{n} b_j\mathbf{u}_j >= \sum_{i=1}^{n} \overline{a}_i < \mathbf{u}_i | \sum_{j=1}^{n} b_j\mathbf{u}_j >=$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{a}_i b_j < \mathbf{u}_i|\mathbf{u}_j >= \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{a}_i b_j \delta_{ij} =$$
$$= \sum_{i=1}^{n}\overline{a}_i b_i.$$

□

Definice 9.15 *Je-li* $\mathbf{u}_1, \dots, \mathbf{u}_n$ *ortonormální báze prostoru* $\mathcal{V}$ *se skalárním součinem a* $\mathbf{x} \in \mathcal{V}$*, pak vyjádření*

$$\mathbf{x} =< \mathbf{u}_1|\mathbf{x} > \mathbf{u}_1+ < \mathbf{u}_2|\mathbf{x} > \mathbf{u}_2 + \cdots + < \mathbf{u}_n|\mathbf{x} > \mathbf{u}_n$$

nazýváme Fourierův rozklad *vektoru* $\mathbf{x}$ *a koeficienty* $< \mathbf{u}_i|\mathbf{x} >$ *nazýváme* Fourierovy koeficienty *vektoru* $\mathbf{x}$ *vzhledem k bázi* $\mathbf{u}_1, \dots, \mathbf{u}_n$.

Ukážeme si ještě následující variantu *Pythagorovy věty.*

Tvrzení 9.16 *Předpokládáme, že* $\mathcal{V}$ *je reálný vektorový prostor se skalárním součinem. Dva vektory* $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ *jsou ortogonální právě když* $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}^2\|$.

Důkaz. Jsou-li vektory $\mathbf{x}, \mathbf{y}$ ortogonální, platí $\mathbf{x} \perp \mathbf{y} = 0$. Potom

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= < \mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y} >= \\ &= < \mathbf{x}|\mathbf{y} > + < \mathbf{x}|\mathbf{y} > + < \mathbf{y}|\mathbf{x} > + < \mathbf{y}|\mathbf{y} >= \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}^2\|. \end{aligned}$$

Naopak, platí-li $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}^2\|$, potom

$$\|\mathbf{x}\|^2 + \|\mathbf{y}^2\| =< \mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y} >=< \mathbf{x}|\mathbf{y} > + < \mathbf{x}|\mathbf{y} > + < \mathbf{y}|\mathbf{x} > + < \mathbf{y}|\mathbf{y} >,$$

a protože v reálném prostoru se skalárním součinem platí $< \mathbf{x}|\mathbf{y} >=< \mathbf{y}|\mathbf{x} >$, plyne odtud

$$2 < \mathbf{x}|\mathbf{y} >= 0,$$

tj. $\mathbf{x} \perp \mathbf{y}$. □

Všimněte si, že v případě komplexního prostoru se skalárním součinem předchozí důkaz ukazuje, že z rovnosti $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}^2\|$ plyne pouze Re $< \mathbf{x}|\mathbf{y} >= 0$, nikoliv $< \mathbf{x}|\mathbf{y} >= 0$. Opačná implikace ale zůstává beze změny, pokud $\mathbf{x} \perp \mathbf{y} = 0$, pak $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}^2\|$.

Gramova-Schmidtova ortogonalizace

Ukázali jsme si, jak snadno lze spočítat souřadnice nějakého vektoru vzhledem k ortonormální bázi a jak snadné je spočítat velikost skalárního součinu dvou vektorů, známe-li souřadnice těchto vektorů vzhledem k nějaké ortonormální bázi. Zbývá se přesvědčit, že v každém vektorovém prostoru se skalárním součinem existuje nějaká ortonormální báze. Snadno se přesvědčíme, že standardní báze je ortonormální báze v reálném aritmetickém prostoru $\mathcal{R}^{n\times 1}$ dimenze n a stejně tak rovněž v komplexním aritmetickém prostoru $\mathcal{C}^{n\times 1}$ dimenze n. Nyní si ukážeme algoritmus, který najde ortonormální bázi v každém prostoru se skalárním součinem. Tento algoritmus se nazývá *Gramova-Schmidtova ortogonalizace.*

V každém vektorovém prostoru $\mathcal{V}$ dimenze n nad libovolným tělesem $\mathbf{T}$ existuje nějaká báze $\mathcal{B}:\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$. Je-li $\mathcal{V}$ prostor se skalárním součinem, znamená to, že těleso skalárů je buď těleso $\mathbf{R}$ reálných čísel nebo těleso $\mathbf{C}$ komplexních čísel. Gramova-Schmidtova ortogonalizace začíná s libovolnou bází $\mathcal{B}:\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$ vektorového prostoru $\mathcal{V}$ se skalárním součinem a sestrojí ortonormální bázi $\mathcal{O}:\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n$ tohoto prostoru, která navíc splňuje podmínku

- pro každé $k=1,\ldots,n$ je posloupnost $\mathcal{O}_k:\mathbf{u}_1,\ldots,\mathbf{u}_k$ ortonormální báze podprostoru $\mathcal{S}_k=\mathcal{L}(\mathbf{x}_1,\ldots,\mathbf{x}_k)$.

Budeme postupovat indukcí podle k. Pro $k=1$ definujeme vektor

$$\mathbf{u}_1=\frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}.$$

Vektor $\mathbf{x}_1\neq\mathbf{0}$, neboť je prvkem báze a tedy lineárně nezávislé posloupnosti. Zlomek je proto dobře definován a $\|\mathbf{u}_1\|=1$, posloupnost $\mathcal{O}_1:\mathbf{u}_1$ je proto ortonormální báze podprostoru $\mathcal{S}_1=\mathcal{L}(\mathbf{x}_1)=\mathcal{L}(\mathbf{u}_1)$.

Předpokládáme nyní, že jsme již sestrojili nějakou ortonormální bázi $\mathcal{O}_k:\mathbf{u}_1,\ldots,\mathbf{u}_k$ podprostoru $\mathcal{S}_k=\mathcal{L}(\mathbf{x}_1,\ldots,\mathbf{x}_k)=\mathcal{L}(\mathbf{u}_1,\ldots,\mathbf{u}_k)$ pro nějaké $k\geq 1$ a $k<n$. Budeme hledat vektor $\mathbf{u}_{k+1}$ ve tvaru

$$\mathbf{u}_{k+1}=\mathbf{x}_{k+1}-\sum_{i=1}^{k}a_i\mathbf{u}_i,$$

kde neznáme skaláry $a_1,\ldots,a_k$. Vektor $\mathbf{u}_{k+1}$ musí být kolmý na všechny vektory $\mathbf{u}_1,\ldots,\mathbf{u}_k$, což platí právě když pro každé $j=1,\ldots,k$ je

$$\begin{aligned} 0 &= <\mathbf{u}_j|\mathbf{u}_{k+1}> = <\mathbf{u}_j|\mathbf{x}_{k+1}-\sum_{i=1}^{k}a_i\mathbf{u}_i> = \\ &= <\mathbf{u}_j|\mathbf{x}_{k+1}> -\sum_{i=1}^{k}a_i<\mathbf{u}_j|\mathbf{u}_i> = <\mathbf{u}_j|\mathbf{x}_{k+1}> -a_j, \end{aligned}$$

tj.

$$a_j \quad = \quad < \mathbf{u}_j | \mathbf{x}_{k+1} > \quad \text{pro všechna} \quad j = 1, \dots, k.$$

Označíme

$$\nu_{k+1} = \|\mathbf{x}_{k+1} - \sum_{i=1}^{k} < \mathbf{u}_i | \mathbf{x}_{k+1} > \mathbf{u}_i\|.$$

Platí $\nu_{k+1} \neq 0$, neboť $\mathbf{x}_{k+1} \notin \mathcal{L}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k)$ podle indukčního předpokladu (rovnost lineárních obalů $\mathcal{L}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k)$) a předpokladu, že $\mathcal{B} : \mathbf{x}_1, \dots, \mathbf{x}_n$ je báze $\mathcal{V}$. Proto platí, že vektor $\mathbf{x}_{k+1} - \sum_{i=1}^{k} < \mathbf{u}_i | \mathbf{x}_{k+1} > \mathbf{u}_i \neq \mathbf{0}$. Můžeme tak položit

$$\mathbf{u}_{k+1} = \frac{\mathbf{x}_{k+1} - \sum_{i=1}^{k} < \mathbf{u}_i | \mathbf{x}_{k+1} > \mathbf{u}_i}{\nu_{k+1}}.$$

Vektor $\mathbf{u}_{k+1}$ je proto kolmý na všechny vektory $\mathbf{u}_j$ pro $j = 1, \dots, k$ a navíc $\|\mathbf{u}_{k+1}\| = 1$. Posloupnost $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}$ je tak ortonormální a proto také lineárně nezávislá podle Tvrzení 9.13.

