Kapitola 1

Úvod

Budeme používat stejné značení jako ve skriptech [2]; jediná výjimka bude značení jednotkové matice, kde se přidržíme spíše skript [4]. Nicméně zde základní značení a pojmy uvedeme znovu, aby tento text nebyl na [2] závislý. V této úvodní kapitole tedy připomeneme základní fakta z lineární algebry, která budeme dále potřebovat.

Symbolem $\mathbb{N}$ značíme množinu všech přirozených čísel, $\mathbb{Z}$ je množina všech celých čísel, $\mathbb{R}$ množina všech reálných čísel a nakonec $\mathbb{C}$ množina všech čísel komplexních. Komplexní číslo $z \in \mathbb{C}$ má obecně tvar $z = x + \mathrm{j}y$, kde $x, y \in \mathbb{R}$, j značí komplexní jednotku ($\mathrm{j}^2 = -1$); x se nazývá reálná, y imaginární část komplexního čísla z a značíme také

$$x = \mathfrak{Re}\, z, \qquad y = \mathfrak{Im}\, z.$$

1.1 Prostory $\mathbb{R}^n$ a $\mathbb{C}^n$

Buď $n \in \mathbb{N}$ dané přirozené číslo. Symbolem $\mathbb{R}^n$, resp. $\mathbb{C}^n$, značíme množinu všech (uspořádaných) reálných, resp. komplexních, čísel. Prvkům z $\mathbb{R}^n$ a $\mathbb{C}^n$ také říkáme *vektory*, podrobněji *algebraické vektory*; v případě prvků z $\mathbb{R}^n$ mluvíme o *reálných vektorech*, zatímco prvky z $\mathbb{C}^n$ nazýváme *vektory komplexní*. Poznamenejme, že je ovšem $\mathbb{R}^n \subset \mathbb{C}^n$. Čtenář se v základním kurzu lineární algebry pravděpodobně setkal především s reálným vektorovým prostorem $\mathbb{R}^n$. Nejzákladnější pojmy, se kterými budeme v dalším pracovat, budou pojmy charakteristického čísla a charakteristického vektoru matice. Jak uvidíme, i v případě reálných matic jsou charakteristická čísla a charakteristické vektory obecně komplexní. V této souvislosti se práci s komplexním prostorem $\mathbb{C}^n$ nemůžeme vyhnout. V dalším tedy budeme pracovat s prostorem $\mathbb{C}^n$ a na případy, kdy se lze omezit na reálný případ $\mathbb{R}^n$, upozorníme.

Na algebraické vektory se v tomto textu budeme vždy dívat jako na *sloupcové vektory*, tj. je-li $\boldsymbol{x} \in \mathbb{C}^n$ (resp. $\boldsymbol{x} \in \mathbb{R}^n$), pak

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

kde pro $i = 1, 2, \dots, n$ je $x_i \in \mathbb{C}$ (resp. $x_i \in \mathbb{R}$); x_i se nazývá *i-tá složka vektoru $\boldsymbol{x}$*. Poznamenejme, že v tom, že vektory chápeme jako sloupcové vektory, se tento text liší např. od skript [6], kde vektory z $\mathbb{R}^n$ jsou naopak uvažovány zásadně jako vektory řádkové. Někdy se také užívá značení $(\boldsymbol{x})_i = x_i$ nebo $\boldsymbol{x}(i) = x_i$. Podobně jako pro matice (jak uvidíme dále, na algebraické vektory bude ostatně užitečné se dívat jako na speciální případ matic), i pro vektory používáme operaci *transpozice*; píšeme

$$\boldsymbol{x}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^T = (x_1, x_2, \dots, x_n)$$

a podobně

$$(x_1, x_2, \dots, x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{x}.$$

Nulový vektor v $\mathbb{C}^n$ (a také v $\mathbb{R}^n$), tj. vektor, jehož všechny složky jsou nulové, značíme symbolem $\mathbf{0}$ (na rozdíl od čísla 0).

1.1.1. Předpokládáme, že čtenář je seznámen s pojmem součtu algebraických vektorů a násobení vektoru skalárem. Skalárem rozumíme reálné nebo komplexní číslo. Zdůrazněme pouze, že v případě $\mathbb{R}^n$ za skaláry bereme pouze reálná čísla, zatímco v případě $\mathbb{C}^n$ čísla komplexní. Na $\mathbb{R}^n$ se tedy díváme jako na *reálný lineární prostor* (lineární prostor nad tělesem reálných čísel $\mathbb{R}$), na $\mathbb{C}^n$ jako na *komplexní lineární prostor* (lineární prostor nad tělesem komplexních čísel $\mathbb{C}$).

Mezi nejzákladnější pojmy lineární algebry patří pojmy lineární kombinace, lineární závislosti a nezávislosti a další pojmy odtud odvozené jako např. pojmy báze a dimenze atd. Připomeňme zde příslušné definice a základní vlastnosti.

Buď $k \in \mathbb{N}$. ***Lineární kombinací vektorů*** $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{C}^n$ *(resp. vektorů* $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{R}^n$*)* ***s koeficienty*** $a_1, a_2, \dots, a_k \in \mathbb{C}$ *(resp.* $a_1, a_2, \dots, a_k \in \mathbb{R}$*) rozumíme výraz*

$$\sum_{i=1}^{k} a_i \boldsymbol{x}_i = a_1 \boldsymbol{x}_1 + a_2 \boldsymbol{x}_2 + \dots + a_k \boldsymbol{x}_k. \tag{1.1}$$

Řekneme, že ***vektor*** $\boldsymbol{x} \in \mathbb{C}^n$ *(resp.* $\boldsymbol{x} \in \mathbb{R}^n$*)* ***je lineární kombinací vektorů*** $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{C}^n$ *(resp.* $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{R}^n$*), jestliže existují koeficienty* $a_1, a_2, \dots, a_k \in \mathbb{C}$ *(resp.* $a_1, a_2, \dots, a_k \in \mathbb{R}$*) tak, že*

$$\boldsymbol{x} = \sum_{i=1}^{k} a_i \boldsymbol{x}_i. \tag{1.2}$$

Lineární kombinaci (1.2) nazýváme ***triviální lineární kombinace*** *(vektorů* $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k$*), jestliže všechny koeficienty* $a_1, a_2, \dots, a_k$ *jsou nulové. V opačném případě, tj. pokud alespoň jeden z koeficientů* $a_1, a_2, \dots, a_k$ *je nenulový, nazýváme (1.2)* ***netriviální lineární kombinace****.*

Lineárním obalem vektorů $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ *(resp.* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$*) rozumíme množinu součtů všech lineárních kombinací těchto vektorů a značíme jej* $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k]$.

Poznamenejme, že je třeba rozlišovat lineární kombinaci v rámci prostoru $\mathbb{C}^n$ a v rámci prostoru $\mathbb{R}^n$. V prvním případě totiž uvažujeme obecně komplexní koeficienty, zatímco ve druhém pouze koeficienty reálné (mohli bychom také mluvit o komplexní nebo reálné lineární kombinaci). Ze stejného důvodu je třeba rozlišovat lineární obal v rámci prostoru $\mathbb{C}^n$ nebo $\mathbb{R}^n$. Lineární obal vektorů $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ v rámci komplexního prostoru $\mathbb{C}^n$ můžeme tedy zapsat ve tvaru

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k] = \left\{ \boldsymbol{x} = \sum_{i=1}^{k} a_i \boldsymbol{x}_i \;\middle|\; a_1, a_2, \ldots, a_k \in \mathbb{C} \right\},$$

zatímco v případě reálného prostoru $\mathbb{R}^n$ je

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k] = \left\{ \boldsymbol{x} = \sum_{i=1}^{k} a_i \boldsymbol{x}_i \;\middle|\; a_1, a_2, \ldots, a_k \in \mathbb{R} \right\};$$

přitom v tomto druhém případě je třeba předpokládat, že vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ jsou reálné. Jelikož je ale $\mathbb{R}^n \subset \mathbb{C}^n$ (tj. na vektory z $\mathbb{R}^n$ se lze zároveň dívat jako na vektory z $\mathbb{C}^n$), v případě lineárního obalu reálných vektorů je většinou potřeba explicite říci, zda se jedná o lineární obal v rámci prostoru $\mathbb{R}^n$ nebo $\mathbb{C}^n$.

Výše jsme definovali triviální lineární kombinaci, ve které jsou všechny koeficienty nulové. Je potřeba rozlišovat mezi triviální lineární kombinací a tzv. *nulovou lineární kombinací*, tj. lineární kombinací, která *má nulový součet*. Triviální lineární kombinace má samozřejmě nulový součet, obecně se ale může stát, že nějaká netriviální lineární kombinace vektorů má nulový součet. Odtud vychází následující definice lineární závislosti a nezávislosti vektorů.

Buď $k \in \mathbb{N}$. *Řekneme, že vektory* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ *(resp.* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \in \mathbb{R}^n$*) jsou* ***lineárně nezávislé****, jestliže pouze triviální lineární kombinace těchto vektorů je nulová (má nulový součet), tj. jestliže platí implikace*

$$\left(\sum_{i=1}^{k} a_i \boldsymbol{x}_i = \boldsymbol{0} \right) \implies (a_1 = a_2 = \cdots = a_k = 0). \tag{1.3}$$

Jestliže vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ *nejsou lineárně nezávislé, pak řekneme, že jsou* ***lineárně závislé****.*

Jak jsme již poznamenali, v případě reálných vektorů můžeme uvažovat reálné nebo komplexní lineární kombinace. V (1.3) v definici lineární nezávislosti můžeme tedy v tomto případě uvažovat buď $a_i \in \mathbb{R}$ nebo také $a_i \in \mathbb{C}$ a mohlo by se zdát, že pojem lineární závislosti reálných vektorů závisí na tom, zda se na ně díváme jako vektory v $\mathbb{R}^n$ nebo jako na vektory z $\mathbb{C}^n$. Ve skutečnosti tomu tak není. Přenecháváme čtenáři aby si rozmyslel, že platí toto tvrzení:

Vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$ jsou lineárně závislé jakožto vektory z $\mathbb{R}^n$ ***právě když*** *jsou lineárně závislé jakožto vektory z $\mathbb{C}^n$.*

Předpokládáme, že jednoduché vlastnosti lineární závislosti a nezávislosti obsažené v následujícím tvrzení čtenář dobře zná ze základního kurzu lineární algebry.

1.1.2 Tvrzení. *Buď $k \in \mathbb{N}$, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ (resp. $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$). Potom platí:*

(1) *Je-li $k = 1$, pak $\boldsymbol{x}_1$ je lineárně závislé právě když $\boldsymbol{x}_1 = \mathbf{0}$.*

(2) *Je-li $k = 2$, pak $\boldsymbol{x}_1, \boldsymbol{x}_2$ jsou lineárně závislé právě když pro nějaké $a \in \mathbb{C}$ (resp. $a \in \mathbb{R}$) je buď $\boldsymbol{x}_1 = a\boldsymbol{x}_2$ nebo $\boldsymbol{x}_2 = a\boldsymbol{x}_1$ (tj. když jeden z těchto vektorů je násobkem druhého).*

(3) *Existuje-li $j \in \{1, 2, \ldots, k\}$ tak, že $\boldsymbol{x}_j = \mathbf{0}$, pak vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ jsou lineárně závislé (tj. mezi lineárně nezávislými vektory nemůže být nulový vektor).*

(4) *Je-li $k > 1$, pak vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ jsou lineárně závislé právě když existuje $j \in$ $\in \{1, 2, \ldots, k\}$ tak, že*

$$\boldsymbol{x}_j \in \left[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_{j+1}, \boldsymbol{x}_{j+2}, \ldots, \boldsymbol{x}_k\right]$$

(tj. právě když některý z daných vektorů je lineární kombinací ostatních).

V dalším budeme občas potřebovat bod (4) z posledního tvrzení mít formulovaný v trochu silnějším tvaru:

1.1.3 Věta. *Buď $k \in \mathbb{N}$. Vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ (resp. $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$) jsou lineárně závislé právě když je buď $\boldsymbol{x}_1 = \mathbf{0}$ nebo existuje $m \in \{2, 3, \ldots, k\}$ tak, že*

$$\boldsymbol{x}_m \in [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{m-1}]$$

(tj mezi vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ existuje alespoň jeden, který je lineární kombinací vektorů ***předchozích****).*

Rozmyslete si, že věta 1.1.3 je opravdu zesílením bodu (4) z tvrzení 1.1.2, tj. z věty 1.1.3 plyne okamžitě bod (4) tvrzení 1.1.2, nikoli obráceně. Užitečné je také si rozmyslet, že z věty 1.1.3 nebo tvrzení 1.1.2 plyne tvrzení, které lze formulovat např. takto:

Přidáme-li k lineárně závislým vektorům jakékoli další vektory, dostaneme opět vektory lineárně závislé. Ubereme-li z lineárně nezávislých vektorů některé (ne všechny), zbylé vektory budou opět lineárně nezávislé.

Připomeňme pojem lineárního podprostoru.

1.1.4 Definice. *Buď $P \subset \mathbb{C}^n$ (resp. $P \in \mathbb{R}^n$), $P \neq \emptyset$. Řekneme, že P je* ***lineární podprostor*** *(krátce podprostor) prostoru $\mathbb{C}^n$ (resp. prostoru $\mathbb{R}^n$) a píšeme $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$), je-li P uzavřený vůči sčítání vektorů a násobení vektoru skalárem, tj. jsou-li splněny podmínky*

(i) $\boldsymbol{x}, \boldsymbol{y} \in P \implies \boldsymbol{x} + \boldsymbol{y} \in P$,

(ii) $\boldsymbol{x} \in P$, $a \in \mathbb{C}$ *(resp.* $a \in \mathbb{R}$*)* $\implies a\boldsymbol{x} \in P$.

Snadno je vidět, že je $\mathbb{C}^n \subset\subset \mathbb{C}^n$ a také $\mathbb{R}^n \subset\subset \mathbb{R}^n$, tj. $\mathbb{C}^n$ a $\mathbb{R}^n$ jsou svými podprostory. Je-li $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$) a přitom $P \neq \mathbb{C}^n$ (resp. $P \neq \mathbb{R}^n$), pak mluvíme o tzv. *vlastním podprostoru*. Extrémní případ vlastního podprostoru je tzv. *triviální podprostor*, tj. podprostor, který obsahuje pouze nulový prvek (rozmyslete si, že každý podprostor obsahuje vždy nulový prvek). Zdůrazněme, že pojem podprostoru závisí na tom, zda se pohybujeme v rámci $\mathbb{C}^n$ nebo pouze v rámci reálného prostoru $\mathbb{R}^n$ — v bodu (ii) definice se totiž podle toho uvažuje buď obecně $a \in \mathbb{C}$ nebo pouze $a \in \mathbb{R}$. Můžete si rozmyslet, že ačkoli je $\mathbb{R}^n \subset \mathbb{C}^n$ [a $\mathbb{R}^n$ je (reálný) lineární prostor!], $\mathbb{R}^n$ *není* podprostor prostoru $\mathbb{C}^n$.

Jsou-li P, Q dva podprostory v $\mathbb{C}^n$ (resp. v $\mathbb{R}^n$) a přitom $P \subset Q$, pak často píšeme $P \subset\subset Q$ nebo podrobněji

$$P \subset\subset Q \subset\subset \mathbb{C}^n \qquad (\text{resp. } P \subset\subset Q \subset\subset \mathbb{R}^n).$$

Přímo z definice lineárního podprostoru je vidět, že podprostor je uzavřený vzhledem k lineárním kombinacím, tj. je-li $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$), $k \in \mathbb{N}$, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in P$, $a_1, a_2, \ldots, a_k \in \mathbb{C}$ (resp. $a_1, a_2, \ldots, a_k \in \mathbb{R}$), pak je vždy

$$\sum_{i=1}^{k} a_i \boldsymbol{x}_i \in P.$$

Důležitá vlastnost lineárního obalu vektorů je ta, že to vždy je lineární podprostor. Přesněji platí:

Buď $P \subset\subset \mathbb{C}^n$ *(resp.* $P \subset\subset \mathbb{R}^n$*),* $k \in \mathbb{N}$, $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in P$. *Potom je*

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k] \subset\subset P.$$

Zdůrazněme, že pokud $P \subset\subset \mathbb{C}^n$, pak zde lineární obal vektorů $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ uvažujeme v rámci prostoru $\mathbb{C}^n$, zatímco v případě $P \subset\subset \mathbb{R}^n$ uvažujeme daný lineární obal v rámci reálného prostoru $\mathbb{R}^n$. Tuto konvenci budeme dodržovat i dále.

Je-li P podprostor v $\mathbb{C}^n$ nebo v $\mathbb{R}^n$ a jestliže pro nějaké vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ je

$$P = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k],$$

pak řekneme, že vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ *generují* P (rozmyslete si, že je potom nutně $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in P$).

1.1.5 Definice. *Buď* $P \subset\subset \mathbb{C}^n$ *(resp.* $P \subset\subset \mathbb{R}^n$*). Řekneme, že vektory* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ *(resp.* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$*) tvoří **bázi*** P*, jestliže*

(i) $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ *jsou lineárně nezávislé,*

(ii) $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ *generují* P.

Jelikož v triviálním podprostoru $\{\mathbf{0}\}$ neexistují lineárně nezávislé vektory, triviální podprostor nemá bázi. Důležité je, že každý netriviální podprostor v $\mathbb{C}^n$ nebo $\mathbb{R}^n$, už bázi má. Platí následující tvrzení.

1.1.6 Věta. *Buď $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$), $P \neq \{\mathbf{0}\}$. Potom P má bázi a každé dvě báze P mají stejný počet prvků.*

Skutečnost, že různé báze téhož (netriviálního) podprostoru mají stejný počet prvků je nejenom zajímavá, ale i důležitá — umožňuje následující definici dimenze podprostoru.

1.1.7 Definice. *Buď P netriviální podprostor v $\mathbb{C}^n$ nebo $\mathbb{R}^n$. Potom **dimenzi** P definujeme jako počet prvků báze P a značíme ji $\dim P$. Dále definujeme $\dim\{\mathbf{0}\} = 0$.*

V následujícím tvrzení jsou shrnuty základní vlastnosti podprostorů, které souvisí s pojmy báze a dimenze.

1.1.8 Věta. *Buď $Q \subset\subset \mathbb{C}^n$ (resp. $Q \subset\subset \mathbb{R}^n$), $P \subset\subset Q$, $\dim Q = m$, $\dim P = k > 0$. Potom platí:*

(a) *k je rovno největšímu počtu lineárně nezávislých vektorů z P.*

(b) *Je-li $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r] = P$ $(r \in \mathbb{N})$, pak $r \geq k$ a existují navzájem různé indexy $i_1, i_2, \ldots, i_r$ $(1 \leq i_j \leq r)$ takové, že vektory $\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \ldots, \boldsymbol{x}_{i_k}$ tvoří bázi P (tj. **z každých vektorů, které generují** P**, lze vybrat bázi** P).*

(c) *Jsou-li vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r \in P$ lineárně nezávislé $(r \in \mathbb{N})$, pak $r \leq k$; je-li $r < k$, pak existují $\boldsymbol{x}_{r+1}, \boldsymbol{x}_{r+2}, \ldots, \boldsymbol{x}_k \in P$ tak, že*

$$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r, \boldsymbol{x}_{r+1}, \ldots, \boldsymbol{x}_k$$

*tvoří bázi P (tj. **lineárně nezávislé vektory lze doplnit na bázi**).*

(d) *Vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in P$ tvoří bázi P právě když jsou lineárně nezávislé.*

(e) *Vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in P$ tvoří bázi P právě když generují P.*

(f) *Je $k \leq m$; přitom $k = m$ právě když $P = Q$.*

1.1.9 Poznámka. Bázi podprostoru jsme definovali dvěma podmínkami — báze je tvořena lineárně nezávislými vektory, které generují daný podprostor P. Všimněme si, že pokud známe dimenzi P a uvažujeme vektory, jejichž počet je roven $\dim P$, pak k ověření toho, že tyto vektory tvoří bázi P, podle bodů (d) a (e) stačí ověřit pouze jednu z definičních podmínek.

1.1.10 Poznámka. Připomeňme pojem tzv. *standardní báze* $\mathbb{R}^n$. Pro $i = 1, 2, \ldots, n$ nechť $\boldsymbol{e}_i$ značí vektor z $\mathbb{R}^n$, jehož i-tá složka je rovna 1 a ostatní jsou nulové, tj.

$$\boldsymbol{e}_1 = (1, 0, \ldots, 0)^T, \quad , \boldsymbol{e}_2 = (0, 1, 0, \ldots, 0)^T, \ldots, \boldsymbol{e}_n = (0, 0, \ldots, 0, 1)^T.$$

Rozmyslete si, že tyto vektory jsou jistě lineárně nezávislé [např. podle bodu (4) tvrzení 1.1.2]. Dále je zřejmé, že je-li $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$, libovolný vektor, pak je

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i, \tag{1.4}$$

tj. lineární obal vektorů $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ je roven $\mathbb{R}^n$ (když uvažujeme pouze reálné lineární kombinace). Vidíme tedy, že $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ tvoří bázi $\mathbb{R}^n$ a

$$\dim \mathbb{R}^n = n.$$

Bázi $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ říkáme *standardní báze.*

Rozmyslete si ještě, že pokud budeme uvažovat obecně komplexní lineární kombinace, pak lineární obal vektorů $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ bude roven $\mathbb{C}^n$. Rovnost (1.4) totiž platí i pro vektor $\boldsymbol{x} \in \mathbb{C}^n$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$; rozdíl je pouze v tom, že x_i jsou potom komplexní. Vektory $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ tvoří tedy zároveň bázi $\mathbb{C}^n$ a je $\dim \mathbb{C}^n = n$. Zajímavé je uvědomit si, že je $\mathbb{R}^n \subset \mathbb{C}^n$, $\mathbb{R}^n \neq \mathbb{C}^n$ a přitom $\dim \mathbb{R}^n = \dim \mathbb{C}^n$. Na první pohled by se mohlo zdát, že tato skutečnost je ve sporu s bodem (f) věty 1.1.8. Ve skutečnosti se ale o spor nejedná, neboť, jak jsme si již dříve všimli, $\mathbb{R}^n$ *není* podprostorem prostoru $\mathbb{C}^n$.

1.1.11 Poznámka. Buď $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$) netriviální podprostor a nechť $\mathcal{B} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ je nějaká báze P. Jelikož vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ generují P, každý vektor $\boldsymbol{y} \in P$ lze napsat ve tvaru

$$\boldsymbol{y} = \sum_{i=1}^{k} y_i \boldsymbol{x}_i, \tag{1.5}$$

kde $y_i \in \mathbb{C}$ (resp. $y_i \in \mathbb{R}$) jsou vhodné koeficienty. Jelikož vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ jsou lineárně nezávislé, jsou koeficienty $y_1, y_2, \ldots, y_k$ určeny jednoznačně (zkuste si rozmyslet jednoduchý důkaz). Těmto jednoznačně určeným koeficientům $y_1, y_2, \ldots, y_k$ se říká *souřadnice vektoru* $\boldsymbol{y}$ *v bázi* $\mathcal{B}$. Vektoru $(y_1, y_2, \ldots, y_k)^T \in \mathbb{C}^k$ [resp. $(y_1, y_2, \ldots, y_k)^T \in \mathbb{R}^k$] se říká *vektor souřadnic* (souřadnicový vektor) vektoru $\boldsymbol{y}$ v bázi $\mathcal{B}$ a někdy se používá značení

$$(\boldsymbol{y})_{\mathcal{B}} = (y_1, y_2, \ldots, y_k)^T.$$

Jeden vektor $\boldsymbol{y} \in P$ má v různých bázích různé souřadnicové vektory.

Všimněte si, že podle rovnosti (1.4) v případě $P = \mathbb{C}^n$ (resp. $P = \mathbb{R}^n$) a standardní báze jsou souřadnice vektoru jednoduše rovny složkám tohoto vektoru.

1.2 Algebra matic

Předpokládáme, že čtenář absolvoval základní kurz lineární algebry, takže jistě se s pojmem matice setkal. Zaveďme zde ale označení, které budeme dále používat a připomeňme některé pojmy a tvrzení týkající matic reálných nebo komplexních čísel.

Jsou-li m, n dvě přirozená čísla, pak maticí *typu* $m \times n$ [nebo typu (m, n)] rozumíme matici, která má m řádků a n sloupců, tj. matici

$$\mathbf{A} = \big(a_{ij}\big)_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

Krátce také píšeme $\mathbf{A} = (a_{ij})$, pokud je zřejmé, jakého typu matice $\mathbf{A}$ je. Číslu a_{ij} říkáme *i-j-tý prvek* matice $\mathbf{A}$, pro který budeme někdy také používat zápis $[\mathbf{A}]_{ij}$ nebo $\mathbf{A}(i, j)$, tj.

$$[\mathbf{A}]_{ij} = \mathbf{A}(i, j) = a_{ij}.$$

Poznamenejme, že prvky matice nemusí být jenom čísla, ale i jiné matematické objekty, např. funkce. Pokud všechny prvky a_{ij} jsou čísla, mluvíme o *číselné matici*; tomuto případu se zde budeme převážně věnovat. Jsou-li všechny prvky matice reálné, mluvíme o reálné matici, komplexní matice má obecně prvky komplexní. Množinu všech komplexních matic typu $m \times n$ značíme symbolem

$$\mathbb{C}^{m\times n},$$

množinu všech reálných matic typu $m \times n$ analogicky symbolem

$$\mathbb{R}^{m\times n}.$$

Je ovšem $\mathbb{R}^{m\times n} \subset \mathbb{C}^{m\times n}$; v dalším budeme většinou mluvit obecně o komplexních maticích a na reálné matice se budeme jednoduše dívat jako na speciální případ matic komplexních.

V dalším budeme často pracovat s jednotlivými sloupci nebo řádky matice a budeme přitom používat následující značení. Je-li $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, pak pro $i = 1, 2, \ldots, m$ značíme

$$\mathbf{A}(i, :) = \big(a_{ij}\big)_{j=1}^{n} = \big(a_{i1}, a_{i2}, \ldots, a_{in}\big),$$

tj. $\mathbf{A}(i, :)$ je i-tý řádek matice $\mathbf{A}$. Pro $j = 1, 2, \ldots, n$ značíme

$$\mathbf{A}(:, j) = \big(a_{ij}\big)_{i=1}^{m} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

j-tý sloupec matice $\mathbf{A}$. Na řádky $\mathbf{A}(i, :)$ se můžeme dívat jako na speciální případy matic o jednom řádku, tj. pro $\mathbf{A} \in \mathbb{C}^{m\times n}$ je $\mathbf{A}(i, :) \in \mathbb{C}^{1\times n}$ [podobně, je-li $\mathbf{A} \in \mathbb{R}^{m\times n}$, pak $\mathbf{A}(i, :) \in \mathbb{R}^{1\times n}$]. Analogicky se na jednotlivé sloupce matice můžeme dívat jako na matice o jednom sloupci, tj. pro $\mathbf{A} \in \mathbb{C}^{m\times n}$ je $\mathbf{A}(:, j) \in \mathbb{C}^{m\times 1}$. Na tyto sloupce se ovšem můžeme zároveň dívat jako na algebraické vektory z $\mathbb{C}^m$, tj. můžeme ztotožnit $\mathbb{C}^{m\times 1}$ a $\mathbb{C}^m$, tj můžeme psát

$$\mathbf{A}(:, j) \in \mathbb{C}^m$$

(podobně v reálném případě). Podobně bychom se mohli na jednotlivé řádky matice dívat jako na algebraické vektory. Při naší konvenci, podle které algebraické vektory chápeme zásadně jako vektory sloupcové, nemůžeme ale přímo ztotožnit $\mathbb{C}^{1\times n}$ a $\mathbb{C}^n$. Přesněji pro $\mathbf{A} \in \mathbb{C}^{m\times n}$ bude

$$\big(\mathbf{A}(i,:)\big)^T = \mathbf{A}(i,:)^T \in \mathbb{C}^n,$$

$i = 1, 2, \ldots, m$ (podobně v reálném případě). V dalším bude často výhodné se naopak na vektory z $\mathbb{C}^n$ (nebo z $\mathbb{R}^n$) dívat jako na matice o jednom sloupci.

1.2.1 Některé speciální typy matic. Až na některé výjimky, budeme v celém textu pracovat se *čtvercovými maticemi*, tj. s maticemi, které mají stejný počet řádků a sloupců. Je-li tento počet roven n, jedná se tedy o matici typu $n \times n$; mluvíme potom o (čtvercové) matici *řádu* n. Podle předchozího značení se jedná o matice z $\mathbb{C}^{n\times n}$ (resp. $\mathbb{R}^{n\times n}$).

Důležitými případy čtvercových matic jsou matice trojúhelníkové a diagonální. Buď $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{n\times n}$. Řekneme, že $\mathbf{A}$ je *horní trojúhelníková matice*, jestliže pro $i, j = 1, 2, \ldots, n$ je

$$a_{ij} = 0 \qquad \text{kdykoli } i > j.$$

Řekneme, že $\mathbf{A}$ je *dolní trojúhelníková matice*, je-li

$$a_{ij} = 0 \qquad \text{kdykoli } i < j.$$

Názorně je $\mathbf{A}$ horní trojúhelníková (resp. dolní trojúhelníková) matice pokud všechny její prvky pod diagonálou (resp. nad diagonálou) jsou nulové. Přitom *diagonálou* (podrobněji *hlavní diagonálou*) čtvercové matice rozumíme tu část matice, která se skládá z tzv. *diagonálních prvků*, tj. prvků a_{ii} ($i = 1, 2, \ldots, n$). (Někdy se termíny diagonální prvek a diagonála používají i v případě matic, které nejsou čtvercové.) Horní trojúhelníková matice je tedy matice tvaru

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1\,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2\,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \ldots & a_{3\,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & 0 & \ldots & 0 & a_{nn} \end{pmatrix},$$

zatímco dolní trojúhelníková matice má tvar

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ a_{31} & a_{32} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix}.$$

Čtvercová matice je *diagonální*, jestliže všechny její prvky mimo (hlavní) diagonálu jsou nulové; je zřejmé, že čtvercová matice je diagonální právě když je zároveň horní i dolní trojúhelníková. Pro diagonální matici $\mathbf{A}$ řádu n, která má na diagonále prvky $\alpha_1, \alpha_2, \ldots, \alpha_n$

(v daném pořadí), budeme často používat zkrácený zápis $\mathbf{A} = \mathrm{diag}[\alpha_1, \alpha_2, \dots, \alpha_n]$, tj.

$$\mathbf{A} = \mathrm{diag}[\alpha_1, \alpha_2, \dots, \alpha_n] = \begin{pmatrix} \alpha_1 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & 0 & \dots & 0 \\ 0 & 0 & \alpha_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_n \end{pmatrix}.$$

Diagonální matici, která má na diagonále samé jedničky značíme $\mathbf{E}$ a nazýváme ji *jednotková matice*, tj.

$$\mathbf{E} = \mathrm{diag}[1, 1, \dots, 1] = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Všimněme si, že z posledního zápisu není zřejmé, jakého řádu je matice $\mathbf{E}$. Symbol $\mathbf{E}$ se skutečně používá pro jednotkové matice různých řádů, přičemž řád této matice je většinou zřejmý z kontextu (v některých učebnicích se pro přesnost pro jednotkovou matici řádu n používá např. zápis $\mathbf{E}_n$ a pod.).

1.2.2. V předchozím jsme se setkali s pojmem transpozice vektoru, kde šlo ovšem pouze o převedení vektoru se sloupcového na řádkový zápis nebo obráceně. Operaci transpozice definujeme obecněji pro matice. Je-li $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, pak matici $\mathbf{B} = (b_{ij}) \in \mathbb{C}^{n\times m}$, kde

$$b_{ij} = a_{ji} \qquad \text{pro } i = 1, 2, \dots, n,\ j = 1, 2, \dots, m,$$

nazýváme *matice transponovaná* k $\mathbf{A}$. Matici transponovanou k $\mathbf{A}$ značíme $\mathbf{A}^T$. V případě čtvercové matice si můžeme transpozici názorně představit jako „překlopení kolem diagonály“. Při našem ztotožnění algebraických vektorů s maticemi o jednom sloupci, transpozice takovýchto matic je ovšem přesně totéž jako transpozice vektoru.

Řekneme, že matice $\mathbf{A}$ je *symetrická*, jestliže

$$\mathbf{A}^T = \mathbf{A}.$$

Jelikož pro $\mathbf{A} \in \mathbb{C}^{m\times n}$ je $\mathbf{A}^T \in \mathbb{C}^{n\times m}$, je zřejmé, že pojem symetrie má smysl pouze pro čtvercové matice.

V případě komplexních matic se častěji než s transponovanými maticemi pracuje s tzv. maticemi konjugovanými. Připomeňme, že je-li $z = x + \mathrm{j}y \in \mathbb{C}$, kde $x, y \in \mathbb{R}$ (j značí imaginární jednotku), pak číslo komplexně sdružené k z je definováno jako $\overline{z} = x - \mathrm{j}y$; pruh zde bude vždy značit komplexně sdruženou hodnotu. Analogické značení budeme používat i pro matice a vektory. Je-li $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, pak tedy

$$\overline{\mathbf{A}} = (\overline{a}_{ij}) \qquad (\in \mathbb{C}^{m\times n})$$

a podobně pro vektory z $\mathbb{C}^n$.

Pro komplexní matici $\mathbf{A} \in \mathbb{C}^{m\times n}$ dále značíme

$$\mathbf{A}^* = \overline{\mathbf{A}^T} = (\overline{\mathbf{A}})^T.$$

Matici $\mathbf{A}^*$ nazýváme *matice konjugovaná* (nebo *hermitovsky adjungovaná*) k matici $\mathbf{A}$; někdy se místo $\mathbf{A}^*$ používá značení $\mathbf{A}^H$. tento pojem má samozřejmě také smysl pro algebraické vektory — je-li $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$, pak je

$$\boldsymbol{x}^* = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n).$$

Matici $\mathbf{A} \in \mathbb{C}^{n\times n}$ nazveme *hermitovská* (nebo *Hermitova*), je-li

$$\mathbf{A}^* = \mathbf{A}.$$

V případě komplexních matic se častěji než se symetrickými maticemi pracuje s maticemi hermitovskými. Můžeme si ovšem všimnout, že pro $\mathbf{A}$ reálné je $\overline{\mathbf{A}} = \mathbf{A}$ (je dokonce pravda, že $\mathbf{A}$ je reálná právě když $\overline{\mathbf{A}} = \mathbf{A}$). Pro reálnou matici $\mathbf{A}$ je tedy jednoduše $\mathbf{A}^* = \mathbf{A}^T$ a reálná matice je hermitovská právě když je symetrická; v reálném případě se proto většinou místo o hermitovských maticích mluví o maticích symetrických.

1.2.3. Mluví-li se o algebře matic, pak se tím myslí, že kromě operací transpozice a konjunkce jsou pro matice definovány operace násobení matice skalárem, součet matic a násobení matice maticí. *Násobení matice* konstantou a *sčítání matic* je definováno podobně analogickým operacím na algebraických vektorech, tj. „po složkách“. Je-li $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, $\alpha \in \mathbb{C}$, pak je $\alpha\mathbf{A} \in \mathbb{C}^{m\times n}$,

$$\alpha\mathbf{A} = (\alpha a_{ij});$$

v případě reálných matic bývá většinou zvykem uvažovat pouze reálné skaláry. Součet matic je definován pro matice stejného typu. Pro $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$, $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$, je $\mathbf{A} + \mathbf{B} \in \mathbb{C}^{m\times n}$,

$$\mathbf{A} + \mathbf{B} = \big(a_{ij} + b_{ij}\big).$$

Velice důležitou operací pro nás v dalším bude především operace násobení matic. Jak uvidíme, při násobení dvou matic záleží na jejich pořadí. Aby byl vůbec definován součin dvou matic, je nutné, aby počet *sloupců* matice „vlevo“ byl stejný jako počet *řádků* matice „vpravo“. Součin matic je definován následujícím způsobem:

> *Buď* $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, $\mathbf{B} = (b_{ij}) \in \mathbb{C}^{n\times k}$ *(*$m, n, k \in \mathbb{N}$*). Potom součinem matic* $\mathbf{A}, \mathbf{B}$ *rozumíme matici* $\mathbf{A} \cdot \mathbf{B} = \mathbf{AB} = \mathbf{C} = (c_{ij}) \in \mathbb{C}^{m\times k}$*, kde*
>
> $$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj} \tag{1.6}$$
>
> *(*$i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, k$*).*

Můžeme si všimnout, že ve vyjádření (1.6) pro i-j-tý prvek součinu $\mathbf{AB}$ se vyskytují pouze prvky i-tého řádku matice $\mathbf{A}$ a prvky j-tého sloupce matice $\mathbf{B}$. Jak ale víme, na i-tý řádek matice $\mathbf{A}$ se můžeme dívat jako na matici typu $1 \times n$ a na j-tý sloupec matice $\mathbf{B}$ jako na matici typu $n \times 1$. Podle právě uvedené definice součinu matic je součin *matic* $\mathbf{A}(i,:)$ a $\mathbf{B}(:,j)$ definován a je to matice typu 1×1, na kterou se ovšem můžeme dívat jednoduše jako na číslo (komplexní nebo reálné). Vidíme, že rovnost (1.6) můžeme psát ve tvaru

$$c_{ij} = \mathbf{A}(i,:)\mathbf{B}(:,j). \tag{1.7}$$

Poznamenejme, že v případě reálných matic je součin napravo v (1.7) roven skalárnímu součinu *vektorů* $(\mathbf{A}(i,:))^T$ a $\mathbf{B}(:,j)$ (nikoli v komplexním případě); pojem skalárního součinu připomeneme podrobněji v odstavci 1.5.

Jelikož na vektory z $\mathbb{C}^n$ se můžeme dívat jako na jednosloupcové matice (viz začátek odstavce 1.2), pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ má smysl součin $\boldsymbol{x}^T\boldsymbol{y}$ jakožto součin matic. Je-li $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T$, $\boldsymbol{y} = (y_1, y_2, \dots, y_n)^T$, pak je

$$\boldsymbol{x}^T\boldsymbol{y} = \sum_{i=1}^{n} x_i y_i;$$

je ovšem $\boldsymbol{x}^T\boldsymbol{y} \in \mathbb{C}$ (opět při konvenci, kdy ztotožňujeme čísla a matice typu 1×1).

Může být zajímavé všimnout si, že jsou-li $\boldsymbol{x}, \boldsymbol{y}$ opět vektory, $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$, $\boldsymbol{y} = (y_1, y_2, \dots, y_m) \in \mathbb{C}^m$ (obecně může být $m \neq n$), pak má také smysl součin $\boldsymbol{x}\boldsymbol{y}^T$. Je potom $\boldsymbol{x}\boldsymbol{y}^T \in \mathbb{C}^{n\times m}$,

$$\boldsymbol{x}\boldsymbol{y}^T = \begin{pmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_m \\ x_2y_1 & x_2y_2 & \dots & x_2y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & \dots & x_ny_m \end{pmatrix}. \tag{1.8}$$

Tomuto součinu se někdy říká *vnější součin*. Přenecháváme čtenáři aby si rozmyslel, že pokud ani jeden z vektorů $\boldsymbol{x}, \boldsymbol{y}$ není nulový, pak matice v (1.8) má hodnost rovnou 1 a naopak každou matici o hodnosti 1 je možné napsat v tomto tvaru (pro pojem hodnosti matice viz dále odstavec 1.2.9).