Zbývá dokázat, že $\mathcal{L}(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}) = \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1})$. Z rovnosti

$$\mathbf{x}_{k+1} = \nu_{k+1} \mathbf{u}_{k+1} + \sum_{i=1}^{k} < \mathbf{u}_i | \mathbf{x}_{k+1} > \mathbf{u}_i$$

vyplývá, že

$$\mathbf{x}_{k+1} \in \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}).$$

Podle indukčního předpokladu dále pro každé $j = 1, \dots, k$ platí

$$\mathbf{x}_j \in \mathcal{L}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k) \subseteq \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}).$$

Platí proto

$$\mathcal{L}(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}) \subseteq \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}).$$

Protože jsou obě posloupnosti $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}$ a $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}$ lineárně nezávislé, mají oba lineární obaly dimenzi $k + 1$ a jsou si proto rovné. Tím je popis a důkaz správnosti Gramovy-Schmidtovy ortogonalizace dokončen.

V případě, že uvažujeme prostory $\mathcal{R}^{m \times 1}$ nebo $\mathcal{C}^{m \times 1}$ se standardním skalárním součinem a euklidovskou normou, můžeme Gramovu-Schmidtovu ortogonalizaci popsat pomocí matic. Je-li $\mathcal{B} : \mathbf{x}_1, \dots, \mathbf{x}_n$ lineárně nezávislá posloupnost v prostoru $\mathcal{C}^{m \times 1}$, dostaneme Gramovou-Schmidtovou ortogonalizací ortonormální bázi

$$\mathbf{u}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} \quad \text{a} \quad \mathbf{u}_{k+1} = \frac{\mathbf{x}_{k+1} - \sum_{i=1}^{k} (\mathbf{u}_i^* \mathbf{x}_{k+1}) \mathbf{u}_i}{\|\mathbf{x}_{k+1} - \sum_{i=1}^{k} (\mathbf{u}_i^* \mathbf{x}_{k+1}) \mathbf{u}_i\|} \quad \text{pro} \quad k = 0, \dots, n - 1.$$

podprostoru $\mathcal{S} = \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ prostoru $\mathcal{C}^{m\times 1}$.

Označíme matice

$$\mathbf{U}_1 = \mathbf{0}_{m\times 1} \quad \text{a} \quad \mathbf{U}_{k+1} = (\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_k) \quad \text{pro} \quad k = 1, \ldots, n-1,$$

a spočítáme, že

$$\mathbf{U}_{k+1}^*\mathbf{x}_{k+1} = \begin{pmatrix} \mathbf{u}_1^*\mathbf{x}_{k+1} \\ \mathbf{u}_2^*\mathbf{x}_{k+1} \\ \vdots \\ \mathbf{u}_k^*\mathbf{x}_{k+1} \end{pmatrix}$$

a

$$\mathbf{U}_{k+1}\mathbf{U}_{k+1}^*\mathbf{x}_{k+1} = \sum_{i=1}^{k} \mathbf{u}_i(\mathbf{u}_i^*\mathbf{x}_{k+1}) = \sum_{i=1}^{k} (\mathbf{u}_i^*\mathbf{x}_{k+1})\mathbf{u}_i$$

pro $k = 1, \ldots, n-1$. Proto platí pro tato k, že

$$\mathbf{x}_{k+1} - \sum_{i=1}^{k} (\mathbf{u}_i^*\mathbf{x}_{k+1})\mathbf{u}_i = \mathbf{x}_{k+1} - \mathbf{U}_{k+1}\mathbf{U}_{k+1}^*\mathbf{x}_{k+1} = (\mathbf{I}_m - \mathbf{U}_{k+1}\mathbf{U}_{k+1}^*)\mathbf{x}_{k+1}.$$

Platí rovněž

$$\mathbf{x}_1 = \mathbf{I}_m\mathbf{x}_1 = (\mathbf{I}_m - \mathbf{U}_1\mathbf{U}_1^*)\mathbf{x}_1,$$

dostáváme tak jednotné vyjádření pro výsledek Gramovy-Schmidtovy ortogonalizace

$$\mathbf{u}_k = \frac{(\mathbf{I}_m - \mathbf{U}_k\mathbf{U}_k^*)\mathbf{x}_k}{\|(\mathbf{I}_m - \mathbf{U}_k\mathbf{U}_k^*)\mathbf{x}_k\|} \quad \text{pro} \quad k = 1, 2, \ldots, n.$$

Klasický Gramův-Schmidtův algoritmus

Je-li dána lineárně nezávislá posloupnost vektorů $\mathbf{x}_1, \ldots, \mathbf{x}_n$, můžeme Gramovu-Schmidtovu ortogonalizaci popsat následujícím algoritmem

pro $k = 1$:

$$\mathbf{u}_1 \leftarrow \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}$$

pro $k = 2, \ldots, n$:

$$\mathbf{u}_k \leftarrow \mathbf{x}_k - \sum_{i=1}^{k-1} (\mathbf{u}_i^*\mathbf{x}_k)\mathbf{u}_i$$

$$\mathbf{u}_k \leftarrow \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.$$

Cvičení 9.4 *Použijte Gramův-Schmidtův algoritmus na několik posloupností lineárně nezávislých vektorů.*

Gramovu-Schmidtovu ortogonalizaci můžeme vyjádřit ještě jiným způsobem pomocí násobení matic. Je-li dána matice $\mathbf{X}_{m\times n} = (\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_n)$ s komplexními prvky a s lineárně nezávislými sloupci, a použijeme-li klasický Gramův-Schmidtův algoritmus na posloupnost $\mathbf{x}_1,\ldots,\mathbf{x}_n$ sloupcových vektorů matice $\mathbf{X}_{m\times n}$, dostaneme ortonormální bázi sloupcového prostoru $\mathcal{S}(\mathbf{X}_{m\times n})$ matice $\mathbf{X}_{m\times n}$:

$$\mathbf{q}_1 = \frac{\mathbf{x}_1}{\nu_1} \quad \text{a} \quad \mathbf{q}_k = \frac{\mathbf{x}_k - \sum_{i=1}^{k-1}(\mathbf{q}_i^*\mathbf{x}_k)\mathbf{q}_i}{\nu_k} \quad \text{pro} \quad k = 2,3,\ldots,n,$$

kde

$$\nu_1 = \|\mathbf{x}_1\| \quad \text{a} \quad \nu_k = \|\mathbf{x}_k - \sum_{i=1}^{k-1}(\mathbf{q}_i^*\mathbf{x}_k)\mathbf{q}_i\| \quad \text{pro} \quad k = 2,3,\ldots,n.$$

Vyjádření vektorů $\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_n$ můžeme přepsat do tvaru

$$\mathbf{x}_1 = \nu_1\mathbf{q}_1 \quad \text{a} \quad \mathbf{x}_k = \sum_{i=1}^{k-1}(\mathbf{q}_i^*\mathbf{x}_k)\mathbf{q}_i + \nu_k\mathbf{q}_k \quad \text{pro} \quad k = 2,3,\ldots,n.$$