Důležitý speciální případ součinu matic je součin matice a vektoru. Je-li $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$, pak součin $\mathbf{A}\boldsymbol{x}$ má smysl (je definován), neboť podle předchozího se na $\boldsymbol{x}$ můžeme dívat jakona matici typu $n \times 1$. Výsledek součinu $\mathbf{A}\boldsymbol{x}$ je matice typu $m \times 1$; při naší konvenci je tedy $\mathbf{A}\boldsymbol{x}$ vektor z $\mathbb{C}^m$. Součin $\mathbf{A}\boldsymbol{x}$ můžeme psát ve tvaru

$$\mathbf{A}\boldsymbol{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}. \tag{1.9}$$

Všimněme si, že tuto rovnost lze psát ve tvaru

$$\begin{aligned} \mathbf{A}\boldsymbol{x} &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \\ &= x_1\mathbf{A}(:,1) + x_2\mathbf{A}(:,2) + \dots + x_n\mathbf{A}(:,n) = \sum_{i=1}^{n} x_j\mathbf{A}(:,j). \end{aligned} \tag{1.10}$$

Odtud především vidíme, že součin $\mathbf{A}\boldsymbol{x}$ je roven lineární kombinaci sloupců matice $\mathbf{A}$, přičemž koeficienty této lineární kombinace jsou rovny složkám vektoru $\boldsymbol{x}$.

Ponecháváme čtenáři aby si rozmyslel, že pro $\boldsymbol{y} \in \mathbb{C}^m$ má podobně smysl součin $\boldsymbol{y}^T\mathbf{A}$ a je $\boldsymbol{y}^T \in \mathbb{C}^{1\times n}$.

1.2.4. Připomeňme alespoň některé jednoduché vlastnosti operací s maticemi. Je zřejmé, že sčítání matic je asociativní a komutativní a např. také platí

$$\alpha(\mathbf{A}+\mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B},$$

kdykoli $\alpha \in \mathbb{C}$ a součet $\mathbf{A}+\mathbf{B}$ má smysl, tj. $\mathbf{A}, \mathbf{B}$ jsou matice téhož typu. Pro matice $\mathbf{A}, \mathbf{B}$ stejného typu dále např. platí

$$(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T, \qquad \overline{(\mathbf{A}+\mathbf{B})} = \overline{\mathbf{A}} + \overline{\mathbf{B}}, \qquad (\mathbf{A}+\mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*.$$

Vlastnosti součinu matic jsou trochu složitější. Důležitá skutečnost je, že *součin matic není komutativní*, tj. v součinu matic obecně záleží na pořadí matic. Pokud součin $\mathbf{AB}$ má smysl, součin $\mathbf{BA}$ nemusí být dokonce ani definován — je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times k}$, pak součin $\mathbf{AB}$ má smysl, ale $\mathbf{BA}$ bude definováno pouze pokud $k = m$. Je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times m}$, pak oba součiny $\mathbf{AB}$ a $\mathbf{BA}$ mají smysl, ale je-li $m \neq n$, o rovnosti $\mathbf{AB} = \mathbf{BA}$ nelze mluvit, neboť potom je $\mathbf{AB} \in \mathbb{C}^{m\times m}$, $\mathbf{BA} \in \mathbb{C}^{n\times n}$, tj. uvedené součiny jsou sice oba čtvercové matice, ale různých řádů.

O rovnosti $\mathbf{AB} = \mathbf{BA}$ by bylo možné mluvit v případě čtvercových matic stejného řádu. Ale ani v tomto případě uvedená rovnost obecně neplatí, tj. ani pro čtvercové matice není násobení komutativní. Jednoduchý protipříklad je třeba následující. Je-li

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 2 & -1 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix},$$

pak můžeme spočítat, že

$$\mathbf{AB} = \begin{pmatrix} -3 & -3 \\ 3 & 3 \end{pmatrix}, \qquad \mathbf{BA} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

tj. $\mathbf{AB} \neq \mathbf{BA}$. Všimněte si zde ještě, že se může snadno stát, že součin dvou nenulových matic je nulový (součin nenulových *čísel* nulový být nemůže). Podobně se může snadno stát, že mocnina nenulové matice je rovna matici nulové. Jednoduchý příklad je třeba matice $\mathbf{A} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$; je $\mathbf{A}^2 = \mathbf{A}\cdot\mathbf{A} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

Násobení matic sice není komutativní, ale je alespoň asociativní. Platí totiž následující tvrzení. Pro úplnost uvádíme jednoduchý důkaz tohoto tvrzení; důkazy většiny dalších tvrzení v tomto stručném opakování většinou vynecháváme.

1.2.5 Tvrzení. *Buďte $m, n, k, r \in \mathbb{N}$, $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times k}$, $\mathbf{C} \in \mathbb{C}^{k\times r}$. Potom*

$$\mathbf{A}\cdot(\mathbf{BC}) = (\mathbf{AB})\cdot\mathbf{C}. \tag{1.11}$$

Důkaz. Nejprve si všimneme, že součiny nalevo i napravo v (1.11) jsou matice stejného typu (a uvedené součiny mají smysl). Podle definice násobení matic je $\mathbf{BC} \in \mathbb{C}^{n\times r}$, takže součin $\mathbf{A}(\mathbf{BC})$ má smysl a je $\mathbf{A}(\mathbf{BC}) \in \mathbb{C}^{m\times r}$. Podobně je $\mathbf{AB} \in \mathbb{C}^{m\times k}$ a dále $(\mathbf{AB})\mathbf{C} \in$ $\in \mathbb{C}^{m\times r}$.

Důkaz rovnosti (1.11) spočívá pouze v podrobném rozepsání uvedených součinů podle definice a vhodném přerovnání pořadí sčítání. Nechť $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$, $\mathbf{C} = (c_{ij})$. Označme $\mathbf{BC} = \mathbf{D} = (d_{ij})$. Jak jsme již poznamenali, je $\mathbf{D} \in \mathbb{C}^{n\times r}$ a podle definice násobení matic je

$$(*) \qquad d_{ij} = \sum_{s=1}^{k} b_{is} c_{sj}$$

$(i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, r)$. Připomeňme, že podle dříve zavedeného značení symbol $\mathbf{A}(i, j) = [\mathbf{A}]_{ij}$ značí i-j-tý prvek matice $\mathbf{A}$. Podle definice násobení matic a podle $(*)$ dostáváme

$$\big[\mathbf{A}\cdot(\mathbf{BC})\big]_{ij} = [\mathbf{AD}]_{ij} = \sum_{h=1}^{n} a_{ih} d_{hj} = \sum_{h=1}^{n} a_{ih} \sum_{s=1}^{k} b_{hs} c_{sj} = \sum_{h=1}^{n}\sum_{s=1}^{k} a_{ih} b_{hs} c_{sj} =$$
$$= \sum_{s=1}^{k}\sum_{h=1}^{n} a_{ih} b_{hs} c_{sj} = \sum_{s=1}^{k}\left(\sum_{h=1}^{n} a_{ih} b_{hs}\right) c_{sj}$$

$(i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, r)$. Pokud si uvědomíme, že (opět podle definice násobení matic)

$$\sum_{h=1}^{n} a_{ih} b_{hs} = [\mathbf{AB}]_{is},$$

vidíme okamžitě, že pro daná i, j je $[\mathbf{A}(\mathbf{BC})]_{ij} = [(\mathbf{AB})\mathbf{C}]_{ij}$, tj. opravdu platí (1.11). □

Následující tři tvrzení uvádíme bez (vesměs jednoduchých) důkazů.

1.2.6 Tvrzení. *Buď* $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B}, \mathbf{C} \in \mathbb{C}^{n\times k}$ *(*$m, n, k \in \mathbb{N}$*),* $\alpha \in \mathbb{C}$. *Potom je*

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}, \qquad \alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B}).$$

1.2.7 Tvrzení. *Buď* $\mathbf{E}$ *jednotková matice řádu* n, $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times n}$. *Potom*

$$\mathbf{A}\cdot\mathbf{E} = \mathbf{A}, \qquad \mathbf{E}\cdot\mathbf{B} = \mathbf{B}.$$

Speciálně: Je-li $\mathbf{A}$ *čtvercová matice řádu* n*, pak*

$$\mathbf{A}\cdot\mathbf{E} = \mathbf{E}\cdot\mathbf{A} = \mathbf{A}.$$

Poznamenejme, že podle tohoto tvrzení jednotková matice $\mathbf{E}$ hraje při násobení matic úlohu jedničky — proto se ji také říká jednotková.

1.2.8 Tvrzení. *Buď* $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times k}$ *(*$m, n, k \in \mathbb{N}$*). Potom*

$$\overline{(\mathbf{AB})} = \overline{\mathbf{A}}\,\overline{\mathbf{B}}, \qquad (\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T, \qquad (\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*.$$

Zde si zvláště všimněme, že transpozice součinu je rovna součinu transpozic, *ale v obráceném pořadí.* Podobně pro konjunkci součinu. Indukcí bychom mohli dostat, že jsou-li $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_r$ matice takové, že součin $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_r$ má smysl, pak

$$(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_r)^T = \mathbf{A}_r^T\mathbf{A}_{r-1}^T\cdots\mathbf{A}_1^T, \qquad (\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_r)^* = \mathbf{A}_r^*\mathbf{A}_{r-1}^*\cdots\mathbf{A}_1^*.$$

1.2.9. Důležitý pojem je pojem hodnosti matice. Připomeňme, že *hodnost matice* $\mathbf{A}$ je definována jako *maximální počet lineárně nezávislých řádků* $\mathbf{A}$. Hodnost matice $\mathbf{A}$ značíme $\operatorname{hod}(\mathbf{A})$. Je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$, pak transpozice řádků $\mathbf{A}$ jsou vektory z $\mathbb{C}^n$ $[\mathbf{A}(i,:)^T \in \mathbb{C}^n]$ — v tomto smyslu můžeme mluvit o lineární závislosti či nezávislosti řádků matice. Je vhodné si uvědomit, že v případě reálné matice její hodnost nezáleží na tom, zda se na ni díváme jako na reálnou nebo komplexní matici. Plyne to z toho, jak jsme si všimli v odstavci 1.1.1, že v případě reálných vektorů jejich lineární závislost nebo nezávislost nezávisí na tom, zda se ne ně díváme jako na reálné nebo komplexní vektory.

Jelikož mezi vektory z $\mathbb{C}^n$ neexistuje více než n lineárně nezávislých vektorů, a řádků matice z $\mathbb{C}^{m\times n}$ je m, vidíme okamžitě, že pro $\mathbf{A} \in \mathbb{C}^{m\times n}$ je vždy

$$\operatorname{hod}(\mathbf{A}) \leq \min\{m, n\}.$$

Pokud zde platí rovnost, tj. pokud pro $\mathbf{A} \in \mathbb{C}^{m\times n}$ je $\operatorname{hod}(\mathbf{A}) = \min\{m, n\}$, pak řekneme, že $\mathbf{A}$ má *plnou hodnost* (nebo že $\mathbf{A}$ je *matice plné hodnosti*).

Čtenář jistě ze základního kurzu algebry zná následující zajímavé a důležité tvrzení:

> *Maximální počet lineárně nezávislých řádků matice je roven maximálnímu počtu lineárně nezávislých sloupců této matice.*

Znamená to tedy, že hodnost matice je také rovna maximálnímu počtu jejich lineárně nezávislých sloupců a vidíme, že platí

$$\operatorname{hod}(\mathbf{A}^T) = \operatorname{hod}(\mathbf{A}).$$

Pokud si uvědomíme, že vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ jsou lineárně nezávislé právě když jsou lineárně nezávislé vektory k nim komplexně sdružené, tj. vektory $\overline{\boldsymbol{x}}_1, \overline{\boldsymbol{x}}_2, \ldots, \overline{\boldsymbol{x}}_k$, pak odtud vidíme, že je také

$$\operatorname{hod}(\mathbf{A}^*) = \operatorname{hod}(\mathbf{A}) = \operatorname{hod}(\overline{\mathbf{A}}).$$

Je-li $\mathbf{A}$ čtvercová matice, $\mathbf{A} \in \mathbb{C}^{n\times n}$, pak řekneme, že $\mathbf{A}$ je *regulární matice*, je-li $\operatorname{hod}(\mathbf{A}) = n$, tj. jestliže její řádky (resp. sloupce) jsou lineárně nezávislé (nebo podle předchozího pokud $\mathbf{A}$ je plné hodnosti). Předpokládáme, že čtenáři je známo, že čtvercová matice je regulární právě když její determinant je nenulový. Definici determinantu zde nebudeme opakovat, odkazujeme čtenáře na základní kurz algebry. Připomeňme ale následující tvrzení týkající se hodnosti součinu matic.

1.2.10 Věta. *Buď* $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times k}$ *($m, n, k \in \mathbb{N}$). Potom*

$$\operatorname{hod}(\mathbf{AB}) \leq \min\bigl\{\operatorname{hod}(\mathbf{A}), \operatorname{hod}(\mathbf{B})\bigr\}.$$

Je-li $\alpha \in \mathbb{C}$, $\alpha \neq 0$, *pak* $\operatorname{hod}(\alpha\mathbf{A}) = \operatorname{hod}(\mathbf{A})$.

Je-li $\operatorname{hod}(\mathbf{B}) = n$ *(tj.* $k \geq n$ *a* $\mathbf{B}$ *má plnou hodnost), pak* $\operatorname{hod}(\mathbf{AB}) = \operatorname{hod}(\mathbf{A})$.

Je-li $m = n = k$ *(tj.* $\mathbf{A}, \mathbf{B}$ *jsou čtvercové matice řádu* n*) a je-li* $\operatorname{hod}(\mathbf{B}) = n$ *(tj.* $\mathbf{B}$ *je regulární), pak*

$$\operatorname{hod}(\mathbf{AB}) = \operatorname{hod}(\mathbf{BA}) = \operatorname{hod}(\mathbf{A}).$$

Zdůrazněme, že např. poslední rovnost je zaručena skutečně pouze za předpokladu, že $\mathbf{B}$ je regulární. Třeba příklad matic $\mathbf{AB}$ z odstavce 1.2.4, pro které $\mathbf{AB} \neq \mathbf{BA}$, zároveň ukazuje, že může být $\operatorname{hod}(\mathbf{AB}) \neq \operatorname{hod}(\mathbf{BA})$.

1.2.11. Čtenář se jistě setkal s pojmem inverzní matice (ke čtvercové matici). Připomeňme zde definici:

Buď $\mathbf{A} \in \mathbb{C}^{n\times n}$. *Řekneme, že* $\mathbf{A}$ ***má inverzní matici****, jestliže existuje matice* $\mathbf{B} \in \mathbb{C}^{n\times n}$ *tak, že*

$$\mathbf{AB} = \mathbf{BA} = \mathbf{E}.$$

Matici $\mathbf{B}$ *potom nazveme* ***inverzní matice*** *k matici* $\mathbf{A}$ *a značíme* $\mathbf{B} = \mathbf{A}^{-1}$.

Pokud $\mathbf{A} \in \mathbb{C}^{n\times n}$ má inverzní matici $\mathbf{A}^{-1}$, pak podle věty 1.2.10 je

$$\min\big\{\mathrm{hod}(\mathbf{A}), \mathrm{hod}(\mathbf{A}^{-1})\big\} \geq \mathrm{hod}(\mathbf{AA}^{-1}) = \mathrm{hod}(\mathbf{E}) = n,$$

tj. je nutně $\mathrm{hod}(\mathbf{A}) = \mathrm{hod}(\mathbf{A}^{-1}) = n$. Vidíme tedy, že pouze regulární matice může mít matici inverzní. Dá se ale ukázat, že platí i obrácené tvrzení, tj. ke každé regulární matici existuje matice inverzní. Dohromady:

Čtvercová matice má matici inverzní právě když je regulární.

V definici inverzní matice se požaduje splnění dvou podmínek — totiž, že platí

$$\mathbf{AA}^{-1} = \mathbf{E} \qquad \text{a zároveň} \qquad \mathbf{A}^{-1}\mathbf{A} = \mathbf{E}.$$

Všimněme si, že pokud použijeme výše uvedené tvrzení o existenci inverzní matice k matici regulární, pak je vidět, že stačí požadovat splnění pouze jedné z těchto dvou podmínek (mluvíme stále o čtvercových maticích). Buďte $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ a předpokládejme např., že je

$$\mathbf{AB} = \mathbf{E}.$$

Opět podle věty 1.2.10 vidíme, že je $\mathrm{hod}(\mathbf{A}) = \mathrm{hod}(\mathbf{B}) = n$, tj. $\mathbf{A}, \mathbf{B}$ jsou regulární. Matice $\mathbf{A}$ má tedy inverzní matici $\mathbf{A}^{-1}$. Potom je ale

$$\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{E} = \mathbf{A}^{-1}(\mathbf{AB}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{EB} = \mathbf{B},$$

tj. $\mathbf{B} = \mathbf{A}^{-1}$ a $\mathbf{B}$ je opravdu inverzní matice k $\mathbf{A}$. Zároveň jsme tak dokázali důležitou skutečnost, že *inverzní matice je určena jednoznačně* (pokud existuje).

Jednoduchý důkaz následujícího tvrzení přenecháváme čtenáři jako cvičení.

Buďte $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ *regulární matice. Potom také* $\mathbf{AB}$ *je regulární a platí*

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Je-li dále $a \in \mathbb{C}$, $a \neq 0$, *pak*

$$(a\mathbf{A})^{-1} = \frac{1}{a}\mathbf{A}^{-1}.$$

Na rovnosti $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ si zvláště všimněte pořadí matic v součinu napravo.

1.2.12. V dalším bude občas užitečné pracovat s maticemi v tzv. *blokovém zápisu*; také se mluví o *blokových maticích* a pod. Naznačíme, co se tímto termínem myslí, aniž bychom uváděli nějakou formální definici. Uvažujme třeba matici $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{5\times 6}$. Tuto matici můžeme „rozdělit svislými a vodorovnými čarami" na části různým způsobem, např. následujícím:

$$\mathbf{A} = \left(\begin{array}{cc|c|ccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ \hline a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{array}\right).$$

Na jednotlivé části $\mathbf{A}$ „oddělené čarami" se můžeme dívat jako na samostatné matice. Označíme-li zde

$$\mathbf{A}_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad \mathbf{A}_{12} = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}, \qquad \mathbf{A}_{13} = \begin{pmatrix} a_{14} & a_{15} & a_{16} \\ a_{24} & a_{25} & a_{26} \end{pmatrix},$$

$$\mathbf{A}_{21} = \begin{pmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix}, \qquad \mathbf{A}_{22} = \begin{pmatrix} a_{33} \\ a_{43} \end{pmatrix}, \qquad \mathbf{A}_{23} = \begin{pmatrix} a_{34} & a_{35} & a_{36} \\ a_{44} & a_{45} & a_{46} \end{pmatrix},$$

$$\mathbf{A}_{31} = \begin{pmatrix} a_{51} & a_{52} \end{pmatrix}, \qquad \mathbf{A}_{32} = \begin{pmatrix} a_{53} \end{pmatrix}, \qquad \mathbf{A}_{33} = \begin{pmatrix} a_{54} & a_{55} & a_{56} \end{pmatrix},$$

můžeme formálně psát

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{pmatrix}$$

nebo stručně $\mathbf{A} = (\mathbf{A}_{ij})$; zde nesmíme zaměňovat symbol $\mathbf{A}_{ij}$, který značí matici, s dříve definovaným symbolem $[\mathbf{A}]_{ij}$, kterým jsme značili i-j-tý prvek matice $\mathbf{A}$ (tj. $[\mathbf{A}]_{ij} = a_{ij}$). Tyto matice $\mathbf{A}_{ij}$ nazýváme *bloky* matice $\mathbf{A}$ (také se někdy mluví o podmaticích a pod.). Všimněme si, že zde všechny matice v jednom „blokovém řádku" mají vždy stejný počet řádků a všechny matice v jednom „blokovém sloupci" mají stejný počet sloupců — pouze za tohoto předpokladu lze obecně z bloků sestavit větší matici.