Poslední vyjádření vektorů $\mathbf{x}_k$ jako lineární kombinace vektorů $\mathbf{q}_1,\ldots,\mathbf{q}_k$ pro $k = 1,\ldots,n$ můžeme vyjádřit pomocí násobení matic jako rovnost

$$(\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_n) = (\mathbf{q}_1|\mathbf{q}_2|\cdots|\mathbf{q}_n)\begin{pmatrix} \nu_1 & \mathbf{q}_1^*\mathbf{x}_2 & \mathbf{q}_1^*\mathbf{x}_3 & \cdots & \mathbf{q}_1^*\mathbf{x}_n \\ 0 & \nu_2 & \mathbf{q}_2^*\mathbf{x}_3 & \cdots & \mathbf{q}_2^*\mathbf{x}_n \\ 0 & 0 & \nu_3 & \cdots & \mathbf{q}_3^*\mathbf{x}_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \nu_n \end{pmatrix}.$$

Tento rozklad matice $\mathbf{X}_{m\times n}$ můžeme zapsat jako $\mathbf{X}_{m\times n} = \mathbf{Q}_{m\times n}\mathbf{R}_{n\times n}$, kde $\mathbf{Q}$ je matice s ortonormálními sloupci a $\mathbf{R}$ je čtvercová horní trojúhelníková matice s kladnými prvky na hlavní diagonále. Lze ukázat, že tyto dvě podmínky určují matice $\mathbf{Q}$ a $\mathbf{R}$ jednoznačně. Tomuto rozkladu matice $\mathbf{X}$ říkáme *QR-faktorizace* matice $\mathbf{X}$.

Cvičení 9.5 *Najděte QR-faktorizaci několika reálných a komplexních matic.*

Příklad 9.17 *Klasický Gramův-Schmidtův algoritmus není numericky stabilní. Použijeme-li jej na posloupnost vektorů*

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 10^{-3} \\ 10^{-3} \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 10^{-3} \\ 0 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 1 \\ 0 \\ 10^{-3} \end{pmatrix},$$

a pokud zaokrouhlujeme každý jednotlivý výpočet na tři platná místa, dostaneme posloupnost vektorů

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 10^{-3} \\ 10^{-3} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ -0,709 \\ -0,709 \end{pmatrix}.$$

Tento výsledek není uspokojivý, protože vektory $\mathbf{u}_2$ *a* $\mathbf{u}_3$ *nejsou příliš ortogonální.*

Modifikovaný Gramův-Schmidtův algoritmus

Numerickou stabilitu Gramovy-Schmidtovy ortogonalizace můžeme vylepšit tím, že přeuspořádáme jednotlivé kroky výpočtu. Z předchozího popisu klasické Gramovy-Schmidtovy ortogonalizace posloupnosti reálných, případně komplexních, vektorů $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ dimenze m dostáváme, že

$$\mathbf{u}_k = \frac{(\mathbf{I} - \mathbf{U}_k\mathbf{U}_k^*)\mathbf{x}_k}{\|(\mathbf{I} - \mathbf{U}_k\mathbf{U}_k^*)\mathbf{x}_k\|}, \quad \text{kde } \mathbf{U}_1 = \mathbf{0} \text{ a } \mathbf{U}_k = (\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_{k-1}) \text{ pro } k > 1.$$

Označíme matice $\mathbf{E}_1 = \mathbf{I}$ a $\mathbf{E}_i = \mathbf{I} - \mathbf{u}_{i-1}\mathbf{u}_{i-1}^*$ pro $i > 1$. Z ortogonality posloupnosti vektorů $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ vyplývá, že

$$\mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1 = \mathbf{I} - \mathbf{u}_1\mathbf{u}_1^* - \mathbf{u}_2\mathbf{u}_2^* - \cdots - \mathbf{u}_{k-1}\mathbf{u}_{k-1}^* = \mathbf{I} - \mathbf{U}_k\mathbf{U}_k^*.$$

Gramovu-Schmidtovu ortogonalizaci můžeme proto vyjádřit ve tvaru

$$\mathbf{u}_k = \frac{\mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{x}_k}{\|\mathbf{E}_k \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{x}_k\|} \quad \text{pro} \quad k = 1, 2, \ldots, n.$$

Odtud vyplývá, že Gramovu-Schmidtovu ortogonalizaci můžeme také uspořádat následovně:

$$\begin{aligned}
\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n &\longrightarrow \mathbf{u}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \\
&\longrightarrow \mathbf{u}_1, \mathbf{E}_2\mathbf{x}_2, \mathbf{E}_2\mathbf{x}_3, \ldots, \mathbf{E}_2\mathbf{x}_n \\
&\longrightarrow \mathbf{u}_1, \mathbf{u}_2, \mathbf{E}_2\mathbf{x}_3, \ldots, \mathbf{E}_2\mathbf{x}_n \\
&\longrightarrow \mathbf{u}_1, \mathbf{u}_2, \mathbf{E}_3\mathbf{E}_2\mathbf{x}_3, \ldots, \mathbf{E}_3\mathbf{E}_2\mathbf{x}_n \\
&\longrightarrow \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{E}_3\mathbf{E}_2\mathbf{x}_4, \ldots, \mathbf{E}_3\mathbf{E}_2\mathbf{x}_n \\
&\ \text{atd.}
\end{aligned}$$

Toto uspořádání vede k následující verzi ortogonalizačního algoritmu, který nazýváme *modifikovaná Gramova-Schmidtova ortogonalizace*. Máme-li dánu lineárně nezávislou posloupnost vektorů $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathcal{C}^{m\times 1}$, pak spočítáme ortonormální posloupnost $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ takto

pro $k = 1$:

$$\mathbf{u}_1 \leftarrow \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|},$$

pro $k > 1$:

$$\begin{aligned} \mathbf{u}_j &\leftarrow \mathbf{E}_k\mathbf{u}_j = \mathbf{u}_j - (\mathbf{u}_{k-1}^*\mathbf{u}_j)\mathbf{u}_{k-1} \quad \text{pro } j = k, k+1, \ldots, n, \\ \mathbf{u}_k &\leftarrow \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}. \end{aligned}$$

V exaktní aritmetice vedou oba algoritmy ke zcela stejnému výsledku. Rozdíl je v případě, kdy nepoužíváme exaktní aritmetiku a počítáme na pevný počet platných míst. Větší numerickou stabilitu takto modifikované Gramovy-Schmidtovy ortogonalizace můžeme demonstrovat v následujícím příkladě.

Příklad 9.18 *Použijeme-li modifikovaný Gramův-Schmidtův algoritmus na posloupnost vektorů* $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ *z Příkladu 9.17 a zaokrouhlujeme-li opět každý mezivýsledek ihned na tři platná místa, dostaneme posloupnost vektorů*

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 10^{-3} \\ 10^{-3} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix},$$

což je výsledek mnohem uspokojivější, než výsledek v Příkladu 9.17.

Ani modifikovaný Gramův-Schmidtův algoritmus není numericky stabilní. Existují příklady, kdy vede k neuspokojivým výsledkům, pokud nepoužíváme exaktní aritmetiku.

Unitární a ortogonální matice

Definice 9.19 *Čtvercová komplexní matice* $\mathbf{U}$ *řádu* n *se nazývá* unitární, *pokud její sloupce tvoří ortonormální bázi prostoru* $\mathcal{C}^{n\times 1}$.

Čtvercová reálná matice $\mathbf{P}$ *řádu* n *se nazývá* ortogonální, *pokud její sloupce tvoří ortonormální bázi prostoru* $\mathcal{R}^{n\times 1}$.

Název *ortogonální matice* je běžně používaný, přestože mnohem vhodnější by byl název *ortonormální matice*.

Věta 9.20 *Pro komplexní čtvercovou matici* $\mathbf{U}$ *řádu* n *je ekvivalentní*

1. $\mathbf{U}$ *je unitární,*

2. *posloupnost řádkových vektorů matice* $\mathbf{U}$ *je ortonormální,*

3. $\mathbf{U}^{-1} = \mathbf{U}^*$,

4. $\|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$ *pro každý vektor* $\mathbf{x} \in \mathcal{C}^{n\times 1}$.

Důkaz.

1 $\Leftrightarrow$ **3**. Matice $\mathbf{U}$ je unitární právě když $(\mathbf{U}_{*i})^*\mathbf{U}_{*j} = \delta_{ij}$. Řádkový vektor $(\mathbf{U}_{*i})^*$ se rovná i-tému řádku matice $\mathbf{U}^*$, proto $(\mathbf{U}^*)_{i*}\mathbf{U}_{*j} = (\mathbf{U}_{*i})^*\mathbf{U}_{*j} = \delta_{ij}$, tj. $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$. Připomeňme si, že δ_{ij} je *Kroneckerův symbol*, který se rovná 1, pokud $i = j$, a je rovný 0, pokud $i \neq j$.

3 $\Leftrightarrow$ **2**. Podle Tvrzení 3.11 platí $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$ právě když $\mathbf{U}\mathbf{U}^* = \mathbf{I}_n$, tj. právě když $\mathbf{U}_{i*}(\mathbf{U}^*)_{*j} = \delta_{ij}$. Protože sloupcový vektor $(\mathbf{U}^*)_{*j}$ se rovná vektoru $(\mathbf{U}_{j*})^*$, dostáváme $\mathbf{U}_{i*}(\mathbf{U}_{j*})^* = \mathbf{U}_{i*}(\mathbf{U}^*)_{*j} = \delta_{ij}$. Číslo $\mathbf{U}_{j*}(\mathbf{U}_{i*})^*$ je komplexně sdružené k čílsu $\mathbf{U}_{i*}(\mathbf{U}_{j*})^* = \delta_{ij}$, poslední rovnost je proto ekvivaletní s tím, že posloupnost řádkových vektorů $\mathbf{U}_{1*}, \mathbf{U}_{2*}, \ldots, \mathbf{U}_{n*}$ ortonormální.

3 $\Rightarrow$ **4**. Je-li $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$ a $\mathbf{x} \in \mathcal{C}^{n\times 1}$, pak

$$\|\mathbf{U}\mathbf{x}\|^2 = (\mathbf{U}\mathbf{x})^*\mathbf{U}\mathbf{x} = \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{x} = \mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2.$$

4 $\Rightarrow$ **1**. Platí-li pro matici $\mathbf{U}$ rovnost $\|\mathbf{U}\mathbf{x}\|^2 = \|\mathbf{x}\|^2$ pro každý vektor $\mathbf{x} \in \mathcal{C}^{n\times 1}$, pak také pro libovolné dva vektory $\mathbf{x}, \mathbf{y} \in \mathcal{C}^{n\times 1}$ platí

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

Pro normu vektoru $\|\mathbf{x}+\mathbf{y}\|^2$ platí

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y})^*(\mathbf{x}+\mathbf{y}) = \mathbf{x}^*\mathbf{x} + \mathbf{y}^*\mathbf{y} + \mathbf{x}^*\mathbf{y} + \mathbf{y}^*\mathbf{x} = \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \mathbf{x}^*\mathbf{y} + \overline{\mathbf{x}^*\mathbf{y}} = \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\cdot \mathrm{Re}(\mathbf{x}^*\mathbf{y}),
\end{aligned}$$

a pro normu vektoru $\mathbf{U}(\mathbf{x}+\mathbf{y})$ podobně dostáváme

$$\begin{aligned}
\|\mathbf{U}(\mathbf{x}+\mathbf{y})\|^2 &= (x+y)^*\mathbf{U}^*\mathbf{U}(\mathbf{x}+\mathbf{y}) = \\
&= \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{x} + \mathbf{y}^*\mathbf{U}^*\mathbf{U}\mathbf{y} + \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y} + \mathbf{y}^*\mathbf{U}^*\mathbf{U}\mathbf{x} = \\
&= \|\mathbf{U}\mathbf{x}\|^2 + \|\mathbf{U}\mathbf{y}\|^2 + \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y} + \overline{\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y}} = \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\cdot \mathrm{Re}(\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y}).
\end{aligned}$$

Z rovnosti $\|\mathbf{U}(\mathbf{x}+\mathbf{y})\|^2 = \|\mathbf{x}+\mathbf{y}\|^2$ tak vyplývá rovnost reálných částí $\mathrm{Re}(\mathbf{x}^*\mathbf{y}) = \mathrm{Re}(\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y})$.

Obdobně dokážeme také rovnost imaginárních částí obou standardních skalárních součinů $\mathbf{x}^*\mathbf{y}$ a $\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y}$, tj. rovnost $\mathrm{Im}(\mathbf{x}^*\mathbf{y}) = \mathrm{Im}(\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y})$. Tentokrát vyjdeme z rovnosti $\|\mathbf{x}+\imath\mathbf{y}\| = \|\mathbf{U}(\mathbf{x}+\imath\mathbf{y})\|$, kde $\imath$ je komplexní jednotka. S využitím rovnosti $\bar{\imath} = -\imath$ dostáváme

$$\begin{aligned}
\|\mathbf{x}+\imath\mathbf{y}\|^2 &= (\mathbf{x}+\imath\mathbf{y})^*(\mathbf{x}+\imath\mathbf{y}) = \\
&= \mathbf{x}^*\mathbf{x} + (\imath\mathbf{y})^*(\imath\mathbf{y}) + \mathbf{x}^*(\imath\mathbf{y}) + (\imath\mathbf{y})^*\mathbf{y} = \\
&= \|\mathbf{x}\|^2 + \bar{\imath}\imath\,\|\mathbf{y}\|^2 + \imath\,\mathbf{x}^*\mathbf{y} + \bar{\imath}\,\mathbf{y}^*\mathbf{x} = \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \imath\,\mathbf{x}^*\mathbf{y} - \imath\,\mathbf{y}^*\mathbf{x} = \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \imath\,(\mathbf{x}^*\mathbf{y} - \overline{\mathbf{x}^*\mathbf{y}}) = \\
&= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\imath\,\mathrm{Im}(\mathbf{x}^*\mathbf{y}).
\end{aligned}$$

Zcela analogicky dokážeme, že rovněž

$$\|\mathbf{U}(\mathbf{x}+\imath\mathbf{y})\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\imath\,\mathrm{Im}(\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y}).$$

Platí proto rovněž rovnost imaginárních částí $\mathrm{Im}(\mathbf{x}^*\mathbf{y}) = \mathrm{Im}(\mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y})$. Z rovnosti reálných a imaginárních částí obou skalárních součinů tak dostáváme rovnost

$$\mathbf{x}^*\mathbf{y} = \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{y}.$$

Nyní už snadno dokončíme důkaz implikace $4 \Rightarrow 1$. Pro libovolné dva indexy $i, j = 1, 2, \ldots, n$ platí

$$(\mathbf{U}_{*i})^*\mathbf{U}_{*j} = \mathbf{e}_i^*\mathbf{U}^*\mathbf{U}\mathbf{e}_j = \mathbf{e}_i^*\mathbf{e}_j = \delta_{ij}.$$

Posloupnost sloupcových vektorů matice $\mathbf{U}$ je proto ortonormální. $\square$

Podobně můžeme také dokázat ekvivalentní podmínky pro ortogonální matice. Stačí všude nahradit vektor $\mathbf{x}^*$ reálným řádkovým vektorem $\mathbf{x}^T$.

Věta 9.21 *Pro reálnou čtvercovou matici* $\mathbf{P}$ *řádu* n *je ekvivalentní*

1. $\mathbf{P}$ *je ortogonální,*
2. *posloupnost řádkových vektorů matice* $\mathbf{P}$ *je ortonormální,*
3. $\mathbf{P}^{-1} = \mathbf{P}^T$,
4. $\|\mathbf{P}\mathbf{x}\| = \|\mathbf{x}\|$ *pro každý vektor* $\mathbf{x} \in \mathcal{R}^{n\times 1}$.

Ortogonální doplněk

Definice 9.22 *Je-li $\mathcal{V}$ vektorový prostor se skalárním součinem a $\mathcal{P}$ podprostor $\mathcal{V}$, pak definujeme* ortogonální doplněk *podprostoru $\mathcal{P}$ jako množinu*

$$\mathcal{P}^{\perp} = \{\mathbf{u} \in \mathcal{V} : < \mathbf{x}|\mathbf{u} >= 0 \quad \text{pro každý vektor} \quad \mathbf{x} \in \mathcal{P}\}.$$

Následující tvrzení obsahuje základní vlastnosti ortogonálního doplňku.