Obecně, je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$ a dále $p, q \in \mathbb{N}$ a $m_i, n_j \in \mathbb{N}$ pro $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$ takové, že

$$m_1 + m_2 + \cdots + m_p = m, \qquad n_1 + n_2 + \cdots + n_q = n,$$

pak matici $\mathbf{A}$ lze rozdělit na bloky (zapsat blokově) ve tvaru

$$\mathbf{A} = (\mathbf{A}_{ij}) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1q} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{p1} & \mathbf{A}_{p2} & \dots & \mathbf{A}_{pq} \end{pmatrix},$$

kde pro $i = 1, 2, \dots, p$, $j = 1, 2, \dots, q$,

$$\mathbf{A}_{ij} \in \mathbb{C}^{m_i\times n_j}$$

jsou příslušné bloky. Jsou-li naopak takovéto matice $\mathbf{A}_{ij}$ dané, lze z nich blokově sestavit matici $\mathbf{A}$.

Speciální případ blokového zápisu je rozdělení matice na řádky nebo na sloupce — s tímto jsme se vlastně již setkali aniž bychom použili název bloky. Připomeňme, že je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$, pak symbolem $\mathbf{A}(i,:)$ značíme i-tý řádek matice $\mathbf{A}$, $\mathbf{A}(:,j)$ značí j-tý sloupec $\mathbf{A}$. Podle právě uvedeného značení můžeme tedy psát

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}(:,1) & \mathbf{A}(:,2) & \dots & \mathbf{A}(:,n) \end{pmatrix} \tag{1.12}$$

a podobně

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}(1,:) \\ \mathbf{A}(2,:) \\ \vdots \\ \mathbf{A}(m,:) \end{pmatrix}.$$

V zápise (1.12) se někdy pro větší srozumitelnost mezi jednotlivé sloupce přidávají svislé čáry, tj. píšeme

$$\mathbf{A} = \left(\begin{array}{c|c|c|c} \mathbf{A}(:,1) & \mathbf{A}(:,2) & \dots & \mathbf{A}(:,n) \end{array} \right).$$

Je-li $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{C}^m$, pak zápis

$$\mathbf{B} = \left(\begin{array}{c|c|c|c} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_k \end{array} \right)$$

má zřejmý význam — $\mathbf{B}$ je matice z $\mathbb{C}^{m\times k}$ jejíž sloupce jsou rovny daným vektorům, $\mathbf{B}(:,j) = \boldsymbol{x}_j$ pro $j = 1, 2, \dots, k$. Podobně můžeme sestavit matici

$$\mathbf{C} = \begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_k^T \end{pmatrix};$$

zde je $\mathbf{C} \in \mathbb{C}^{k\times m}$ a řádky $\mathbf{C}$ jsou rovny transpozicím vektorů $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k$ (rozmyslete si, že zde je $\mathbf{C} = \mathbf{B}^T$).

Uveďme ještě jedno značení, které se v souvislosti s blokovým zápisem matic používá. Je-li $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{m\times n}$, $i_1, i_2, j_1, j_2 \in \mathbb{N}$,

$$1 \le i_1 \le i_2 \le m, \qquad 1 \le j_1 \le j_2 \le n,$$

pak symbolem $\mathbf{A}(i_1{:}i_2, j_1{:}j_2)$ rozumíme blok, který z matice $\mathbf{A}$ dostaneme tak, že uvažujeme řádky od i_1 do i_2 a sloupce od j_1 do j_2 (přesněji příslušné části těchto řádků a sloupců); je tedy např.

$$\mathbf{A}(3{:}5, 2{:}3) = \begin{pmatrix} a_{32} & a_{33} \\ a_{42} & a_{43} \\ a_{52} & a_{53} \end{pmatrix}.$$

Formálně zapsáno

$$\mathbf{A}(i_1{:}i_2, j_1{:}j_2) = (a_{ij})_{\substack{i=i_1,i_1+1,\dots,i_2 \\ j=j_1,j_1+1,\dots,j_2}}.$$

Poznamenejme, že při tomto značení je symbol $\mathbf{A}(i,:)$ vlastně zkratka pro $\mathbf{A}(i{:}i, 1{:}n)$ a symbol $\mathbf{A}(:,j)$ je zkratka pro $\mathbf{A}(1{:}m, j{:}j)$ (když $\mathbf{A} \in \mathbb{C}^{m\times n}$).

1.2.13. Za vhodných předpokladů lze blokově rozdělené matice „blokově násobit“. Uvažujme matice $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times k}$ a předpokládejme, že $\mathbf{A}$ a $\mathbf{B}$ jsou rozděleny na bloky $\mathbf{A}_{ij}, \mathbf{B}_{ij}$,

$$\mathbf{A} = \big(\mathbf{A}_{ij}\big)_{\substack{i=1,2,\dots,p\\ j=1,2,\dots,q}}, \qquad \mathbf{B} = \big(\mathbf{B}_{ij}\big)_{\substack{i=1,2,\dots,q\\ j=1,2,\dots,r}},$$

kde $p, q, r \in \mathbb{N}$ (všimněte si, že q je společné pro obě matice $\mathbf{A}, \mathbf{B}$). Předpokládejme dále, že rozdělení na bloky je takové, že pro každé

$$i = 1, 2, \dots, p, \qquad j = 1, 2, \dots, r, \qquad s = 1, 2, \dots, q$$

má součin $\mathbf{A}_{is}\mathbf{B}_{sj}$ smysl, tj. počet sloupců matice $\mathbf{A}_{is}$ je stejný jako počet řádků matice $\mathbf{B}_{sj}$. Potom se dá ukázat (podrobným rozepsáním), že součin $\mathbf{AB}$ lze blokově zapsat ve tvaru

$$\mathbf{AB} = (\mathbf{C}_{ij}) = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \dots & \mathbf{C}_{1r} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \dots & \mathbf{C}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{p1} & \mathbf{C}_{p2} & \dots & \mathbf{C}_{pr} \end{pmatrix}, \tag{1.13}$$

kde

$$\mathbf{C}_{ij} = \sum_{s=1}^{q} \mathbf{A}_{is}\mathbf{B}_{sj} = \mathbf{A}_{i1}\mathbf{B}_{1j} + \mathbf{A}_{i2}\mathbf{B}_{2j} + \dots + \mathbf{A}_{iq}\mathbf{B}_{qj}. \tag{1.14}$$

Rozmyslete si, že součet napravo v (1.14) má smysl, tj. všechny součiny $\mathbf{A}_{is}\mathbf{B}_{sj}$ jsou matice téhož typu. Dále si všimněte, že rovnost (1.14) je formálně podobná rovnosti (1.6) v definici součinu matic — pouze místo násobení čísel se zde jedná o násobení matic (příslušných bloků).

Jeden ze speciálních případů blokového násobení matic je, když matice napravo v součinu je rozdělena na sloupce a na matici nalevo se díváme jako na jeden blok. Je-li $\mathbf{A} \in \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times k}$, pak pro každé $j = 1, 2, \dots, k$ je $\mathbf{B}(:, j) \in \mathbb{C}^{n\times 1}$ a součin $\mathbf{A}\cdot\mathbf{B}(:, j)$ má smysl a je to opět matice o jednom sloupci, $\mathbf{A}\cdot\mathbf{B}(:, j) \in \mathbb{C}^{m\times 1}$ [nebo jednoduše vektor z $\mathbb{C}^m$; vlastně se na součin $\mathbf{A}\cdot\mathbf{B}(:, j)$ můžeme dívat jako na součin matice a vektoru — viz též rovnost (1.9) z odstavce 1.2.3]. Podle předchozího je

$$\begin{aligned} \mathbf{AB} &= \mathbf{A}\cdot\Big(\ \mathbf{B}(:,1)\ \big|\ \mathbf{B}(:,2)\ \big|\ \dots\ \big|\ \mathbf{B}(:,k)\ \Big) = \\ &= \Big(\ \mathbf{A}\cdot\mathbf{B}(:,1)\ \big|\ \mathbf{A}\cdot\mathbf{B}(:,2)\ \big|\ \dots\ \big|\ \mathbf{A}\cdot\mathbf{B}(:,k)\ \Big). \end{aligned} \tag{1.15}$$

Zapsáno jinak, pro $j = 1, 2, \dots, k$ máme

$$\big(\mathbf{AB}\big)(:, j) = \mathbf{A}\cdot\mathbf{B}(:, j), \tag{1.16}$$

tj., řečeno slovy: *j-tý sloupec součinu* $\mathbf{A}, \mathbf{B}$ *je roven součinu matice* $\mathbf{A}$ *a j-tého sloupce* $\mathbf{B}$. Tuto vlastnost součinu matic budeme dále používat. Zkuste si rozmyslet, že k důkazu rovnosti (1.16) [tj. i (1.15)] není třeba použít obecné tvrzení o násobení blokově rozdělených matic, ale je přímo vidět z definice součinu matic.

Podobně bychom mohli zjistit, že jednotlivé řádky součinu $\mathbf{AB}$ jsou rovny součinům jednotlivých řádků matice $\mathbf{A}$ a (celé) matice $\mathbf{B}$,

$$\big(\mathbf{AB}\big)(i,:) = \mathbf{A}(i,:)\cdot\mathbf{B}$$

($\mathbf{A}\in\mathbb{C}^{m\times n}$, $\mathbf{B}\in\mathbb{C}^{n\times k}$, $i=1,2,\dots,m$).

1.2.14. V dalším se také setkáme s pojmem *blokově trojúhelníkové matice* a *blokově diagonální matice* (též se mluví o *kvazidiagonální matici*). Buď $\mathbf{A}$ čtvercová matice, $\mathbf{A}\in\mathbb{C}^{n\times n}$ a předpokládejme, že $\mathbf{A}$ je rozdělena na bloky ve tvaru

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \square & \dots & \square \\ \square & \mathbf{A}_2 & \dots & \square \\ \vdots & \vdots & \ddots & \vdots \\ \square & \square & \dots & \mathbf{A}_k \end{pmatrix}, \tag{1.17}$$

kde $\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_k$ jsou *čtvercové* matice řádů $n_1,n_2,\dots,n_k$ (je potom ovšem $n_1+n_2+ {}+\dots+n_k=n$). Zde symboly $\square$ označují další bloky, na které je $\mathbf{A}$ rozdělena (tyto bloky obecně nejsou čtvercové — rozmyslete si, jaké typy budou tyto bloky mít). Části matice $\mathbf{A}$, která se skládá z bloků $\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_k$ se říká *bloková diagonála* (kvazidiagonála). Jestliže buď všechny bloky „nad“ nebo všechny bloky „pod“ kvazidiagonálou jsou nulové, pak řekneme, že matice $\mathbf{A}$ je *blokově trojúhelníková* (dolní nebo horní blokově trojúhelníková). Připomeňme známý fakt, že je-li $\mathbf{A}$ [tvaru (1.17)] blokově trojúhelníková, pak pro determinant matice $\mathbf{A}$ platí

$$\det\mathbf{A} = \det\mathbf{A}_1\cdot\det\mathbf{A}_2\cdots\det\mathbf{A}_k, \tag{1.18}$$

což je vlastně analogie skutečnosti, že determinant trojúhelníkové matice je roven součinu diagonálních prvků této matice (předpokládáme, že čtenář je seznámen s pojmem determinantu matice a se základními vlastnostmi determinantu — viz např. skripta [6]).

Jestliže $\mathbf{A}$ tvaru (1.17) má dokonce všechny bloky mimo blokovou diagonálu nulové, pak řekneme, že $\mathbf{A}$ je *kvazidiagonální* nebo *blokově diagonální* (stále za předpokladu, že bloky $\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_k$ jsou čtvercové). Poznamenejme, že pojmy blokově trojúhelníková nebo blokově diagonální matice jsou relativní, neboť závisí na tom, jaké rozdělení matice na bloky uvažujeme. Dále ještě poznamenejme, že trojúhelníková matice nebo diagonální matice jsou nyní pouze speciální případy blokově trojúhelníkové matice nebo kvazidiagonální matice. Je-li $\mathbf{A}$ tvaru (1.17) blokově diagonální, pak ještě používáme zkrácený zápis

$$\mathbf{A} = \operatorname{diag}\big[\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_k\big].$$

Buď dále $\mathbf{B}$ také kvazidiagonální matice o stejném počtu diagonálních bloků jako $\mathbf{A}$,

$$\mathbf{B} = \operatorname{diag}\big[\mathbf{B}_1,\mathbf{B}_2,\dots,\mathbf{B}_k\big]$$

a předpokládejme, že pro každé $j=1,2,\dots,k$ jsou matice $\mathbf{A}_j,\mathbf{B}_j$ stejného řádu. Potom ovšem $\mathbf{A}$ a $\mathbf{B}$ mají stejný řád a součin $\mathbf{AB}$ má smysl. Dá se ukázat, že v tomto případě je $\mathbf{AB}$ také blokově diagonální, přičemž

$$\mathbf{AB} = \operatorname{diag}\big[\mathbf{A}_1\mathbf{B}_1,\mathbf{A}_2\mathbf{B}_2,\dots,\mathbf{A}_k\mathbf{B}_k\big]$$

[zkuste si rozmyslet, jak tato rovnost plyne z rovností (1.13), (1.14) pro násobení obecných blokově rozdělených matic]. Speciálně pro $\mathbf{A}$ daného tvaru je

$$\mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A} = \operatorname{diag}\big[\mathbf{A}_1^2, \mathbf{A}_2^2, \dots, \mathbf{A}_k^2\big]$$

nebo obecněji

$$\mathbf{A}^r = \underbrace{\mathbf{A} \cdot \mathbf{A} \cdots \mathbf{A}}_{r\text{-krát}} = \operatorname{diag}\big[\mathbf{A}_1^r, \mathbf{A}_2^r, \dots, \mathbf{A}_k^r\big]$$

pro $r \in \mathbb{N}$. Jelikož, jak jsme již poznamenali, diagonální matice je speciální případ blokově diagonální matice, pro diagonální matice

$$\mathbf{A} = \operatorname{diag}[a_1, a_2, \dots, a_n], \qquad \mathbf{B} = \operatorname{diag}[b_1, b_2, \dots, b_n]$$

$(a_i, b_i \in \mathbb{C})$ platí

$$\begin{aligned} \mathbf{AB} &= \operatorname{diag}[a_1 b_1, a_2 b_2, \dots, a_n b_n], \\ \mathbf{A}^r &= \operatorname{diag}[a_1^r, a_2^r, \dots, a_n^r]. \end{aligned}$$

K důkazu těchto rovností ovšem opět není třeba mluvit o součinu blokově diagonálních matic, neboť jsou snadno vidět přímo z definice součinu matic.

1.3 Matice zobrazení a podobnost matic

Mnohdy bývá užitečné se na matice dívat jako na na lineární zobrazení nebo v případě čtvercových matic jako na operátory. Vyšetřování matic tak lze převést na obecnější případ vyšetřování lineárních zobrazení nebo operátorů. I když dále budeme mluvit většinou pouze o maticích, zastavme se alespoň velice stručně u pojmu lineárního zobrazení a jeho vztahu k maticím, zvláště v souvislosti s násobením matic a tzv. podobnosti matic. Čtenář si možná ze základního kurzu lineární algebry pamatuje, že existuje vztah mezi některými vlastnostmi lineárních zobrazení a obecnými vlastnostmi řešení soustav lineárních algebraických rovnic. Jelikož řešení soustav není obsahem našeho textu, tímto vztahem se zde nebudeme zabývat.

Předpokládáme, že čtenář je dostatečně seznámen s obecným pojmem *zobrazení*. Připomeňme pouze, že jsou-li M, N dvě množiny, pak zápisem

$$A\colon M \to N$$

se rozumí, že A je nějaké zobrazení množiny M do množiny N. Pokud pro $x \in M$ je $A(x) = y$ $(\in N)$, pak také někdy píšeme

$$A\colon x \mapsto y$$

nebo dokonce krátce $x \mapsto y$.

Z předchozího ještě připomeňme, že zápisem $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$) myslíme, že P je lineární podprostor v $\mathbb{C}^n$ (resp. v $\mathbb{R}^n$); viz definici 1.1.4. Dále např. zápisem $P \subset\subset Q \subset\subset \mathbb{C}^n$ se myslí, že P, Q jsou lineární podprostory v $\mathbb{C}^n$ a $P \subset Q$.

Podobně jako v předchozím, stále zde předpokládáme, že m, n jsou nějaká daná přirozená čísla.

1.3.1 Definice. *Nechť $P \subset\subset \mathbb{C}^n$, $Q \subset\subset \mathbb{C}^m$ (resp. $P \subset\subset \mathbb{R}^n$, $Q \subset\subset \mathbb{R}^m$), $A\colon P \to Q$. Řekneme, že zobrazení A je* ***lineární****, jestliže*

(i) *A je* ***aditivní****, tj. pro každé $\boldsymbol{x}, \boldsymbol{y} \in P$ je $A(\boldsymbol{x}+\boldsymbol{y}) = A(\boldsymbol{x}) + A(\boldsymbol{y})$,*

(ii) *A je* ***homogenní****, tj. pro každé $\boldsymbol{x} \in P$ a každé $a \in \mathbb{C}$ (resp. $a \in \mathbb{R}$) je $A(a\boldsymbol{x}) = aA(\boldsymbol{x})$.*

Množinu všech lineárních zobrazení $A\colon P \to Q$ značíme $\mathscr{L}(P,Q)$. Je-li speciálně $m = n$ a $Q = P$, pak značíme $\mathscr{L}(P,P) = \mathscr{L}(P)$. Prvky z $\mathscr{L}(P)$, tj. lineární zobrazení $A\colon P \to P$, nazýváme ***lineární operátory*** *na P.*

Poznamenejme, že reálný případ od komplexního, tj. případ $P \subset\subset \mathbb{C}^n$, $Q \subset\subset \mathbb{C}^m$ a $P \subset\subset \mathbb{R}^n$, $Q \subset\subset \mathbb{R}^m$, se liší pouze tím, že v rovnosti $A(a\boldsymbol{x}) = aA(\boldsymbol{x})$ v bodu (ii) uvažujeme v komplexním případě $a \in \mathbb{C}$, zatímco v reálném případě jenom $a \in \mathbb{R}$. Pokud dále nebude řečeno nic jiného, budeme stále uvažovat obecně komplexní případ, tj.

$$P \subset\subset \mathbb{C}^n, \qquad Q \subset\subset \mathbb{C}^m.$$

Můžeme si všimnout, že z bodů (i), (ii) v definici okamžitě plyne, že v lineárním zobrazení je obraz lineární kombinace roven stejné lineární kombinaci obrazů, tj. je-li

$$A \in \mathscr{L}(P,Q), \qquad \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in P, \qquad a_1, a_2, \ldots, a_k \in \mathbb{C}$$

($k \in \mathbb{N}$), pak

$$A\left(\sum_{i=1}^{k} a_i \boldsymbol{x}_i\right) = \sum_{i=1}^{k} a_i A(\boldsymbol{x}_i). \tag{1.19}$$

1.3.2 Poznámka. Přestože předpokládáme, že čtenář je seznámen s obecným pojmem zobrazení, připomeňme alespoň některé další základní věci. Nechť M, N jsou nějaké dvě neprázdné množiny a uvažujme zobrazení

$$A\colon M \to N.$$

Řekneme, že A je *prosté zobrazení*, jestliže pro každé $x, y \in M$, $x \neq y$, je také $A(x) \neq A(y)$ [zapsáno ekvivalentně, pro $x, y \in M$ platí implikace $A(x) = A(y) \implies x = y$]. Řekneme, že A je *zobrazení na N*, jestliže pro každé $z \in N$ existuje $x \in M$ tak, že $A(x) = z$.