Tvrzení 9.23 *Předpokládáme, že $\mathcal{P}$ je podprostor vektorového prostoru $\mathcal{V}$ se skalárním součinem. Potom platí*

1. $\mathcal{P}^{\perp}$ *je podprostor prostoru* $\mathcal{V}$,
2. $\mathcal{P}^{\perp} \cap \mathcal{P} = \{\mathbf{0}\}$,
3. *je-li* $\mathbf{u}_1, \ldots, \mathbf{u}_k$ *ortonormální báze podprostoru* $\mathcal{P}$ *a* $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ *ortonormální báze podprostoru* $\mathcal{P}^{\perp}$, *pak* $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ *je ortonormální báze prostoru* $\mathcal{V}$,
4. $\dim \mathcal{P}^{\perp} = \dim \mathcal{V} - \dim \mathcal{P}$,
5. $\mathcal{P}^{\perp\perp} = \mathcal{P}$.

Důkaz. 1. Jsou-li $\mathbf{u}, \mathbf{v} \in \mathcal{P}^{\perp}$ a a libovolný skalár, pak pro každý vektor $\mathbf{x} \in \mathcal{P}$ platí

$$< \mathbf{x}|\mathbf{u} + \mathbf{v} >=< \mathbf{x}|\mathbf{u} > + < \mathbf{x}|\mathbf{v} >= 0$$

a podobně

$$< \mathbf{x}|a\mathbf{u} >= a < \mathbf{x}|\mathbf{u} >= 0.$$

Ortogonální doplněk $\mathcal{P}^{\perp}$ je proto podprostor $\mathcal{V}$.

2. Je-li $\mathbf{u} \in \mathcal{P}^{\perp} \cap \mathcal{P}$, pak $< \mathbf{u}|\mathbf{u} >= 0$, tj. $\mathbf{u} = \mathbf{0}$.

3. Posloupnost $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ je ortonormální, a proto lineárně nezávislá podle Tvrzení 9.13. Dokážeme, že je bází prostoru $\mathcal{V}$. Uděláme to sporem. Kdyby platilo $\mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n) \neq \mathcal{V}$, doplnili bychom posloupnost $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ do báze prostoru $\mathcal{V}$ přidáním nějakých vektorů $\mathbf{x}_1, \ldots, \mathbf{x}_m$. Na lineárně nezávislou posloupnost $\mathbf{u}_1, \ldots, \mathbf{u}_k$, $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n, \mathbf{x}_1, \ldots, \mathbf{x}_m$ bychom použili Gramovu-Schmidtovu ortogonalizaci. Dostali bychom ortonormální posloupnost $\mathbf{u}_1, \ldots, \mathbf{u}_k$, $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$, $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Pro vektor $\mathbf{v}_1$ by pak platilo $< \mathbf{u}_1|\mathbf{v}_1 >= 0$ pro $i = 1, \ldots, k$ a tedy také

$$< \sum_{i=1}^{k} a_i\mathbf{u}_i|\mathbf{v}_1 >= \sum_{i=1}^{k} \overline{a}_i < \mathbf{u}_i|\mathbf{v}_1 >= 0,$$

a proto také $\mathbf{v}_1 \in \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_k)^\perp = \mathcal{P}^\perp$. To je ale ve sporu s tím, že $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ je báze $\mathcal{P}^\perp$ a $\mathbf{v}_1 \notin \mathcal{L}(\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n)$. Posloupnost $\mathbf{u}_1, \ldots, \mathbf{u}_k$, $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ je proto bází celého prostoru $\mathcal{V}$.

4. Podle části 3. dostáváme $\dim \mathcal{P}^\perp = n - k = \dim \mathcal{V} - \dim \mathcal{P}$.

5. Pro každý vektor $\mathbf{u} \in \mathcal{P}$ a $\mathbf{x} \in \mathcal{P}^\perp$ platí

$$< \mathbf{x}|\mathbf{u} >= \overline{< \mathbf{u}|\mathbf{x} >} = 0,$$

proto rovněž $\mathbf{u} \in \mathcal{P}^{\perp\perp}$, a tedy $\mathcal{P} \subseteq \mathcal{P}^{\perp\perp}$. Vzhledem k 4. platí také

$$\dim \mathcal{P}^{\perp\perp} = \dim \mathcal{V} - \dim \mathcal{P}^\perp = \dim \mathcal{V} - \dim \mathcal{V} + \dim \mathcal{P} = \dim \mathcal{P},$$

proto z právě dokázané inkluze vyplývá $\mathcal{P} = \mathcal{P}^{\perp\perp}$. $\Box$

Následující *věta o ortogonálním rozkladu* dává do souvislosti čtyři základní podprostory určené maticí a ortogonální doplněk.

Věta 9.24 *Pro každou matice* $\mathbf{A} \in \mathcal{R}^{m\times n}$ *platí*

- $\mathcal{S}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T)$,
- $\mathcal{N}(\mathbf{A})^\perp = \mathcal{S}(\mathbf{A}^T)$.

Podobně pro každou matici $\mathbf{B} \in \mathcal{C}^{m\times n}$ *platí*

- $\mathcal{S}(\mathbf{B})^\perp = \mathcal{N}(\mathbf{B}^*)$,
- $\mathcal{N}(\mathbf{B})^\perp = \mathcal{S}(\mathbf{B}^*)$.

Důkaz. Vektor $\mathbf{z} \in \mathcal{S}(\mathbf{A})$ právě když $\mathbf{z} = \mathbf{A}\mathbf{y}$ pro nějaký vektor $\mathbf{y} \in \mathcal{R}^{n\times 1}$. Platí tedy, že vektor $\mathbf{x} \in \mathcal{S}(\mathbf{A})^\perp$ právě když pro každý vektor $\mathbf{y} \in \mathcal{R}^{n\times 1}$ platí $< \mathbf{A}\mathbf{y}|\mathbf{x} >= 0$, což je právě když $\mathbf{y}^T\mathbf{A}^T\mathbf{x} = 0$, neboli právě když $< \mathbf{y}|\mathbf{A}^T\mathbf{x} >= 0$. Poslední rovnost nastává právě když $\mathbf{A}^T\mathbf{x} = 0$, neboli právě když $\mathbf{x} \in \mathcal{N}(\mathbf{A}^T)$. Tím je dokázána první z rovností.

Druhou rovnost dostaneme použitím první rovnosti na transponovanou matici $\mathbf{A}^T$ a z Tvrzení 9.23.5. Dostáváme tak

$$\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^{T^T}) = \mathcal{S}(\mathbf{A}^T)^\perp$$

a proto také

$$\mathcal{N}(\mathbf{A})^\perp = \mathcal{S}(\mathbf{A}^T).$$

Analogické rovnosti pro komplexní matici $\mathbf{B} \in \mathcal{C}^{m\times n}$ dokážeme stejným postupem, pouze všude nahradíme transponované matice $\mathbf{C}^T$ maticemi $\mathbf{C}^*$. $\Box$

Předpokládejme nyní, že $\mathbf{A} \in \mathcal{R}^{m\times n}$ je reálná matice a její hodnost $r(\mathbf{A}) = \dim \mathcal{S}(\mathbf{A}) = r$. Zvolme nějakou ortonormální bázi $\mathcal{B} : \mathbf{u}_1, \dots, \mathbf{u}_r$ sloupcového prostoru $\mathcal{S}(\mathbf{A})$ matice $\mathbf{A}$ a dále zvolme ortonormální bázi $\mathcal{B}' : \mathbf{u}_{r+1}, \dots, \mathbf{u}_m$ ortogonálního doplňku $\mathcal{S}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T)$. Posloupnost vektorů $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_m$ je potom ortonormální báze prostoru $\mathcal{R}^{m\times 1}$ podle Tvrzení 9.23.3. Matice $\mathbf{U} = (\mathbf{u}_1|\mathbf{u}_2|\cdots|\mathbf{u}_m)$ je potom regulární podle Věty 9.21.

Dále zvolíme ortonormální bázi $\mathbf{v}_1, \dots, \mathbf{v}_r$ prostoru $\mathcal{S}(\mathbf{A}^T)$ a ortonormální bázi $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ jeho ortogonálního doplňku $\mathcal{S}(\mathbf{A}^T)^\perp = \mathcal{N}(\mathbf{A})$. Matice $\mathbf{V} = (\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n)$ je potom rovněž regulární, ze stejného důvodu jako matice $\mathbf{U}$.

Nyní budeme uvažovat součin $\mathbf{R} = \mathbf{U}^T\mathbf{A}\mathbf{V}$. Je-li $\mathbf{R} = (r_{ij})$, pak $r_{ij} = \mathbf{u}_i^T\mathbf{A}\mathbf{v}_j$. Protože $\mathbf{u}_i^T\mathbf{A} = \mathbf{0}$ pro $i = r+1, \dots, m$ a $\mathbf{A}\mathbf{v}_j = \mathbf{0}$ pro $j = r+1, \dots, n$, dostáváme pro matici $\mathbf{R}$ vyjádření

$$\mathbf{R} = \mathbf{U}^T\mathbf{A}\mathbf{V} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{u}_1^T\mathbf{A}\mathbf{v}_r & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots & \ddots & \vdots \\ \mathbf{u}_r^T\mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{u}_r^T\mathbf{A}\mathbf{v}_r & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Protože $\mathbf{U}^{-1} = \mathbf{U}^T$ a $\mathbf{V}^{-1} = \mathbf{V}^T$ podle Věty 9.21, můžeme matici $\mathbf{A}$ vyjádřit ve tvaru

$$\mathbf{A} = \mathbf{U}\mathbf{R}\mathbf{V}^T = \mathbf{U}\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T.$$

Dále platí $r(\mathbf{C}) = r$, neboť

$$r(\mathbf{C}) = r\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = r(\mathbf{U}^T\mathbf{A}\mathbf{V}) = r(\mathbf{A}) = r,$$

protože součin libovolné matice s regulární maticí nemění hodnost této matice podle Úlohy 6.4 a Úlohy 6.5. Takto získanému rozkladu $\mathbf{A} = \mathbf{U}\mathbf{R}\mathbf{V}^T$ matice $\mathbf{A}$ říkáme *URV-faktorizace* matice $\mathbf{A}$.

Cvičení 9.6 *Najděte URV-faktorizaci několika matic.*

Každá množina ortonormálních bází čtyř základních podprostorů určených maticí $\mathbf{A}$ určuje URV-faktorizaci matice $\mathbf{A}$. Platí to také naopak, každá URV-faktorizace matice $\mathbf{A}$ určuje ortonormální báze všech čtyř podprostorů.

Máme-li totiž ortonormální matice $\mathbf{U} = (\mathbf{U}_1|\mathbf{U}_2)$ řádu m, $\mathbf{V} = (\mathbf{V}_1|\mathbf{V}_2)$ řádu n a regulární matici $\mathbf{C}$ řádu r, pro které platí

$$\mathbf{A} = \mathbf{URV}^T = \mathbf{U} \begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T,$$

pak podle Úlohy 6.4 platí

$$\mathcal{S}(\mathbf{A}) = \mathcal{S}(\mathbf{URV}^T) = \mathcal{S}(\mathbf{UR}) = \mathcal{S}(\mathbf{U}_1\mathbf{C}|\mathbf{0}) = \mathcal{S}(\mathbf{U}_1\mathbf{C}) = \mathcal{S}(\mathbf{U}_1).$$

Proto sloupcové vektory matice $\mathbf{U}_1$ tvoří ortonormální bázi sloupcového prostoru $\mathcal{S}(\mathbf{A})$ matice $\mathbf{A}$. Všechny sloupcové vektory matice $\mathbf{U}_2$ leží v ortogonálním doplňku $\mathcal{S}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T)$ a protože $\dim \mathcal{S}(\mathbf{A})^\perp = m - \dim \mathcal{S}(\mathbf{A})$, musí tvořit ortonormální bázi $\mathcal{S}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T)$.

Vynásobíme-li libovolnou matici $\mathbf{B}$ zleva regulární maticí (tj. provedeme-li posloupnost elementárních řádkových úprav matice $\mathbf{B}$), nezměníme nulový prostor $\mathcal{N}(\mathbf{B})$ matice $\mathbf{B}$. Platí proto

$$\begin{aligned} \mathcal{N}(\mathbf{A}) &= \mathcal{N}(\mathbf{URV}^T) = \mathcal{N}(\mathbf{RV}^T) = \mathcal{N}\begin{pmatrix} \mathbf{CV}_1^T \\ \mathbf{0} \end{pmatrix} = \mathcal{N}(\mathbf{CV}_1^T) = \\ &= \mathcal{N}(\mathbf{V}_1^T) = \mathcal{S}(\mathbf{V}_1)^\perp = \mathcal{S}(\mathbf{V}_2), \end{aligned}$$

a proto také $\mathcal{S}(\mathbf{A}^T) = \mathcal{N}(\mathbf{A})^\perp = \mathcal{S}(\mathbf{V}_2)^\perp = \mathcal{S}(\mathbf{V}_1)$. Tím jsme dokončili důkaz následující *věty o URV-faktorizaci.*

Věta 9.25 *Je-li $\mathbf{A} \in \mathcal{R}^{m\times n}$ reálná matice hodnosti r, pak existují ortogonální matice $\mathbf{U}$ řádu m, ortogonální matice $\mathbf{V}$ řádu n a regulární matice $\mathbf{C}$ řádu r, pro které platí*

$$\mathbf{A} = \mathbf{URV}^T = \mathbf{U} \begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T,$$

a dále

- *prvních r sloupců matice $\mathbf{U}$ tvoří ortonormální bázi $\mathcal{S}(\mathbf{A})$,*
- *posledních $m - r$ sloupců matice $\mathbf{U}$ tvoří ortonormální bázi $\mathcal{N}(\mathbf{A}^T)$,*
- *prvních r sloupců matice $\mathbf{V}$ tvoří ortonormální bázi $\mathcal{S}(\mathbf{A}^T)$,*
- *posledních $n - r$ sloupců matice $\mathbf{V}$ tvoří ortonormální bázi $\mathcal{N}(\mathbf{A})$.*

Jestliže naopak jsou $\mathbf{U}$ *ortogonální matice řádu* m, $\mathbf{V}$ *ortogonální matice řádu* n *a* $\mathbf{C}$ *regulární matice řádu* r, *pro které platí*

$$\mathbf{A} = \mathbf{URV}^T = \mathbf{U}\begin{pmatrix} \mathbf{C}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathbf{V}^T,$$

pak matice $\mathbf{U}$ *a* $\mathbf{V}$ *mají uvedené čtyři vlastnosti.*

Později si ukážeme silnější verzi URV-faktorizace, které se říká SDV-rozklad a jeho geometrický význam.

Ortogonální projekce

Je-li $\mathcal{M}$ podprostor vektorového prostoru $\mathcal{V}$ se skalárním součinem, platí podle Tvrzení 9.23 a Tvrzení 6.22

$$\begin{aligned} \dim(\mathcal{M}+\mathcal{M}^\perp) &= \dim\mathcal{M} + \dim\mathcal{M}^\perp - \dim\mathcal{M}\cap\mathcal{M}^\perp = \\ &= \dim\mathcal{M} + \dim\mathcal{M}^\perp = \dim\mathcal{V}, \end{aligned}$$

proto $\mathcal{M}+\mathcal{M}^\perp = \mathcal{V}$. Pro každý vektor $\mathbf{x}\in\mathcal{V}$ tak existuje vyjádření $\mathbf{x} = \mathbf{m}+\mathbf{n}$, kde $\mathbf{m}\in\mathcal{M}$ a $\mathbf{n}\in\mathcal{M}^\perp$. Toto vyjádření je určené jednoznačně, neboť je-li $\mathbf{x} = \mathbf{m}'+\mathbf{n}'$ jiné vyjádření splňující podmínky $\mathbf{m}'\in\mathcal{M}$ a $\mathbf{n}'\in\mathcal{M}^\perp$, pak odečtením obou vyjádření vektoru $\mathbf{x}$ dostáváme rovnost $\mathbf{0} = (\mathbf{m}-\mathbf{m}') + (\mathbf{n}-\mathbf{n}')$, tj. $\mathbf{m}-\mathbf{m}' = -(\mathbf{n}-\mathbf{n}')$. Protože $\mathbf{m}-\mathbf{m}'\in\mathcal{M}$ a $-(\mathbf{n}-\mathbf{n}')\in\mathcal{M}^\perp$, dostáváme z rovnosti $\mathcal{M}\cap\mathcal{M}^\perp = \{\mathbf{0}\}$, že $\mathbf{m}-\mathbf{m}' = -(\mathbf{n}-\mathbf{n}') = \mathbf{0}$, tj. $\mathbf{m} = \mathbf{m}'$ a $\mathbf{n} = \mathbf{n}'$.