Je-li $M_1 \subset M$, pak *obraz množiny* M_1 při zobrazení A značíme $A(M_1)$ a definujeme rovností

$$A(M_1) = \bigl\{A(x) \bigm| x \in M_1\bigr\}.$$

Vidíme, že $A\colon M \to N$ je zobrazení na, právě když $A(M) = N$.

Je-li $N_1 \subset N$, pak *vzor množiny* N_1 při zobrazení A značíme $A^{-1}(N_1)$ a definujeme rovností

$$A^{-1}(N_1) = \bigl\{x \in M \bigm| A(x) \in N_1\bigr\}.$$

Všimněme si, že pro $M_1 \subset M$ je $A(M_1) \subset N$ a pro $N_1 \subset N$ je $A^{-1}(N_1) \subset M$.

Buď $A\colon M \to N$ a předpokládejme, že A je zobrazení prosté a na (zobrazení, které je prosté a na, se také nazývá *vzájemně jednoznačné zobrazení*). Jelikož tedy $A(M) = N$, pro každé $y \in N$ existuje $x \in M$ tak, že $A(x) = y$. Jelikož A je prosté, vidíme (přímo z definice prostého zobrazení), že toto x je určeno jednoznačně. Pro dané $y \in N$ označme $A^{-1}(y) = x$ [kde tedy x je takové, že $A(x) = y$]. Tímto ale definujeme zobrazení

$$A^{-1} \colon N \to M.$$

Toto zobrazení se nazývá *inverzní zobrazení* k zobrazení A. Vidíme, že přitom pro každé $x \in M$ je

$$A^{-1}\big(A(x)\big) = x$$

a podobně pro každé $y \in N$ je

$$A\big(A^{-1}(y)\big) = y.$$

Všimněme si ještě, že je-li zobrazení $A \colon M \to N$ prosté a na, pak pro $N_1 \subset N$ je vzor množiny N_1, $A^{-1}(N_1)$, jednoduše rovno obrazu N_1 při zobrazení A^{-1}.

1.3.3 Poznámka. V poznámce 1.1.11 jsme definovali pojem souřadnic a souřadnicového vektoru při dané bázi. Je-li $\mathcal{B} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ báze P (předpokládáme, že podprostor P není triviální), pak pro $\boldsymbol{y} \in P$ jsme symbolem $(\boldsymbol{y})_\mathcal{B}$ značili vektor souřadnic v bázi $\mathcal{B}$. Připomeňme, že je potom

$$(\boldsymbol{y})_\mathcal{B} = (y_1, y_2, \ldots, y_k)^T \in \mathbb{C}^k,$$

a

$$\boldsymbol{y} = \sum_{i=1}^{k} y_i \boldsymbol{x}_i; \tag{$*$}$$

vektor souřadnic je bodem $\boldsymbol{y}$ určen jednoznačně. Je-li naopak $(y_1, y_2, \ldots, y_k)^T \in \mathbb{C}^k$ libovolný vektor, pak je rovností $(*)$ určen (jednoznačně) bod $\boldsymbol{y} \in P$, pro který je $(\boldsymbol{y})_\mathcal{B} = (y_1, y_2, \ldots, y_k)^T$. Nyní vidíme, že definujeme-li zobrazení $A \colon P \to \mathbb{C}^k$ rovností

$$A(\boldsymbol{y}) = (\boldsymbol{y})_\mathcal{B}, \tag{1.20}$$

pak toto zobrazení je vzájemně jednoznačné. Přenecháváme čtenáři aby ověřil, že toto zobrazení A je lineární. Tomuto zobrazení se říká *souřadnicový izomorfismus*. Podobně v reálném případě.

Obecně řekneme, že lineární zobrazení $A \in \mathscr{L}(P, Q)$ je *izomorfní zobrazení* (krátce *izomorfismus*), jestliže je vzájemně jednoznačné (tj. prosté a na). Pokud existuje izomorfismus $A \colon P \to Q$, pak řekneme, že *prostory* P, Q *jsou izomorfní*. Připomeňme, že $P \subset\subset \mathbb{C}^n$, $Q \subset\subset \mathbb{C}^m$ jsou izomorfní *právě když* mají stejnou dimenzi, $\dim P = \dim Q$ (podobně pro případ podprostorů v $\mathbb{R}^n$ a $\mathbb{R}^m$). V algebraickém smyslu bývá zvykem izomorfní prostory ztotožňovat. V předchozím jsme „přirozeně“ ztotožnili např. prostory $\mathbb{C}^n$ a $\mathbb{C}^{n \times 1}$. Rozmyslete si, že se vlastně jednalo o ztotožnění ve smyslu izomorfismu.

1.3.4. Mezi lineárními zobrazeními lze přirozeným způsobem definovat sčítání zobrazení, násobek zobrazení skalárem a skládání (násobek) dvou lineárních zobrazení. Jsou-li $A, B \in \mathscr{L}(P, Q)$, $a \in \mathbb{C}$, pak definujeme součet $A + B$ tak, že položíme

$$\big(A + B\big)(\boldsymbol{x}) = A(\boldsymbol{x}) + B(\boldsymbol{x})$$

pro $\boldsymbol{x} \in P$, a podobně položíme

$$\big(aA\big)(\boldsymbol{x}) = aA(\boldsymbol{x}).$$

Je ovšem $(A+B)\colon P \to Q$, $aA\colon P \to Q$ a zobrazení $(A+B)$ a aA jsou lineární (ověřte).

Je-li dále $V \subset\subset \mathbb{C}^k$ ($k \in \mathbb{N}$), $A \in \mathscr{L}(Q,V)$, $B \in \mathscr{L}(P,Q)$, pak pro každé $\boldsymbol{x} \in P$ položíme

$$\big(AB\big)(\boldsymbol{x}) = A\big(B(\boldsymbol{x})\big).$$

Potom je $AB\colon P \to V$ a opět je snadno vidět, že AB je zobrazení lineární, tj. $AB \in \in \mathscr{L}(P,V)$. Složené zobrazení AB se nazývá *součin zobrazení* A, B (také píšeme $AB = = A \cdot B$).

Mezi lineárními zobrazeními z $\mathscr{L}(P)$ má zvláštní význam tzv. *identické zobrazení* (krátce *identita*), tj. zobrazení $I\colon P \to P$, které je definováno rovností

$$I(\boldsymbol{x}) = \boldsymbol{x} \qquad (\boldsymbol{x} \in P)$$

(že takto definované zobrazení je lineární, je ovšem zřejmé). V poznámce 1.3.2 jsme připomněli obecný pojem inverzního zobrazení. Přímo z definice inverzního zobrazení je vidět, že $A \in \mathscr{L}(P)$ má inverzní zobrazení A^{-1} právě když A je izomorfismus; potom je

$$A \cdot A^{-1} = I = A^{-1} \cdot A. \tag{1.21}$$

Dále je zřejmé, že pro každé $A \in \mathscr{L}(P)$ je

$$A \cdot I = A = I \cdot A. \tag{1.22}$$

1.3.5. Konečně připomeňme pojem matice lineárního zobrazení. Nechť opět $P \subset\subset \mathbb{C}^n$, $Q \subset\subset \mathbb{C}^m$ jsou netriviální podprostory. Nechť $\mathcal{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$ je nějaká báze P a $\mathcal{B} = = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_r\}$ je báze Q (je tedy $\dim P = k$, $\dim Q = r$). Je-li $A \in \mathscr{L}(P,Q)$, pak *maticí zobrazení* A vzhledem k bázím $\mathcal{A}, \mathcal{B}$ rozumíme matici ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}}$ typu $r \times k$ takovou, že její j-tý ($j = 1, 2, \ldots, k$) sloupec je roven souřadnicovému vektoru $(A(\boldsymbol{a}_j))_{\mathcal{B}}$ (viz poznámku 1.1.11). Formálně můžeme psát

$${}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}} = \Big(\ \big(A(\boldsymbol{a}_1)\big)_{\mathcal{B}} \,\Big|\, \big(A(\boldsymbol{a}_2)\big)_{\mathcal{B}} \,\Big|\, \ldots \,\Big|\, \big(A(\boldsymbol{a}_k)\big)_{\mathcal{B}}\ \Big). \tag{1.23}$$

Je-li ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}} = (a_{ij})$, pak podle definice souřadného vektoru lze obraz vektoru $\boldsymbol{a}_j$ při zobrazení A psát ve tvaru

$$A(\boldsymbol{a}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{b}_i. \tag{1.24}$$

Buď nyní $\boldsymbol{x} \in P$ libovolný. Je-li $(\boldsymbol{x})_{\mathcal{A}} = (x_1, x_2, \ldots, x_k)^T$, je tedy

$$\boldsymbol{x} = \sum_{j=1}^{k} x_j \boldsymbol{a}_j$$

a (vzhledem k linearitě A) z (1.24) dostáváme

$$A(\boldsymbol{x}) = A\Big(\sum_{j=1}^{k} x_j \boldsymbol{a}_j\Big) = \sum_{j=1}^{k} x_j A(\boldsymbol{a}_j) = \sum_{j=1}^{k} x_j \sum_{i=1}^{r} a_{ij} \boldsymbol{b}_j = \sum_{i=1}^{r} \boldsymbol{b}_i \sum_{j=1}^{k} a_{ij} x_j.$$

Tato rovnost ovšem znamená, že i-tá souřadnice vektoru $A(\boldsymbol{x})$ v bázi $\mathcal{B}$ je rovna součtu $\sum_{j=1}^{k} a_{ij} x_j$, což je ovšem i-tá složka součinu matice ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}}$ a souřadného vektoru $(\boldsymbol{x})_{\mathcal{A}}$. Zapsáno jinak

$$\big(A(\boldsymbol{x})\big)_{\mathcal{B}} = {}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}} \cdot (\boldsymbol{x})_{\mathcal{A}}. \tag{1.25}$$

Podle této rovnosti tedy každé lineární zobrazení $A \in \mathscr{L}(P, Q)$ můžeme reprezentovat pomocí jeho matice. Je-li naopak dána matice $\mathbf{A} \in \mathbb{C}^{r \times k}$, pak můžeme definovat zobrazení $A\colon P \to Q$ tak, že pro $\boldsymbol{x} \in P$ položíme

$$\big(A(\boldsymbol{x})\big)_{\mathcal{B}} = \mathbf{A} \cdot (\boldsymbol{x})_{\mathcal{A}}.$$

Snadno je vidět, že toto zobrazení je lineární. Dostáváme tak *vzájemně jednoznačný vztah mezi lineárními zobrazeními z $\mathscr{L}(P, Q)$ a maticemi z $\mathbb{C}^{r \times k}$* (vše při pevně daných bázích prostorů P, Q). Rozmyslete si zvláště případ $P = \mathbb{C}^n$, $Q = \mathbb{C}^m$. Pokud nyní $\mathcal{A}$ bude standardní báze v $\mathbb{C}^n$, $\mathcal{B}$ standardní báze v $\mathbb{C}^m$, pak rovnost (1.25) má jednoduše tvar

$$A(\boldsymbol{x}) = {}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}} \cdot \boldsymbol{x}. \tag{1.26}$$

Podobně v reálném případě.

V případě $Q = P$ a lineárního zobrazení $A \in \mathscr{L}(P)$ bývá zvykem většinou uvažovat v P jednu bázi (rozmyslete si, že teoreticky bychom mohli v P uvažovat dvě různé báze, v jedné bychom vyjadřovali vzory, v druhé obrazy při daném zobrazení). Je-li $\mathcal{A}$ taková báze, pak matici zobrazení A vzhledem k bázi $\mathcal{A}$ podle předchozího značíme ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$. Bude-li zřejmé, vzhledem k jaké bázi matici zobrazení uvažujeme, můžeme krátce psát $\mathbf{A}$ (častý je případ $P = Q = \mathbb{C}^n$, kdy v $\mathbb{C}^n$ uvažujeme standardní bázi).

Jednoduchý ale důležitý vztah je mezi operacemi na maticích a na lineárních zobrazeních. Zhruba se dá říci, že operacím na lineárních zobrazeních odpovídají stejné operace na maticích těchto zobrazení. Uvažujme $P \subset\subset \mathbb{C}^n$, $Q \subset\subset \mathbb{C}^m$, $V \subset\subset \mathbb{C}^k$ a nechť $\mathcal{A}$ je nějaká báze P, $\mathcal{B}$ báze Q a $\mathcal{C}$ báze V (předpokládáme, že podprostory P, Q, V jsou netriviální). Potom platí:

(a) *Je-li $A \in \mathscr{L}(P, Q)$, $a \in \mathbb{C}$, pak matice zobrazení aA při bázích $\mathcal{A}, \mathcal{B}$ je rovna $a \cdot {}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}}$.*

(b) *Nechť $A, B \in \mathscr{L}(P, Q)$, $C = A + B$. Potom*

$${}^{\mathcal{B}}\mathbf{C}^{\mathcal{A}} = {}^{\mathcal{B}}\mathbf{A}^{\mathcal{A}} + {}^{\mathcal{B}}\mathbf{B}^{\mathcal{A}}$$

(slovy: matice součtu zobrazení je rovna součtu matic těchto zobrazení).

(c) *Buďte $A \in \mathscr{L}(Q, V)$, $B \in \mathscr{L}(P, Q)$, $C = AB$ [je $AB \in \mathscr{L}(P, V)$]. Potom*

$${}^{\mathcal{C}}\mathbf{C}^{\mathcal{A}} = {}^{\mathcal{C}}\mathbf{A}^{\mathcal{B}} \cdot {}^{\mathcal{B}}\mathbf{B}^{\mathcal{A}}$$

(slovy: matice součinu zobrazení je rovna součinu matic těchto zobrazení).

(d) *Matice identického zobrazení* $I \in \mathscr{L}(P)$ *(v bázi* $\mathcal{A}$*) je rovna jednotkové matici* $\mathbf{E}$ *řádu* $r = \dim P$.

(e) *Matice zobrazení* $A \in \mathscr{L}(P)$ *je regulární právě když* A *je prosté. Potom je*

$$^{\mathcal{A}}\big(\mathbf{A}^{-1}\big)^{\mathcal{A}} = \big({}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}\big)^{-1}$$

[$^{\mathcal{A}}\big(\mathbf{A}^{-1}\big)^{\mathcal{A}}$ *značí matici zobrazení* A^{-1}].

Všimněme si třeba, jak jednoduše plyne bod (c) z asociativity násobení matic (viz tvrzení 1.2.5) a násobku zobrazení jako složeného zobrazení. V daném případě pro každé $\boldsymbol{x} \in P$ je

$$\begin{aligned} {}^{\mathcal{C}}\mathbf{C}^{\mathcal{A}} \cdot (\boldsymbol{x})_{\mathcal{A}} = \big(C(\boldsymbol{x})\big)_{\mathcal{C}} = \Big(A\big(B(\boldsymbol{x})\big)\Big)_{\mathcal{C}} = \\ = {}^{\mathcal{C}}\mathbf{A}^{\mathcal{B}} \cdot \big(B(\boldsymbol{x})\big)_{\mathcal{B}} = {}^{\mathcal{C}}\mathbf{A}^{\mathcal{B}} \cdot \big({}^{\mathcal{B}}\mathbf{B}^{\mathcal{A}} \cdot (\boldsymbol{x})_{\mathcal{A}}\big) = \big({}^{\mathcal{C}}\mathbf{A}^{\mathcal{B}} \cdot {}^{\mathcal{B}}\mathbf{B}^{\mathcal{A}}\big) \cdot (\boldsymbol{x})_{\mathcal{A}}, \end{aligned}$$

tj. opravdu ${}^{\mathcal{C}}\mathbf{C}^{\mathcal{A}} = {}^{\mathcal{C}}\mathbf{A}^{\mathcal{B}} \cdot {}^{\mathcal{B}}\mathbf{B}^{\mathcal{A}}$ (zde by ještě bylo potřeba si uvědomit, že pokud pro nějaké dvě matice $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$ je $\mathbf{A}\boldsymbol{x} = \mathbf{B}\boldsymbol{x}$ pro každé $\boldsymbol{x} \in \mathbb{C}^n$, pak je nutně $\mathbf{A} = \mathbf{B}$ — k tomu si např. stačí rozmyslet, jak vypadá součin $\mathbf{A}\boldsymbol{x}$ když $\boldsymbol{x}$ je vektor standardní báze).

Přenecháváme čtenáři aby si podrobně rozmyslel body (d) a (e). V souvislosti s bodem (e) si také zkuste rozmyslet, že lineární zobrazení $A \in \mathscr{L}(P)$ je prosté právě když je na nebo také právě když $A(\boldsymbol{x}) = \mathbf{0}$ pouze pro $\boldsymbol{x} = \mathbf{0}$.

1.3.6. V souvislosti s dalším nás bude především zajímat tzv. matice přechodu a s ní související pojem podobnosti matic. Mnohé lineární geometrické transformace (např. v počítačové grafice) odpovídají zobrazení jednoho předmětu v různých souřadných soustavách, což odpovídá změně báze (kromě posunutí). Buď opět $P \subset\subset \mathbb{C}^n$ netriviální podprostor (v jednoduchých geometrických aplikacích se často uvažuje reálný případ a $P = \mathbb{R}^2$ nebo $P = \mathbb{R}^3$, v jiných aplikacích je ale užitečné uvažovat obecný případ reálný nebo dokonce i komplexní). Nechť

$$\mathcal{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}, \qquad \mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k\}$$

jsou nějaké dvě báze P. Je-li $A \in \mathscr{L}(P)$ nějaké lineární zobrazení, pak můžeme uvažovat dvě matice tohoto zobrazení — vzhledem k bázi $\mathcal{A}$ a vzhledem k bázi $\mathcal{B}$, tj. matice ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$ a ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}}$. Vzniká otázka, zda z ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$ lze nějakým způsobem získat ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}}$ a obráceně.

V předchozím jsme říkali, že v případě lineárního operátoru z $\mathscr{L}(P)$ uvažujeme většinou matici zobrazení vzhledem k jedné bázi. Pro identický operátor I uvažujme nyní matici tohoto zobrazení vzhledem ke dvěma bázím $\mathcal{A}, \mathcal{B}$. Podle předchozího značení je [viz (1.23)]

$$^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} = \Big(\ (\boldsymbol{a}_1)_{\mathcal{B}}\ \big|\ (\boldsymbol{a}_2)_{\mathcal{B}}\ \big|\ \ldots\ \big|\ (\boldsymbol{a}_k)_{\mathcal{B}}\ \Big)$$

a podobně

$$^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} = \Big(\ (\boldsymbol{b}_1)_{\mathcal{A}}\ \big|\ (\boldsymbol{b}_2)_{\mathcal{A}}\ \big|\ \ldots\ \big|\ (\boldsymbol{b}_k)_{\mathcal{A}}\ \Big)$$

Matici ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}$ říkáme *matice přechodu od báze* $\mathcal{A}$ *k bázi* $\mathcal{B}$ (a matici ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$ je matice přechodu od $\mathcal{B}$ k $\mathcal{A}$). Všimněme si nejprve, že je

$$^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} = \mathbf{E}, \tag{1.27}$$

tj. matice ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}$ a ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$ jsou navzájem inverzní; odtud je také vidět, že matice přechodu je vždy regulární (zkuste ověřit přímo). Použijeme-li rovnost (1.25), pak pro $\boldsymbol{x} \in P$ dostáváme

$$(\boldsymbol{x})_{\mathcal{B}} = \big(I(\boldsymbol{x})\big)_{\mathcal{B}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot \big(\boldsymbol{x}\big)_{\mathcal{A}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot \big(I(\boldsymbol{x})\big)_{\mathcal{A}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} \cdot (\boldsymbol{x})_{\mathcal{B}},$$

tj. $({}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}) \cdot (\boldsymbol{x})_{\mathcal{B}} = (\boldsymbol{x})_{\mathcal{B}}$ pro každé $\boldsymbol{x} \in P$, tj. platí (1.27).