Definice 9.26 *Vektor* $\mathbf{m}\in\mathcal{M}$ *nazýváme* ortogonální projekce *vektoru* $\mathbf{x}$ *na podprostor* $\mathcal{M}$ *a označujeme jej* $\mathbf{P}_\mathcal{M}(\mathbf{x})$. *Zobrazení* $\mathbf{P}_\mathcal{M}:\mathcal{V}\to\mathcal{M}$ *přiřazující každému vektoru* $\mathbf{x}\in\mathcal{V}$ *jeho ortogonální projekci* $\mathbf{P}_\mathcal{M}(\mathbf{x})$ *na podprostor* $\mathcal{M}$ *nazýváme* ortogonální projekce *prostoru* $\mathcal{V}$ *na podprostor* $\mathcal{M}$.

Lemma 9.27 *Zobrazení* $\mathbf{P}_\mathcal{M}:\mathcal{V}\to\mathcal{M}$ *je lineární.*

Důkaz. Jsou-li $\mathbf{x},\mathbf{y}\in\mathcal{V}$ libovolné vektory, $\mathbf{P}_\mathcal{M}(\mathbf{x}) = \mathbf{m}\in\mathcal{M}$, $\mathbf{P}_\mathcal{M}(\mathbf{y}) = \mathbf{p}\in\mathcal{M}$, pak existují vektory $\mathbf{n},\mathbf{q}\in\mathcal{M}^\perp$ takové, že $\mathbf{x} = \mathbf{m}+\mathbf{n}$ a $\mathbf{y} = \mathbf{p}+\mathbf{q}$. Potom $\mathbf{x}+\mathbf{y} = (\mathbf{m}+\mathbf{p}) + (\mathbf{n}+\mathbf{q})$, a protože $\mathbf{m}+\mathbf{p}\in\mathcal{M}$ a $\mathbf{n}+\mathbf{q}\in\mathcal{M}^\perp$, plyne odtud $\mathbf{P}_\mathcal{M}(\mathbf{x}+\mathbf{y}) = \mathbf{m}+\mathbf{p} = \mathbf{P}_\mathcal{M}(\mathbf{x}) + \mathbf{P}_\mathcal{M}(\mathbf{y})$.

Podobně z rovnosti $\mathbf{x} = \mathbf{m}+\mathbf{n}$ plyne rovnost $a\mathbf{x} = a\mathbf{m}+a\mathbf{n}$ pro každý skalár a. Protože $a\mathbf{m}\in\mathcal{M}$ a $a\mathbf{p}\in\mathcal{M}^\perp$, plyne z definice ortogonální projekce $\mathbf{P}_\mathcal{M}$, že $\mathbf{P}_\mathcal{M}(a\mathbf{x}) = a\mathbf{m} = a\mathbf{P}_\mathcal{M}(\mathbf{x})$. $\square$

Ortogonální projekci vektoru $\mathbf{x} \in \mathcal{V}$ na podprostor $\mathcal{M} \subseteq \mathcal{V}$ najdeme tak, že napřed sestrojíme ortonormální bázi $\mathbf{u}_1, \ldots, \mathbf{u}_k$ podprostoru $\mathcal{M}$, ortonormální bázi $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ ortogonálního doplňku $\mathcal{M}^\perp$, a poté najdeme souřadnice vektoru $\mathbf{x}$ vzhledem k ortonormální bázi $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ celého prostoru $\mathcal{V}$. Tyto souřadnice dostaneme z Fourierova vyjádření

$$\mathbf{x} = \sum_{i=1}^{n} < \mathbf{u}_i | \mathbf{x} > \mathbf{u}_i$$

vektoru $\mathbf{x}$ vzhledem k bázi $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$. Platí proto

$$\mathbf{P}_{\mathcal{M}}(\mathbf{x}) = \sum_{i=1}^{k} < \mathbf{u}_i | \mathbf{x} > \mathbf{u}_i .$$

Cvičení 9.7 *Najděte ortogonální projekce několika vektorů na některé podprostory prostoru $\mathcal{R}^{n\times 1}$ a $\mathcal{C}^{n\times 1}$.*

Ortogonální projekce $\mathbf{P}_{\mathcal{M}}(\mathbf{x})$ vektoru $\mathbf{x}$ na podprostor $\mathcal{M}$ má následující speciální vlastnost.

Tvrzení 9.28 *Je-li $\mathcal{M}$ podprostor vektorového prostoru $\mathcal{V}$ se skalárním součinem a $\mathbf{x} \in \mathcal{V}$, pak pro každý vektor $\mathbf{b} \in \mathcal{M}$ různý od ortogonální projekce $\mathbf{P}_{\mathcal{M}}(\mathbf{x})$ platí*

$$\|\mathbf{x} - \mathbf{P}_{\mathcal{M}}(\mathbf{x})\| < \|\mathbf{x} - \mathbf{b}\|.$$

Důkaz. Vektor $\mathbf{P}_{\mathcal{M}}(\mathbf{x}) - \mathbf{b}$ leží v podprostoru $\mathcal{M}$, a platí proto, že $(\mathbf{x} - \mathbf{P}_{\mathcal{M}}(\mathbf{x})) \perp (\mathbf{P}_{\mathcal{M}}(\mathbf{x}) - \mathbf{b})$. Podle Pythagorovy věty, tj. Tvrzení 9.16, platí

$$\|\mathbf{x} - \mathbf{b}\|^2 = \|\mathbf{x} - \mathbf{P}_{\mathcal{M}}(\mathbf{x})\|^2 + \|\mathbf{P}_{\mathcal{M}}(\mathbf{x}) - \mathbf{b}\|^2 < \|\mathbf{x} - \mathbf{P}_{\mathcal{M}}(\mathbf{x})\|^2,$$

protože předpokládáme $\mathbf{P}_{\mathcal{M}}(\mathbf{x}) \neq \mathbf{b}$. $\square$

Definice 9.29 *Reálné číslo $\|\mathbf{x} - \mathbf{P}_{\mathcal{M}}(\mathbf{x})\|$ nazýváme* vzdálenost vektoru $\mathbf{x}$ od podprostoru $\mathcal{M}$.

Cvičení 9.8 *Spočtěte vzdálenosti několika vektorů od některých podprostorů prostoru $\mathcal{R}^{n\times 1}$ a $\mathcal{C}^{n\times 1}$.*

Elementární ortogonální projektory a reflektory

Nyní se budeme zabývat speciálním případem, kdy $\mathcal{V} = \mathcal{C}^{n\times 1}$ je komplexní unitární prostor se standardním skalárním součinem a podprostor $\mathcal{M}$

prostoru $\mathcal{C}^{n\times 1}$ je definován jako ortogonální doplněk lineárního obalu $\mathcal{L}(\mathbf{u})$, kde vektor $\mathbf{u} \in \mathcal{C}^{n\times 1}$ je jednotkový vektor, tj. $\|\mathbf{u}\| = 1$. V tomto případě existuje jednoduchý způsob, jak spočítat ortogonální projekci libovolného vektoru $\mathbf{x} \in \mathcal{R}^{n\times 1}$ na podprostor $\mathcal{M} = \mathcal{L}(\mathbf{u})^{\perp}$.

Označíme si $\mathbf{P} = \mathbf{I}_n - \mathbf{u}\mathbf{u}^*$ a spočítáme vektory $\mathbf{P}\mathbf{x}$ a $(\mathbf{I}_n - \mathbf{P})\mathbf{x}$. Platí

$$(\mathbf{I}_n - \mathbf{P})\mathbf{x} = \mathbf{x} - (\mathbf{I}_n - \mathbf{u}\mathbf{u}^*)\mathbf{x} = \mathbf{x} - \mathbf{x} + \mathbf{u}(\mathbf{u}^*\mathbf{x}) = (\mathbf{u}^*\mathbf{x})\mathbf{u} \in \mathcal{L}(\mathbf{u}).$$

Dále

$$\mathbf{u}^*(\mathbf{P}\mathbf{x}) = \mathbf{u}^*(\mathbf{I}_n - \mathbf{u}\mathbf{u}^*)\mathbf{x} = \mathbf{u}^*\mathbf{x} - (\mathbf{u}^*\mathbf{u})(\mathbf{u}^*\mathbf{x}) = \mathbf{u}^*\mathbf{x} - \|\mathbf{u}\| \cdot \mathbf{u}^*\mathbf{x} = 0,$$

tj. vektor $\mathbf{P}\mathbf{x} \in \mathcal{L}(\mathbf{u})^{\perp} = \mathcal{M}$. Protože platí $\mathbf{x} = (\mathbf{I}_n - \mathbf{P})\mathbf{x} + \mathbf{P}\mathbf{x}$, dostáváme, že pro ortogonální projekci $\mathbf{P}_{\mathcal{M}}(\mathbf{x})$ vektoru $\mathbf{x}$ na podprostor $\mathcal{M}$ platí

$$\mathbf{P}_{\mathcal{M}}(\mathbf{x}) = \mathbf{P}\mathbf{x}.$$

Definice 9.30 *Matici* $\mathbf{P} = \mathbf{I}_n - \mathbf{u}\mathbf{u}^*$ *nazýváme* elementární ortogonální projektor *určený jednotkovým vektorem* $\mathbf{u} \in \mathcal{C}^{n\times 1}$.

Pokud $\mathbf{0} \neq \mathbf{v} \in \mathcal{C}^{n\times 1}$ je libovolný vektor, označíme $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$ jednotkový vektor stejné délky a stejného směru jako vektor $\mathbf{v}$. Protože $\mathcal{L}(\mathbf{v}) = \mathcal{L}(\mathbf{u})$, platí rovněž $\mathcal{L}(\mathbf{v})^{\perp} = \mathcal{L}(\mathbf{u})^{\perp} = \mathcal{M}$ a ortogonální projekci libovolného vektoru $\mathbf{x}$ na podprostor $\mathcal{M}$ spočítáme pomocí elementárního ortogonálního projektoru určeného vektorem $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$

$$\mathbf{P} = \mathbf{I}_n - \mathbf{u}\mathbf{u}^* = \mathbf{I}_n - \frac{\mathbf{v}\mathbf{v}^*}{\|\mathbf{v}\|^2} = \mathbf{I}_n - \frac{\mathbf{v}\mathbf{v}^*}{\mathbf{v}^*\mathbf{v}}.$$

Podobně najdeme také matici, pomocí které spočítáme vektor $\mathcal{S}_{\mathcal{M}}(\mathbf{x})$ symetrický k vektoru $\mathbf{x}$ vzhledem k podprostoru $\mathcal{M} = \mathcal{L}(\mathbf{u})^{\perp}$. Víme totiž, že pokud je $\mathbf{u}$ jednotkový vektor, pak ortogonální projekce $\mathbf{x}$ na podprostor $\mathcal{M}$ se rovná $\mathbf{P}_{\mathcal{M}}(\mathbf{x}) = \mathbf{P}\mathbf{x}$. Vektor $\mathcal{S}_{\mathcal{M}}(\mathbf{x})$ symetrický k vektoru $\mathbf{x}$ vzhledem k podprostoru $\mathcal{M}$ splňuje rovnost

$$\begin{aligned} \mathbf{S}_{\mathcal{M}}(\mathbf{x}) - \mathbf{x} &= 2(\mathbf{P}_{\mathcal{M}}(\mathbf{x}) - \mathbf{x}) = 2(\mathbf{P}\mathbf{x} - \mathbf{x}) = 2(\mathbf{I}_n - \mathbf{u}\mathbf{u}^*)\mathbf{x} - 2\mathbf{x} = \\ &= -(2\mathbf{u}\mathbf{u}^*)\mathbf{x}, \end{aligned}$$

tj.

$$\mathbf{S}_{\mathcal{M}}(\mathbf{x}) = \mathbf{x} - (2\mathbf{u}\mathbf{u}^*)\mathbf{x} = (\mathbf{I}_n - 2\mathbf{u}\mathbf{u}^*)\mathbf{x}.$$

Definice 9.31 *Je-li* $\mathbf{u} \in \mathcal{C}^{n\times 1}$ *jednotkový vektor, pak matici*

$$\mathbf{R} = \mathbf{I}_n - 2\mathbf{u}\mathbf{u}^*$$

nazýváme elementární ortogonální reflektor *určený vektorem* $\mathbf{u}$.

Podobně jako v případě ortogonální projekce na podprostor $\mathcal{M} = \mathcal{L}(\mathbf{u})^\perp$ dostaneme, že pokud $\mathbf{v} \in \mathcal{C}^{n\times 1}$ je libovolný nenulový vektor, pak pro libovolný vektor $\mathbf{x} \in \mathcal{C}^{n\times 1}$ spočítáme vektor $\mathbf{S}_\mathcal{M}(\mathbf{x})$ symetrický k vektoru $\mathbf{x}$ vzhledem k podprostoru $\mathcal{M} = \mathcal{L}(\mathbf{v})^\perp$ jako

$$\mathbf{S}_\mathcal{M}(\mathbf{x}) = (\mathbf{I}_n - 2\frac{\mathbf{u}\mathbf{u}^*}{\mathbf{u}^*\mathbf{u}})\mathbf{x}.$$

Metoda nejmenších čtverců

Máme-li soustavu lineárních rovnic $\mathbf{Ax} = \mathbf{b}$ s reálnými (komplexními) koeficienty, která nemá žádné řešení, můžeme najít "přibližné řešení" této soustavy následujícím způsobem. Z předchozí teorie víme, že soustava $\mathbf{Ax} = \mathbf{b}$ je řešitelná právě když $\mathbf{b} \in \mathcal{S}(\mathbf{A})$, tj. právě když vektor $\mathbf{A}$ leží ve sloupcovém prostoru matice $\mathbf{A}$. Soustava je tedy neřešitelná právě když $\mathbf{b} \notin \mathcal{S}(\mathbf{A})$. Ve sloupcovém prostoru $\mathcal{S}(\mathbf{A})$ matice $\mathbf{A}$ existuje podle Tvrzení 9.28 vektor, který je blíže k vektoru $\mathbf{b}$, než kterýkoliv jiný vektor tohoto sloupcového prostoru $\mathcal{S}(\mathbf{A})$. Tímto vektorem je ortogonální projekce vektoru $\mathbf{b}$ na podprostor $\mathcal{S}(\mathbf{A}) = \mathcal{M}$. Tuto ortogonální projekci vektoru $\mathbf{b}$, vektor $\mathbf{P}_\mathcal{M}(\mathbf{b})$, nazýváme *řešení soustavy* $\mathbf{Ax} = \mathbf{b}$ *metodou nejmenších čtverců.*

Pokud nejsou sloupcové vektory matice soustavy $\mathbf{A}$ lineárně nezávislé, existují různá vyjádření jednoznačně určené ortogonální projekce $\mathcal{P}_\mathcal{M}(\mathbf{b})$ vektoru pravých stran $\mathbf{b}$ na sloupcový prostor $\mathcal{S}(\mathbf{A})$ matice $\mathbf{A}$ coby lineární kombinace sloupcových vektorů matice $\mathbf{A}$.

Cvičení 9.9 *Vyřešte několik neřešitelných soustav lineárních rovnic metodou nejmenších čtverců.*