Nyní si všimněme vztahu mezi maticemi zobrazení ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$ a ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}}$. Opět stačí použít rovnost (1.25). Pro $\boldsymbol{x} \in P$ máme

$$\begin{aligned}{}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} \cdot (\boldsymbol{x})_{\mathcal{B}} = \big(A(\boldsymbol{x})\big)_{\mathcal{B}} = \big(I\big(A(\boldsymbol{x})\big)\big)_{\mathcal{B}} &= {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot \big(A(\boldsymbol{x})\big)_{\mathcal{A}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot (\boldsymbol{x})_{\mathcal{A}} = \\ &= {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot \big(I(\boldsymbol{x})\big)_{\mathcal{A}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} \cdot (\boldsymbol{x})_{\mathcal{B}},\end{aligned}$$

odtud vidíme, že

$${}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}.$$

Jelikož matice ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}$ a ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$ jsou navzájem inverzní, můžeme tuto rovnost psát např. ve tvaru

$$(1.28) \qquad {}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} = \big({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}\big)^{-1} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}.$$

Každá matice přechodu je podle předchozího regulární, ale není těžké si uvědomit, že naopak každá regulární matice je maticí přechodu mezi nějakými dvěma bázemi. Buď $\mathcal{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$ nějaká báze P a buď $\mathbf{B} = (b_{ij}) \in \mathbb{C}^{k\times k}$ regulární matice. Pro $j = 1, 2, \ldots, k$ označme

$$(*) \qquad \boldsymbol{b}_j = \sum_{i=1}^{k} b_{ij}\boldsymbol{a}_i.$$

Vzhledem k regularitě $\mathbf{B}$ je vidět, že vektory $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k$ jsou lineárně nezávislé (rozmyslete si) a $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k\}$ je báze P. Rovnost $(*)$ znamená, že $(\boldsymbol{b}_j)_{\mathcal{A}} = \mathbf{B}(:, j)$ a podle definice matice přechodu je tedy $\mathbf{B} = {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$. Opravdu je tedy každá regulární matice nějakou maticí přechodu. Na základě této skutečnosti dostáváme rovnost (když $\mathcal{A}$ je stále nějaká pevně daná báze $P \subset\subset \mathbb{C}^n$, $\dim P = k$, $A \in \mathscr{L}(P)$, $\mathbf{A} = {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$)

$$(1.29) \qquad \big\{{}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{B}^{\mathcal{B}} \;\big|\; \mathcal{B} \text{ je báze } P\big\} = \big\{\mathbf{X}^{-1}\mathbf{A}\mathbf{X} \;\big|\; \mathbf{X} \in \mathbb{C}^{k\times k} \text{ je regulární}\big\}.$$

Podobně v reálném případě. Vidíme, že je-li $\mathbf{A}$ čtvercová matice, pak

> *všechny matice tvaru* $\mathbf{X}^{-1}\mathbf{A}\mathbf{X}$ *reprezentují tentýž lineární operátor, pouze vyjádřený v různých bázích.*

Odtud vychází následující definice.

1.3.7 Definice. *Řekneme, že matice* $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ *jsou* ***podobné*** *a píšeme* $\mathbf{A} \sim \mathbf{B}$, *jestliže existuje regulární matice* $\mathbf{X}$ *řádu* n *tak že*

$$(1.30) \qquad \mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}.$$

Rovnost (1.30) *se nazývá* ***podobnostní transformace*** *a matici* $\mathbf{X}$ *se říká* ***transformační matice*** *(nebo* ***transformující matice****).*

1.3.8 Poznámka. V rovnosti (1.30) uvažujeme matici $\mathbf{X}$ obecně komplexní. Vzniká přirozená otázka, zda v případě, že $\mathbf{A}, \mathbf{B}$ jsou reálné, lze uvažovat matici $\mathbf{X}$ také jenom reálnou. V této souvislosti na chvíli zaveďme pojem podobnosti v reálném nebo v komplexním oboru. Řekneme, že matice $\mathbf{A}, \mathbf{B}$ jsou *podobné v reálném oboru* nebo *podobné v komplexním oboru*, pokud transformující matice $\mathbf{X}$ v (1.30) je reálná nebo komplexní. Jsou-li matice podobné v reálném oboru, pak jsou samozřejmě podobné i v oboru komplexním, neboť reálná matice je speciální případ matice komplexní. Jsou-li $\mathbf{A}, \mathbf{B}$ obě reálné, pak ale platí i obrácená implikace, jak plyne z následujícího tvrzení, u kterého uvádíme i zajímavý důkaz. Toto tvrzení dává kladnou odpověď na výše položenou otázku, tj. v případě reálných matic $\mathbf{A}, \mathbf{B}$ se můžeme omezit na reálné transformační matice $\mathbf{X}$.

1.3.9 Věta. *Jsou-li $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ podobné v komplexním oboru, pak jsou také podobné v reálném oboru.*

Důkaz. Předpokládejme, že $\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$, kde $\mathbf{X} = (x_{ij}) \in \mathbb{C}^{n\times n}$ je nějaká regulární matice. Nechť $\mathbf{Y} = (y_{ij})$, $\mathbf{Z} = (z_{ij})$ jsou reálná a imaginární část matice $\mathbf{X}$, tj.

$$y_{ij} = \mathfrak{Re}\, x_{ij}, \qquad z_{ij} = \mathfrak{Im}\, x_{ij}$$

$(i, j = 1, 2, \ldots, n)$, zapsáno krátce, $\mathbf{Y} = \mathfrak{Re}\,\mathbf{X}$, $\mathbf{Z} = \mathfrak{Im}\,\mathbf{X}$. Je ovšem $\mathbf{Y}, \mathbf{Z} \in \mathbb{R}^{n\times n}$, $\mathbf{X} = \mathbf{Y} + \mathrm{j}\mathbf{Z}$. Rovnost $\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$ můžeme napsat ve tvaru

$$\mathbf{X}\mathbf{B} = \mathbf{A}\mathbf{X} \tag{1.31}$$

(vynásobíme původní rovnost maticí $\mathbf{X}$ zleva). Všimněme si, že je

$$\mathbf{X}\mathbf{B} = (\mathbf{Y} + \mathrm{j}\mathbf{Z})\mathbf{B} = \mathbf{Y}\mathbf{B} + \mathrm{j}\mathbf{Z}\mathbf{B}, \qquad \mathbf{A}\mathbf{X} = \mathbf{A}(\mathbf{Y} + \mathrm{j}\mathbf{Z}) = \mathbf{A}\mathbf{Y} + \mathrm{j}\mathbf{A}\mathbf{Z}.$$

Jelikož matice $\mathbf{Y}, \mathbf{Z}$ jsou reálné a podle předpokladu $\mathbf{A}, \mathbf{B}$ také, dostáváme odtud a z (1.31), že

$$\mathbf{Y}\mathbf{B} = \mathbf{A}\mathbf{Y}, \qquad \mathbf{Z}\mathbf{B} = \mathbf{A}\mathbf{Z}. \tag{1.32}$$

Kdyby alespoň jedna z matic $\mathbf{Y}, \mathbf{Z}$ byla regulární, pak by už přímo odtud bylo vidět, že $\mathbf{A}, \mathbf{B}$ jsou podobné v reálném oboru. Obecně ale $\mathbf{Y}, \mathbf{Z}$ mohou být obě singulární i když $\mathbf{X} = \mathbf{Y} + \mathrm{j}\mathbf{Z}$ je regulární (nalezněte jednoduchý příklad takových matic řádu 2).

Definujme funkci p komplexní proměnné x

$$p(x) = \det(\mathbf{Y} + x\mathbf{Z}).$$

Přímo z definice determinantu je vidět, že p je polynom. Jelikož $p(\mathrm{j}) = \det(\mathbf{Y} + \mathrm{j}\mathbf{Z}) = \det \mathbf{X} \neq 0$ ($\mathbf{X}$ je regulární), polynom p není identicky nulový a má tedy pouze konečně mnoho nulových bodů. Existuje proto $x_0 \in \mathbb{R}$ tak, že $p(x_0) \neq 0$, tj. $\det(\mathbf{Y} + x_0\mathbf{Z}) \neq 0$. Matice

$$\mathbf{T} = \mathbf{Y} + x_0\mathbf{Z}$$

je tedy regulární a přitom reálná. Z rovností (1.32) nyní dostáváme

$$\mathbf{T}\mathbf{B} = (\mathbf{Y} + x_0\mathbf{Z})\mathbf{B} = \mathbf{Y}\mathbf{B} + x_0\mathbf{Z}\mathbf{B} = \mathbf{A}\mathbf{Y} + x_0\mathbf{A}\mathbf{Z} = \mathbf{A}(\mathbf{Y} + x_0\mathbf{Z}) = \mathbf{A}\mathbf{T}.$$

Jelikož $\mathbf{T}$ je regulární, vynásobením této rovnosti maticí $\mathbf{T}^{-1}$ zleva dostáváme rovnost $\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ a vidíme, že opravdu $\mathbf{A}, \mathbf{B}$ jsou podobné v reálném oboru. $\square$

Všimněme si ještě, že podobnost matic je ekvivalence.

1.3.10 Tvrzení. *Buďte* $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{C}^{n\times n}$. *Potom*

(i) $\mathbf{A} \sim \mathbf{A}$,

(ii) $\mathbf{A} \sim \mathbf{B} \implies \mathbf{B} \sim \mathbf{A}$,

(iii) $\mathbf{A} \sim \mathbf{B}, \mathbf{B} \sim \mathbf{C} \implies \mathbf{A} \sim \mathbf{C}$.

Důkaz. Vlastnost (i) je zřejmá, neboť je jistě $\mathbf{A} = \mathbf{E}^{-1}\mathbf{A}\mathbf{E}$.

Předpokládejme, že $\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$. Potom

$$\left(\mathbf{X}^{-1}\right)^{-1}\mathbf{B}\mathbf{X}^{-1} = \mathbf{X}\mathbf{B}\mathbf{X}^{-1} = \mathbf{X}(\mathbf{X}^{-1}\mathbf{A}\mathbf{X})\mathbf{X}^{-1} = \mathbf{E}\mathbf{A}\mathbf{E} = \mathbf{A}$$

a vidíme, že platí (ii).

Dokažme bod (iii). Předpokládejme, že

$$\mathbf{B} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}, \qquad \mathbf{C} = \mathbf{Y}^{-1}\mathbf{B}\mathbf{Y}$$

pro nějaké regulární matice $\mathbf{X}, \mathbf{Y}$. S použitím tvrzení o matici inverzní k součinu matic (viz odstavec 1.2.11) dostaneme

$$\mathbf{C} = \mathbf{Y}^{-1}\mathbf{B}\mathbf{Y} = \mathbf{Y}^{-1}(\mathbf{X}^{-1}\mathbf{A}\mathbf{X})\mathbf{Y} = (\mathbf{Y}^{-1}\mathbf{X}^{-1})\mathbf{A}(\mathbf{X}\mathbf{Y}) = (\mathbf{X}\mathbf{Y})^{-1}\mathbf{A}(\mathbf{X}\mathbf{Y})$$

a tedy opravdu $\mathbf{A} \sim \mathbf{C}$. □

Jelikož podobnost matic je ekvivalence, lze $\mathbb{C}^{n\times n}$ (nebo $\mathbb{R}^{n\times n}$) rozložit na třídy podobných matic. V předchozím jsme si již všimli, že matice v každé této třídě vždy reprezentují tentýž lineární operátor vyjádřený v různých bázích, přičemž na transformující matici $\mathbf{X}$ se lze dívat jako na matici přechodu [viz rovnost (1.29)].

1.4 Soustavy lineárních rovnic

Řešení soustav lineárních rovnic není tématem tohoto textu, nicméně dále budeme potřebovat znát obecné vlastnosti soustav lineárních rovnic, zvláště podmínky řešitelnosti a pod. Připomeňme proto alespoň základní fakta.

Je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$, pak *soustavu lineárních (algebraických) rovnic s maticí* $\mathbf{A}$ *a pravou stranou* $\boldsymbol{b}$ můžeme zapsat ve tvaru

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1.33}$$

kde $\boldsymbol{x} \in \mathbb{C}^n$ je neznámá [že se jedná o soustavu m rovnic o n neznámých je zřejmé po rozepsání rovnosti (1.33) do jednotlivých složek, když si uvědomíme, že součin $\mathbf{A}\boldsymbol{x}$ má tvar (1.9)]. Je-li matice $\mathbf{A}$ reálná, mluvíme o *soustavě s reálnými koeficienty*. Je-li navíc i vektor pravých stran $\boldsymbol{b}$ reálný, mluvíme o *reálné soustavě*; v tomto případě většinou uvažujeme pouze reálná řešení (tj. $\boldsymbol{x} \in \mathbb{R}^n$). Je-li speciálně $\boldsymbol{b} = \mathbf{0}$, říkáme, že daná soustava je *homogenní*, jinak mluvíme o *nehomogenní* soustavě.

Je zřejmé, že homogenní soustava má vždy řešení, totiž *nulové řešení* $\boldsymbol{x} = \mathbf{0}$; tomuto řešení se také říká *triviální řešení*. Dále je jednoduše vidět, že množina všech řešení homogenní soustavy $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ je podprostor v $\mathbb{C}^n$ (je-li $\mathbf{A}$ typu $m \times n$) nebo v $\mathbb{R}^n$, je-li $\mathbf{A}$ reálná a uvažujeme-li pouze reálná řešení. Jsou-li totiž $\boldsymbol{x}, \boldsymbol{y}$ dvě řešení, tj. $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, $\mathbf{A}\boldsymbol{y} = \mathbf{0}$, pak je $\mathbf{A}(\boldsymbol{x} + \boldsymbol{y}) = \mathbf{A}\boldsymbol{x} + \mathbf{A}\boldsymbol{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, tj. součet dvou řešení dané soustavy je také řešení a podobně je vidět, že násobek řešení skalárem opět tuto soustavu řeší; množina všech řešení je tedy opravdu lineární podprostor. Dokonce je známo, že dimenze množiny všech řešení homogenní soustavy závisí pouze na počtu neznámých (tj. v našem případě na n) a na hodnosti matice této soustavy — platí následující tvrzení.

1.4.1 Věta. *Buď $\mathbf{A} \in \mathbb{C}^{m\times n}$ a buď V množina všech řešení homogenní soustavy $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Potom je*

$$V \neq \emptyset, \qquad V \subset\subset \mathbb{C}^{m\times n}, \qquad \dim V = n - \mathrm{hod}(\mathbf{A}).$$

Poznamenejme, že v reálném případě platí zcela analogické tvrzení pro řešení reálná:

1.4.2 Věta. *Je-li $\mathbf{A} \in \mathbb{R}^{m\times n}$, V množina všech **reálných** řešení homogenní soustavy $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, pak*

$$V \neq \emptyset, \qquad V \subset\subset \mathbb{R}^{n}, \qquad \dim V = n - \mathrm{hod}(\mathbf{A}).$$

Všimneme si, že pokud $\mathrm{hod}\,\mathbf{A} = n$, tj. pokud $m \geq n$ a $\mathbf{A}$ má plnou hodnost (viz odstavec 1.2.9), pak je $\dim V = 0$, tj. v tomto (a právě v tomto) případě homogenní soustava má pouze jediné řešení — řešení triviální. Jinak je $\dim V > 0$ a tedy homogenní soustava má řešení nekonečně mnoho.

Nehomogenní soustava obecně řešení mít nemusí. Je-li $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$, pak rovnost (1.33) můžeme napsat ve tvaru [viz rovnost (1.10)]

$$x_1\mathbf{A}(:,1) + x_2\mathbf{A}(:,2) + \cdots + x_n\mathbf{A}(:,n) = \boldsymbol{b}. \tag{1.34}$$

Odtud vidíme, že *soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má řešení právě když vektor pravých stran $\boldsymbol{b}$ je lineární kombinací sloupců matice* $\mathbf{A}$. V trochu jiném tvaru se tomuto tvrzení říká Frobeniova věta. Tzv. *rozšířenou matici soustavy* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ dostaneme tak, že ke sloupcům matice $\mathbf{A}$ přidáme sloupec $\boldsymbol{b}$; v blokovém zápise má rozšířená matice této soustavy tvar $\big(\mathbf{A} \mid \boldsymbol{b}\big)$.

1.4.3 Věta (Frobeniova věta)**.** *Buď $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$. Potom soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má řešení právě když*

$$\mathrm{hod}(\mathbf{A}) = \mathrm{hod}\Big(\big(\mathbf{A} \mid \boldsymbol{b}\big)\Big)$$

(tj. právě když hodnost matice soustavy a hodnost rozšířené matice soustavy jsou stejné).

Zcela analogické tvrzení platí v reálném případě, tj. pokud $\mathbf{A}, \boldsymbol{b}$ jsou reálné a pokud uvažujeme pouze reálná řešení.

1.4.4. Množina všech řešení nehomogenní soustavy už netvoří lineární prostor (např. součet dvou řešení nehomogenní soustavy už řešení této soustavy není — je-li $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, $\mathbf{A}\boldsymbol{y} = \boldsymbol{b}$, pak $\mathbf{A}(\boldsymbol{x} + \boldsymbol{y}) = \mathbf{A}\boldsymbol{x} + \mathbf{A}\boldsymbol{y} = \boldsymbol{b} + \boldsymbol{b} = 2\boldsymbol{b} \neq \boldsymbol{b}$ pokud $\boldsymbol{b} \neq \mathbf{0}$; podobně pro násobek řešení). Všimněme si ale, že množina všech řešení nehomogenní soustavy, pokud je neprázdná, je „*posunutý podprostor*“.

Uvažujme nehomogenní soustavu $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ (kde $\boldsymbol{b} \neq \mathbf{0}$). Tzv. *příslušnou homogenní soustavou* rozumíme homogenní soustavu se stejnou maticí $\mathbf{A}$, tj. soustavu $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Buď $\boldsymbol{x}$ řešení nehomogenní soustavy, $\boldsymbol{y}$ řešení příslušné homogenní soustavy, tj. $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, $\mathbf{A}\boldsymbol{y} = \mathbf{0}$. Potom součet $\boldsymbol{x} + \boldsymbol{y}$ opět řeší původní nehomogenní soustavu — je totiž

$$\mathbf{A}(\boldsymbol{x} + \boldsymbol{y}) = \mathbf{A}\boldsymbol{x} + \mathbf{A}\boldsymbol{y} = \boldsymbol{b} + \mathbf{0} = \boldsymbol{b}.$$

Je-li $\boldsymbol{x}$ řešení nehomogenní soustavy $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ pak je dále vidět (rozmyslete si), že *každé* řešení této soustavy lze napsat ve tvaru $\boldsymbol{x} + \boldsymbol{y}$, kde $\boldsymbol{y}$ je vhodné řešení příslušné homogenní soustavy. Platí tedy:

> *Buď $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$ a předpokládejme, že soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má řešení. Nechť M značí množinu všech řešení této soustavy, V množinu všech řešení příslušné homogenní soustavy. Potom pro každé $\boldsymbol{y} \in M$ je*
>
> $$M = \boldsymbol{y} + V = \{\boldsymbol{y} + \boldsymbol{x} \mid \boldsymbol{x} \in V\}. \tag{1.35}$$

Zcela analogicky v případě reálné soustavy a množiny všech reálných řešení.

Jak jsme si už uvědomili, je-li $\operatorname{hod}(\mathbf{A}) = n$, homogenní soustava $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ má pouze triviální řešení. Pokud nehomogenní soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má řešení, pak např. z (1.35) můžeme vidět, že v tomto případě má řešení jediné. Pokud je $m > n$, pak ani v případě $\operatorname{hod}(\mathbf{A}) = n$ nemusí mít nehomogenní soustava s maticí $\mathbf{A}$ řešení. Pro čtvercové matice, tj. v případě $m = n$, se ale už něco takového stát nemůže. Připomeňme následující důležité tvrzení týkající se řešitelnosti soustav se čtvercovou maticí.

1.4.5 Věta. *Buď $\mathbf{A} \in \mathbb{C}^{n\times n}$ (resp. $\mathbf{A} \in \mathbb{R}^{n\times n}$). Potom následující podmínky jsou ekvivalentní:*

(a) $\mathbf{A}$ *je regulární;*

(b) *řádky $\mathbf{A}$ jsou lineárně nezávislé;*

(c) *sloupce $\mathbf{A}$ jsou lineárně nezávislé;*

(d) $\det \mathbf{A} \neq 0$*;*

(e) *existuje $\mathbf{A}^{-1}$;*

(f) *pro každé $\boldsymbol{x} \in \mathbb{C}^n$ (resp. každé $\boldsymbol{x} \in \mathbb{R}^n$), $\boldsymbol{x} \neq \mathbf{0}$, je $\mathbf{A}\boldsymbol{x} \neq \mathbf{0}$;*

(g) *pro každé $\boldsymbol{b} \in \mathbb{C}^n$ (resp. každé $\boldsymbol{b} \in \mathbb{R}^n$) soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má řešení;*

(h) *pro každé $\boldsymbol{b} \in \mathbb{C}^n$ (resp. každé $\boldsymbol{b} \in \mathbb{R}^n$) soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má právě jedno řešení;*

(i) *pro každé $\boldsymbol{b} \in \mathbb{C}^n$ (resp. každé $\boldsymbol{b} \in \mathbb{R}^n$) soustava $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ má nejvýše jedno řešení.*

Připomeňme, že v případě soustavy se čtvercovou regulární maticí $\mathbf{A}$ můžeme (jediné) řešení soustavy $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ psát ve tvaru

$$\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}.$$

Toto vyjádření řešení je elegantní, ve skutečnosti se ale pro praktické hledání řešení, zvláště v případě větších soustav, příliš nehodí. Metodami řešení soustav lineárních rovnic se v tomto textu nebudeme zabývat.

1.5 Skalární součin a ortogonalita

Nebudeme zde mluvit o obecném pojmu skalárního součinu v obecném lineárním prostoru — pro naše účely se stačí omezit na tzv. *standardní skalární součin* v $\mathbb{C}^n$ nebo $\mathbb{R}^n$. Pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^T$, značíme skalární součin vektorů $\boldsymbol{x}, \boldsymbol{y}$ symbolem $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ a definujeme rovností

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i \overline{y}_i \tag{1.36}$$

(připomeňme, že $\overline{y}_i$ značí číslo komplexně sdružené k y_i). Pro označení skalárního součinu zde volíme symbol $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, nikoli $\boldsymbol{xy}$ nebo $\boldsymbol{x} \cdot \boldsymbol{y}$ proto, abychom odlišili skalární součin od maticového součinu — připomeňme, že na vektory z $\mathbb{C}^n$ se můžeme dívat jako na jednosloupcové matice. Poznamenejme, že pomocí maticového součinu můžeme skalární součin vyjádřit ve tvaru

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \cdot \overline{\boldsymbol{y}} = (x_1, x_2, \ldots, x_n) \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \vdots \\ \overline{y}_n \end{pmatrix} ; \tag{1.37}$$

vyjádření skalárního součinu v tomto tvaru budeme dále často používat.

Jelikož pro $z \in \mathbb{C}$ je $z\overline{z} = |z|^2 \geq 0$ (a $|z|^2 = 0$ pouze pro $z = 0$), vidíme, že pro $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$ je

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^T \cdot \overline{\boldsymbol{x}} = \sum_{i=1}^{n} x_i \overline{x}_i = \sum_{i=1}^{n} |x_i|^2 \geq 0, \tag{1.38}$$

přičemž $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ právě když $\boldsymbol{x} = \boldsymbol{0}$. Základní vlastnosti skalárního součinu v $\mathbb{C}^n$ jsou následující:

(i) *Pro každé* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ *je* $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$.

(ii) *Pro každé* $\boldsymbol{x} \in \mathbb{C}^n$ *je* $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ *a* $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ *právě když* $\boldsymbol{x} = \boldsymbol{0}$.

(iii) *Pro každé* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{C}^n$ *je*

$$\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle , \qquad \langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle .$$

(iv) *Pro každé* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$, $a \in \mathbb{C}$, *je*

$$\langle a\boldsymbol{x}, \boldsymbol{y} \rangle = a \langle \boldsymbol{x}, \boldsymbol{y} \rangle , \qquad \langle \boldsymbol{x}, a\boldsymbol{y} \rangle = \overline{a} \langle \boldsymbol{x}, \boldsymbol{y} \rangle .$$

Snadné ověření těchto vlastností přenecháváme čtenáři.

Jelikož pro reálné číslo a je $\overline{a} = a$, pro reálné vektory $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ je jednoduše

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i .$$

Právě s tímto tvarem skalárního součinu v $\mathbb{R}^n$ se čtenář pravděpodobně setkal v základním kurzu lineární algebry.

Pro $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$ definujeme *normu vektoru* $\boldsymbol{x}$ jako

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\boldsymbol{x}^T \cdot \overline{\boldsymbol{x}}} = \sqrt{\textstyle\sum_{i=1}^n |x_i|^2}. \tag{1.39}$$

Je-li $\boldsymbol{x} \in \mathbb{R}^n$, pak je ovšem

$$\|\boldsymbol{x}\| = \sqrt{\textstyle\sum_{i=1}^n x_i^2},$$

tj. v reálném případě se jedná o eukleidovskou normu.

Připomeňme, že pro $\boldsymbol{x} \in \mathbb{C}^n$ značíme $\boldsymbol{x}^* = \overline{\boldsymbol{x}}^T = \overline{\boldsymbol{x}^T}$. Můžeme si všimnout, že pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ je

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{(\boldsymbol{x}^* \boldsymbol{y})}. \tag{1.40}$$

Jelikož $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ je vždy reálné, je $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^* \boldsymbol{x}$ a můžeme také psát

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^* \boldsymbol{x}}. \tag{1.41}$$

Připomeňme dvě důležité vlastnosti normy odvozené od skalárního součinu.

1.5.1 Věta. *Pro každé $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ platí*

(a) ***Schwarzova nerovnost***

$$\big|\langle \boldsymbol{x}, \boldsymbol{y} \rangle\big| \le \|\boldsymbol{x}\| \|\boldsymbol{y}\|,$$

(b) ***trojúhelníková nerovnost***

$$\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|.$$

V dalším budeme potřebovat především pojem ortogonality (kolmosti) vektorů, který lze definovat právě pomocí skalárního součinu.

1.5.2 Definice. *Vektory $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ nazveme **ortogonální**, je-li $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. Jsou-li $\boldsymbol{x}, \boldsymbol{y}$ ortogonální, píšeme $\boldsymbol{x} \perp \boldsymbol{y}$.*

*Buď $k \in \mathbb{N}$, $k \ge 2$. Řekneme, že vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{C}^n$ jsou **ortogonální**, jestliže jsou **po dvou ortogonální**, tj. jestliže pro každé $i, j = 1, 2, \dots, k$, $i \ne j$, je $\boldsymbol{x}_i \perp \boldsymbol{x}_j$. Jestliže je navíc $\|\boldsymbol{x}_i\| = 1$ pro každé $i = 1, 2, \dots, k$, pak řekneme, že vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k$ jsou **ortonormální**.*

Vzhledem k rovnosti (1.40) vidíme, že

$$\boldsymbol{x} \perp \boldsymbol{y} \qquad \text{právě když} \qquad \boldsymbol{x}^* \boldsymbol{y} = 0. \tag{1.42}$$

Dále vidíme, že vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{C}^n$ jsou ortonormální právě když pro každé $i, j =$ $= 1, 2, \dots, k$ je

$$\boldsymbol{x}_i^* \boldsymbol{x}_j = \begin{cases} 0, & i \ne j, \\ 1, & i = j. \end{cases} \tag{1.43}$$

1.5.3 Poznámka. Pro ortogonální vektory platí Pythagorova věta, tj. pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$, $\boldsymbol{x} \perp \boldsymbol{y}$, je

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2. \tag{1.44}$$

Za předpokladu $\boldsymbol{x} \perp \boldsymbol{y}$ je totiž $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{y}, \boldsymbol{x}\rangle = 0$ a tedy opravdu

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle + \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{y}, \boldsymbol{x}\rangle + \langle \boldsymbol{y}, \boldsymbol{y}\rangle = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

V reálném případě má rovnost (1.44) zřejmý geometrický význam. Analogie této rovnosti platí i pro více než dva vektory. Indukcí lze ukázat, že jsou-li $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ ortogonální, pak

$$\left\|\sum_{i=1}^{k} \boldsymbol{x}_i\right\|^2 = \sum_{i=1}^{k} \|\boldsymbol{x}_i\|^2.$$

Důležitá vlastnost ortogonálních vektorů je lineární nezávislost.

1.5.4 Věta. *Ortogonální nenulové vektory jsou lineárně nezávislé.*
Speciálně: Ortonormální vektory jsou lineárně nezávislé.

Důkaz. Nechť $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ jsou ortogonální a nenulové a předpokládejme, že pro nějaké $a_1, a_2, \ldots, a_k \in \mathbb{C}$ je

$$\sum_{i=1}^{k} a_i \boldsymbol{x}_i = \mathbf{0}.$$

Potom pro každé $j = 1, 2, \ldots, k$ je

$$0 = \langle \boldsymbol{x}_j, \mathbf{0}\rangle = \left\langle \boldsymbol{x}_j, \sum_{i=1}^{k} a_i \boldsymbol{x}_i \right\rangle = \sum_{i=1}^{k} \overline{a}_i \langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle = \overline{a}_j \langle \boldsymbol{x}_j, \boldsymbol{x}_j\rangle = \overline{a}_j \|\boldsymbol{x}_j\|^2,$$

neboť pro $i \neq j$ je $\langle \boldsymbol{x}_j, \boldsymbol{x}_i\rangle = 0$. Jelikož $\boldsymbol{x}_j \neq \mathbf{0}$, je $\|\boldsymbol{x}_j\|^2 > 0$ a tedy $\overline{a}_j = 0$, tj. také $a_j = 0$. Vidíme, že pouze triviální lineární kombinace vektorů $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ má nulový součet, tj. tyto vektory jsou lineárně nezávislé.

Druhá část tvrzení plyne okamžitě z první, neboť ortonormální vektory jsou podle definice nenulové. □

V poznámce 1.1.10 jsme připomněli pojem standardní báze $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$, což byla báze $\mathbb{R}^n$ a zároveň báze $\mathbb{C}^n$. Jak je snadno vidět, při uvažovaném standardním skalárním součinu je tato báze ortonormální. V $\mathbb{R}^n$ a $\mathbb{C}^n$ tedy existuje ortonormální báze. Podle věty 1.1.6 každý netriviální podprostor v $\mathbb{R}^n$ nebo $\mathbb{C}^n$ má bázi. Vzniká otázka, zda každý netriviální podprostor v $\mathbb{R}^n$ nebo $\mathbb{C}^n$ má dokonce ortonormální bázi. Následující tvrzení dává kladnou odpověď. I když důkaz tohoto tvrzení není úplně krátký, uvádíme jej zde, neboť obsahuje algoritmus, pomocí kterého lze ortonormální bázi nalézt. Připomeňme, že zápis $P \subset\subset \mathbb{C}^n$ nebo $P \subset\subset \mathbb{R}^n$ říká, že P je podprostor v $\mathbb{C}^n$ nebo $\mathbb{R}^n$, a že symbolem $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k]$ značíme lineární obal vektorů $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$, přičemž by mělo být z kontextu zřejmé, zda se jedná o lineární obal v rámci prostoru $\mathbb{C}^n$ nebo $\mathbb{R}^n$ (viz odstavec 1.1.1).

1.5.5 Věta. *Nechť $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{C}^n$ jsou lineárně nezávislé ($k, n \in \mathbb{N}$, $k \leq n$). Potom existují ortonormální vektory $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k \in \mathbb{C}^n$ tak, že pro každé $r = 1, 2, \ldots, k$ je*

$$[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r] = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r]. \tag{1.45}$$

Důkaz. Vektory $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ budeme konstruovat postupně. Postup (algoritmus), který přitom použijeme, se nazývá *Grammův-Schmidtův ortogonalizační proces* nebo *Grammova-Schmidtova ortogonalizace.*

Nejprve položíme

$$\boldsymbol{y}_1 = \frac{1}{\|\boldsymbol{x}_1\|}\boldsymbol{x}_1 \tag{1.46}$$

(je určitě $\|\boldsymbol{x}_1\| \neq 0$, neboť podle předpokladu jsou vektory $\boldsymbol{x}_i$ lineárně nezávislé a tedy nenulové). Potom je

$$\|\boldsymbol{y}_1\| = \left\|\frac{1}{\|\boldsymbol{x}_1\|}\boldsymbol{x}_1\right\| = \frac{1}{\|\boldsymbol{x}_1\|}\|\boldsymbol{x}_1\| = 1$$

a zároveň zřejmě $[\boldsymbol{y}_1] = [\boldsymbol{x}_1]$ (rozmyslete si znova, jak vypadá lineární obal jednoho vektoru).

Nyní hledáme vektor $\boldsymbol{y}_2$ s požadovanou vlastností. Nejprve nalezneme nějaký vektor $\boldsymbol{y}_2'$ takový, že $\boldsymbol{y}_2' \perp \boldsymbol{y}_1$ a přitom $[\boldsymbol{y}_1, \boldsymbol{y}_2'] = [\boldsymbol{x}_1, \boldsymbol{x}_2]$. Vektor $\boldsymbol{y}_2'$ budeme hledat ve tvaru

$$\boldsymbol{y}_2' = \boldsymbol{x}_2 - a_1\boldsymbol{y}_1, \tag{1.47}$$

kde $a_1 \in \mathbb{C}$ je vhodná konstanta. Tuto konstantu určíme z podmínky $\boldsymbol{y}_2' \perp \boldsymbol{y}_1$, tj. z podmínky [viz (1.42)]

$$0 = \boldsymbol{y}_1^*\boldsymbol{y}_2' = \boldsymbol{y}_1^*(\boldsymbol{x}_2 - a_1\boldsymbol{y}_1) = \boldsymbol{y}_1^*\boldsymbol{x}_2 - a_1\boldsymbol{y}_1^*\boldsymbol{y}_1 = \boldsymbol{y}_1^*\boldsymbol{x}_2 - a_1$$

(je $\boldsymbol{y}_1^*\boldsymbol{y}_1 = \|\boldsymbol{y}_1\|^2 = 1$). Podmínka $\boldsymbol{y}_2' \perp \boldsymbol{y}_1$ je tedy splněna *právě když* volíme

$$a_1 = \boldsymbol{y}_1^*\boldsymbol{x}_2. \tag{1.48}$$

Ukažme, že je $[\boldsymbol{y}_1, \boldsymbol{y}_2'] = [\boldsymbol{x}_1, \boldsymbol{x}_2]$. Připomeňme známé tvrzení, že jsou-li $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k$ nějaké vektory (v $\mathbb{C}^n$ nebo $\mathbb{R}^n$) a přičteme-li k nějakému z těchto vektorů lineární kombinaci vektorů ostatních, lineární obal těchto vektorů se nezmění — např. je tedy vždy

$$[\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k] = \left[\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_{k-1}, \boldsymbol{q}_k + \sum_{i=1}^{k-1}\lambda_i\boldsymbol{q}_i\right],$$

kde $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ jsou libovolné skaláry (z $\mathbb{C}$ nebo $\mathbb{R}$; rozmyslete si snadný důkaz tohoto tvrzení). Jelikož $\boldsymbol{y}_2'$ má tvar (1.47), vidíme odtud okamžitě, že je

$$[\boldsymbol{y}_1, \boldsymbol{y}_2'] = [\boldsymbol{y}_1, \boldsymbol{x}_2] = [\boldsymbol{x}_1, \boldsymbol{x}_2]$$

(rozmyslete si ještě, proč je $[\boldsymbol{y}_1, \boldsymbol{x}_2] = [\boldsymbol{x}_1, \boldsymbol{x}_2]$). Podle volby a_1 jsou vektory $\boldsymbol{y}_1, \boldsymbol{y}_2'$ ortogonální, ale nemusí být ortonormální, neboť může být $\|\boldsymbol{y}_2^*\| \neq 1$. Je ale jistě $\boldsymbol{y}_2' \neq \boldsymbol{0}$ — to je vidět např. přímo z (1.47), neboť $\boldsymbol{x}_1, \boldsymbol{x}_2$ jsou lineárně nezávislé a $\boldsymbol{y}_1$ je nenulový násobek $\boldsymbol{x}_1$ (jinak řečeno, $\boldsymbol{y}_2'$ je netriviální lineární kombinace lineárně nezávislých vektorů — koeficient u $\boldsymbol{x}_2$ je roven jedné). Nyní stačí položit

$$\boldsymbol{y}_2 = \frac{1}{\|\boldsymbol{y}_2'\|}\boldsymbol{y}_2'.$$

Potom už jsou vektory $\boldsymbol{y}_1, \boldsymbol{y}_2$ ortonormální a přitom $[\boldsymbol{y}_1, \boldsymbol{y}_2] = [\boldsymbol{x}_1, \boldsymbol{x}_2]$.

Dále v důkazu postupujeme indukcí. Předpokládejme, že pro nějaké r přirozené, $1 \leq \leq r < k$, už máme sestrojeny ortonormální vektory $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r$ takové, že pro každé $s = 1, 2, \ldots, r$ je

$$[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_s] = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_s].$$

Podobně jako jsme sestrojili vektory $\boldsymbol{y}_2', \boldsymbol{y}_2$, sestrojíme nyní vektory $\boldsymbol{y}_{r+1}', \boldsymbol{y}_{r+1}$. Uvažujme $\boldsymbol{y}_{r+1}'$ ve tvaru

$$\boldsymbol{y}_{r+1}' = \boldsymbol{x}_{r+1} - \sum_{i=1}^{r} a_i \boldsymbol{y}_i, \tag{1.49}$$

kde $a_i \in \mathbb{C}$. Konstanty a_i opět určíme z podmínek kolmosti, tj. z podmínek

$$\boldsymbol{y}_{r+1}' \perp \boldsymbol{y}_j \qquad \text{pro } j = 1, 2, \ldots, r,$$

které lze napsat ve tvaru

$$0 = \boldsymbol{y}_j^* \boldsymbol{y}_{r+1}' = \boldsymbol{y}_j^* \left(\boldsymbol{x}_{r+1} - \sum_{i=1}^{r} a_i \boldsymbol{y}_i \right) = \boldsymbol{y}_j^* \boldsymbol{x}_{r+1} - \sum_{i=1}^{r} a_i \boldsymbol{y}_j^* \boldsymbol{y}_i = \boldsymbol{y}_j^* \boldsymbol{x}_{r+1} - a_j$$

(podle předpokladu pro $i \neq j$ je $\boldsymbol{y}_j^* \boldsymbol{y}_i = 0$ a dále $\boldsymbol{y}_j^* \boldsymbol{y}_j = 1$). Podmínky kolmosti budou tedy splněny právě když volíme

$$a_j = \boldsymbol{y}_j^* \boldsymbol{x}_{r+1} \qquad \text{pro } j = 1, 2, \ldots, r. \tag{1.50}$$

Pokud vektor $\boldsymbol{y}_{r+1}'$ je zkonstruován tímto způsobem, pak vektory

$$\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r, \boldsymbol{y}_{r+1}'$$

jsou ortogonální. Přitom s použitím výše uvedeného tvrzení o zachování lineárního obalu můžeme vidět, že je

$$[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r, \boldsymbol{y}_{r+1}'] = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{r+1}].$$

Odtud dále vidíme, že je $\boldsymbol{y}_{r+1}' \neq \boldsymbol{0}$ — např. tak, že si uvědomíme, že je

$$\dim[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r, \boldsymbol{y}_{r+1}'] = \dim[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{r+1}] = r + 1$$

(podle předpokladu jsou $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ lineárně nezávislé). Položíme-li nakonec

$$\boldsymbol{y}_{r+1} = \frac{1}{\|\boldsymbol{y}_{r+1}'\|}\boldsymbol{y}_{r+1}', \tag{1.51}$$

vidíme, že $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{r+1}$ mají všechny požadované vlastnosti. $\square$

1.5.6 Poznámka. Shrňme, že realizace Grammovy-Schmidtovy ortogonalizace spočívá v postupném použití rovností (1.46), (1.50), (1.49) a (1.51). Uvažovali jsme přitom obecně komplexní případ. Jsou-li vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ reálné, pak je ale vidět, že konstanty a_i a postupně konstruované vektory $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ jsou také reálné. Místo $\boldsymbol{y}_j^*$ můžeme potom všude jednoduše psát $\boldsymbol{y}_j^T$. Příslušné lineární obaly potom můžeme uvažovat v rámci $\mathbb{R}^n$ — věta 1.5.5 tak zůstane v platnosti když $\mathbb{C}^n$ nahradíme prostorem $\mathbb{R}^n$.

1.5.7 Poznámka. Podle věty 1.1.6 každý netriviální podprostor v $\mathbb{C}^n$ nebo $\mathbb{R}^n$ má bázi. Podle věty 1.5.5 můžeme tuto bázi pomocí Grammova-Schmidtova ortogonalizačního procesu zortogonalizovat a dostaneme tak ortonormální bázi tohoto podprostoru. Krátce řečeno, *každý netriviální podprostor v $\mathbb{C}^n$ nebo $\mathbb{R}^n$ má ortonormální bázi.* Jelikož jeden podprostor má různé báze, ortogonalizací můžeme dostat různé ortonormální báze téhož podprostoru (s výjimkou jednodimenzionálního podprostoru nekonečně mnoho). Např. tedy standardní báze ani zdaleka není jediná ortonormální báze $\mathbb{C}^n$ nebo $\mathbb{R}^n$.

1.5.8 Poznámka. V bodu (c) věty 1.1.8 se mluví o doplnění lineárně nezávislých vektorů na bázi. V dalším budeme potřebovat vědět, že podobně lze ortonormální vektory doplnit na ortonormální bázi. Buď $P \subset\subset \mathbb{C}^n$ (nebo $P \subset\subset \mathbb{R}^n$), $\dim P = k > 0$, a nechť $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r \in P$ jsou nějaké ortonormální vektory ($1 \le r < k$). Podle věty 1.5.4 jsou tyto vektory lineárně nezávislé, takže podle bodu (c) věty 1.1.8 existují $\boldsymbol{x}_{r+1}, \boldsymbol{x}_{r+2}, \ldots, \boldsymbol{x}_k \in P$ tak, že vektory

$$\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r, \boldsymbol{x}_{r+1}, \boldsymbol{x}_{r+2}, \ldots, \boldsymbol{x}_k$$

tvoří bázi P. Tyto vektory lze ortogonalizovat pomocí Grammovy-Schmidtovy ortogonalizace a dostaneme ortonormální bázi P. Přitom je z konstrukce vidět, že v této ortogonalizaci vektory $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_r$ zůstanou nezměněny [viz rovnosti (1.50) a (1.49)]. Příslušné tvrzení o doplnění ortonormálních vektorů na ortonormální bázi můžeme tedy formulovat např. v následujícím tvaru:

> *Buď $P \subset\subset \mathbb{C}^n$ (resp. $P \subset\subset \mathbb{R}^n$),* $\dim P = k > 1$*. Buď dále r přirozené, $1 \le$ $\le r < k$ a nechť vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r \in P$ jsou ortonormální. Potom existují vektory $\boldsymbol{x}_{r+1}, \boldsymbol{x}_{r+2}, \ldots, \boldsymbol{x}_k \in P$ tak, že*
>
> $$\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_r, \boldsymbol{x}_{r+1}, \boldsymbol{x}_{r+2}, \ldots, \boldsymbol{x}_k$$
>
> *jsou ortonormální.*

Později toto tvrzení použijeme zvláště v případě, kdy $k = n$, tj. kdy je $P = \mathbb{C}^n$ (nebo $P = \mathbb{R}^n$).

1.5.9. Připomeňme, že pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ je

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \overline{\boldsymbol{y}} = \overline{(\boldsymbol{x}^* \boldsymbol{y})}$$

[viz (1.37), (1.40)]. Buď nyní $\mathbf{A} \in \mathbb{C}^{m \times n}$, $\boldsymbol{x} \in \mathbb{C}^n$, $\boldsymbol{y} \in \mathbb{C}^m$. Potom je $\mathbf{A}\boldsymbol{x} \in \mathbb{C}^m$, skalární součin $\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{y} \rangle$ je definovaný a podle právě uvedených rovností je (když ještě použijeme tvrzení 1.2.8)

$$\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{y} \rangle = (\mathbf{A}\boldsymbol{x})^T \overline{\boldsymbol{y}} = \boldsymbol{x}^T \mathbf{A}^T \overline{\boldsymbol{y}} = \overline{\big(\boldsymbol{x}^* (\mathbf{A}^* \boldsymbol{y})\big)} = \langle \boldsymbol{x}, \mathbf{A}^* \boldsymbol{y} \rangle,$$

krátce

$$\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathbf{A}^* \boldsymbol{y} \rangle \tag{1.52}$$

kdykoli $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\boldsymbol{x} \in \mathbb{C}^n$, $\boldsymbol{y} \in \mathbb{C}^m$. Je-li matice $\mathbf{A}$ reálná, pak je $\mathbf{A}^* = \mathbf{A}^T$ a tuto rovnost můžeme psát ve tvaru

$$\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \left\langle \boldsymbol{x}, \mathbf{A}^T \boldsymbol{y} \right\rangle . \tag{1.53}$$

Připomeňme, že pokud pro $\mathbf{A} \in \mathbb{C}^{n\times n}$ je $\mathbf{A}^* = \mathbf{A}$, pak říkáme, že matice $\mathbf{A}$ je hermitovská a že reálná hermitovská matice je symetrická. Je-li nyní v předchozím $m = n$ a $\mathbf{A}$ je hermitovská, je tedy $\mathbf{A}^* = \mathbf{A}$ a podle (1.52) je

$$\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathbf{A}\boldsymbol{y} \rangle$$

pro každé $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$; podobně je-li $\mathbf{A}$ reálná symetrická matice. V dalším bude užitečné vědět, že tato rovnost charakterizuje hermitovské (nebo reálné symetrické) matice, tj. že platí následující tvrzení.

1.5.10 Věta. *Buď $\mathbf{A} \in \mathbb{C}^{n\times n}$. Potom $\mathbf{A}$ je hermitovská právě když pro každé $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ je*

$$\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathbf{A}\boldsymbol{y} \rangle . \tag{1.54}$$

Je-li $\mathbf{A}$ reálná, pak $\mathbf{A}$ je symetrická právě když (1.54) platí pro každé $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

Důkaz. Právě jsme si všimli, že je-li $\mathbf{A}$ hermitovská, pak (1.54) platí pro každé $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$. Stačí tedy dokázat obrácenou implikaci.

Předpokládejme, že (1.54) platí pro každé $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$. Nechť $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ jsou vektory standardní báze v $\mathbb{C}^n$ (viz poznámku 1.1.10). Snadno je vidět, že pro $i = 1, 2, \ldots, n$ je $\mathbf{A}\boldsymbol{e}_i = \mathbf{A}(:, i)$ (rozepište si). Pro $i, j = 1, 2, \ldots, n$ tedy platí [když $\mathbf{A} = (a_{ij})$]

$$\langle \mathbf{A}\boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \left\langle \mathbf{A}(:, i), \boldsymbol{e}_j \right\rangle = \mathbf{A}(:, i)^T \cdot \boldsymbol{e}_j = a_{ji} \tag{$*$}$$

(je $\overline{\boldsymbol{e}}_j = \boldsymbol{e}_j$) a podobně

$$\langle \boldsymbol{e}_i, \mathbf{A}\boldsymbol{e}_j \rangle = \left\langle \boldsymbol{e}_i, \mathbf{A}(:, j) \right\rangle = \boldsymbol{e}_i^T\, \overline{\mathbf{A}(:, j)} = \overline{a}_{ij}. \tag{$**$}$$

Z (1.54) pro $\boldsymbol{x} = \boldsymbol{e}_i$, $\boldsymbol{y} = \boldsymbol{e}_j$, a z rovností ($*$), ($**$) okamžitě dostáváme, že $\overline{a}_{ij} = a_{ji}$, tj. $\mathbf{A}^* = \mathbf{A}$, tj. $\mathbf{A}$ je hermitovská. Jelikož vektory $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ jsou reálné, dostáváme odtud i druhou část tvrzení. □

V různých aplikacích se často pracuje s tzv. ortonormálními maticemi; v komplexním případě s tzv. maticemi unitárními.

1.5.11 Definice. *Buď $\mathbf{A} \in \mathbb{C}^{n\times n}$. Řekneme, že $\mathbf{A}$ je **unitární matice**, jestliže řádkové vektory $\mathbf{A}(1, :)^T, \mathbf{A}(2, :)^T, \ldots, \mathbf{A}(n, :)^T$ této matice jsou ortonormální (jakožto vektory z $\mathbb{C}^n$). Reálnou unitární matici také nazýváme **matice ortonormální**.*

Poznamenejme, že v reálném případě se někdy místo ortonormální matice používá název matice ortogonální. Základní vlastnosti unitárních matic jsou shrnuty v následujícím tvrzení.

1.5.12 Věta. *Buď* $\mathbf{A} \in \mathbb{C}^{n\times n}$. *Potom následující podmínky jsou ekvivalentní:*

(i) $\mathbf{A}$ *je unitární,*

(ii) $\mathbf{A}^* = \mathbf{A}^{-1}$,

(iii) *sloupcové vektory* $\mathbf{A}(:,1), \mathbf{A}(:,2), \dots, \mathbf{A}(:,n)$ *matice* $\mathbf{A}$ *jsou ortonormální,*

(iv) *pro každé* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ *je* $\langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$,

(v) *pro každé* $\boldsymbol{x} \in \mathbb{C}^n$ *je* $\|\mathbf{A}\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

DŮKAZ. V odstavci 1.2.11 jsme si uvědomili, že k ověření, že dvě čtvercové matice $\mathbf{A}, \mathbf{B}$ jsou navzájem inverzní, stačí ověřit např., že $\mathbf{AB} = \mathbf{E}$. Podle definice násobení matic je

$$(\mathbf{AA}^*)(i,j) = \mathbf{A}(i,:)\mathbf{A}^*(:,j)$$

[viz rovnost (1.7)]. Přitom je $\mathbf{A}^*(:,j) = \overline{\mathbf{A}(j,:)}$, tj.

$$(*) \qquad (\mathbf{AA}^*)(i,j) = \mathbf{A}(i,:)\overline{\mathbf{A}(j,:)} = \Big\langle \big(\mathbf{A}(i,:)\big)^T, \big(\mathbf{A}(j,:)\big)^T \Big\rangle.$$

Jelikož $\mathbf{AA}^* = \mathbf{E}$ právě když $(\mathbf{AA}^*)(i,j) = 0$ pro $i \neq j$ a $(\mathbf{AA}^*)(i,i) = 1$, z $(*)$ vidíme, že $\mathbf{A}^* = \mathbf{A}^{-1}$ právě když řádkové vektory matice $\mathbf{A}$ jsou ortonormální. Dokázali jsme ekvivalenci (i) $\iff$ (ii).

Zcela analogicky je vidět, že $\mathbf{A}^*\mathbf{A} = \mathbf{E}$ právě když sloupcové vektory $\mathbf{A}$ jsou ortonormální, tj. platí (ii) $\iff$ (iii).

Předpokládejme, že je $\mathbf{A}^* = \mathbf{A}^{-1}$. Pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ potom s použitím rovnosti (1.52) dostáváme

$$\langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \mathbf{A}^*\mathbf{A}\boldsymbol{y}\rangle = \big\langle \boldsymbol{x}, \mathbf{A}^{-1}\mathbf{A}\boldsymbol{y}\big\rangle = \langle \boldsymbol{x}, \mathbf{E}\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle\,.$$

Platí tedy (ii) $\implies$ (iv).

Nyní předpokládejme, že je splněna podmínka (iv) a nechť $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ je standardní báze $\mathbb{C}^n$ (viz poznámka 1.1.10). Potom pro $i = 1, 2, \dots, n$ je $\mathbf{A}\boldsymbol{e}_i = \mathbf{A}(:,i)$ a pro $i, j = 1, 2, \dots, n$ podle (iv) máme

$$\big\langle \mathbf{A}(:,i), \mathbf{A}(:,j)\big\rangle = \langle \mathbf{A}\boldsymbol{e}_i, \mathbf{A}\boldsymbol{e}_j\rangle = \langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Vidíme, že platí (iv) $\implies$ (iii).

Je-li splněna podmínka (iv), pak pro $\boldsymbol{x} \in \mathbb{C}^n$ je

$$\|\mathbf{A}\boldsymbol{x}\| = \sqrt{\langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle} = \|\boldsymbol{x}\|,$$

tj. platí také (v).

Nyní zbývá ukázat např. platnost implikace (v) $\implies$ (iv). Norma vektoru vyjádřena (definována) pomocí skalárního součinu. Ale skalární součin se dá naopak vyjádřit pomocí normy. Přenecháváme čtenáři aby ověřil (např. rozepsáním pravé strany), že pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ platí (připomeňme, že j značí imaginární jednotku)

$$(*) \qquad \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \frac{1}{4}\Big(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + \mathrm{j}\|\boldsymbol{x}+\mathrm{j}\boldsymbol{y}\|^2 - \mathrm{j}\|\boldsymbol{x}-\mathrm{j}\boldsymbol{y}\|^2 \Big)$$

[poznamenejme, že v reálném případě, tj. pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, je jednodušeji $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \frac{1}{4}(\|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x} - \boldsymbol{y}\|^2)$]. Platí-li (v), pak hodnota výrazu vpravo v (∗∗) se nezmění, když místo $\boldsymbol{x}, \boldsymbol{y}$ budeme všude psát $\mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y}$ [využijeme ještě např. toho, že $\mathbf{A}\boldsymbol{x} + \mathrm{j}\mathbf{A}\boldsymbol{y} = \mathbf{A}(\boldsymbol{x} + \mathrm{j}\boldsymbol{y})$] a z (∗∗) tedy vidíme, že platí $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y} \rangle$. Tvrzení je dokázáno. □

1.5.13. Pro případ reálných matic můžeme větu 1.5.12 vyslovit v následujícím tvaru; připomeňme, že reálnou unitární matici bývá zvykem nazývat matice ortonormální.

Buď $\mathbf{A} \in \mathbb{R}^{n\times n}$. Potom následující podmínky jsou ekvivalentní:

(i) $\mathbf{A}$ *je ortonormální,*

(ii) $\mathbf{A}^T = \mathbf{A}^{-1}$,

(iii) *sloupcové vektory* $\mathbf{A}(:,1), \mathbf{A}(:,2), \ldots, \mathbf{A}(:,n)$ *matice* $\mathbf{A}$ *jsou ortonormální,*

(iv) *pro každé* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *je* $\langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$,

(v) *pro každé* $\boldsymbol{x} \in \mathbb{R}^n$ *je* $\|\mathbf{A}\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

Věta 1.5.12 má mnohé zajímavé a důležité důsledky. Uveďme alespoň některé z nich.

1.5.14 Důsledek. *Je-li $\mathbf{A} \in \mathbb{C}^{n\times n}$ unitární matice, pak*

$$|\det \mathbf{A}| = 1.$$

DŮKAZ. Víme, že pro čtvercovou matici $\mathbf{A}$ je $\det(\mathbf{A}^T) = \det \mathbf{A}$. Není těžké si rozmyslet, že $\det \overline{\mathbf{A}} = \overline{\det \mathbf{A}}$. Odtud

$$\det \mathbf{A}^* = \det \overline{\mathbf{A}^T} = \overline{\det \mathbf{A}^T} = \overline{\det \mathbf{A}}.$$

Nyní vidíme, že je-li $\mathbf{A}$ unitární, pak

$$|\det \mathbf{A}|^2 = \det \mathbf{A}\, \overline{\det \mathbf{A}} = \det \mathbf{A} \det \mathbf{A}^* = \det \mathbf{A} \det \mathbf{A}^{-1} = \det(\mathbf{A}\mathbf{A}^{-1}) = \det \mathbf{E} = 1,$$

takže opravdu $|\det \mathbf{A}| = 1$. □

1.5.15 Důsledek. *Buď $\mathbf{A} \in \mathbb{C}^{n\times n}$ unitární a nechť vektory $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{C}^n$ jsou ortonormální. Potom také vektory $\mathbf{A}\boldsymbol{x}_1, \mathbf{A}\boldsymbol{x}_2, \ldots, \mathbf{A}\boldsymbol{x}_n$ jsou ortonormální.*

Toto tvrzení plyne okamžitě z vlastností (iv), (v) z věty 1.5.12.

1.5.16 Důsledek. *Součin unitárních matic (stejného řádu) je unitární.*

DŮKAZ. Buďte $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ unitární. Podle bodu (iii) věty 1.5.12 jsou sloupcové vektory $\mathbf{B}(:,1), \mathbf{B}(:,2), \ldots, \mathbf{B}(:,n)$ ortonormální a předchozího důsledku jsou ortonormální také vektory $\mathbf{A}\cdot\mathbf{B}(:,1), \mathbf{A}\cdot\mathbf{B}(:,2), \ldots, \mathbf{A}\cdot\mathbf{B}(:,n)$. Ale vektory $\mathbf{A}\cdot\mathbf{B}(:,j)$ nejsou nic jiného než sloupcové vektory součinu $\mathbf{AB}$ [viz rovnost (1.16)], takže podle věty 1.5.12 je opravdu matice $\mathbf{AB}$ unitární. □

1.5.17 Poznámka. V odstavci 1.3 jsme mluvili o lineárním zobrazení a matici lineárního zobrazení; v případě lineárního zobrazení nějakého lineárního prostoru do sebe jsme mluvili o lineárním operátoru. Připomeňme, že je-li $\mathbf{A} \in \mathbb{C}^{n\times n}$, pak zobrazení $A\colon \mathbb{C}^n \to \mathbb{C}^n$ definované rovností

$$A(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x} \tag{1.55}$$

pro $\boldsymbol{x} \in \mathbb{C}^n$, je lineární zobrazení na $\mathbb{C}^n$. Matice $\mathbf{A}$ je přitom maticí zobrazení A, přesněji matice zobrazení A vzhledem ke standardní bázi $\mathbb{C}^n$. Je-li $\mathbf{A}$ reálná, můžeme se přitom omezit na $\boldsymbol{x}$ reálné a na A se dívat jako na lineární zobrazení na $\mathbb{R}^n$. Víme, že naopak každé lineární zobrazení A na $\mathbb{C}^n$ (nebo na $\mathbb{R}^n$) lze napsat ve tvaru (1.55), kde $\mathbf{A}$ je vhodná čtvercová matice.

Je-li $A\colon \mathbb{C}^n \to \mathbb{C}^n$ (resp. $A\colon \mathbb{R}^n \to \mathbb{R}^n$) lineární zobrazení, pak A nazveme *izometrie* (*izometrické zobrazení*), jestliže A zachovává velikost vektorů, tj. jestliže pro každé $\boldsymbol{x} \in \mathbb{C}^n$ (resp. $\boldsymbol{x} \in \mathbb{R}^n$) je

$$\big\|A(\boldsymbol{x})\big\| = \|\boldsymbol{x}\|. \tag{1.56}$$

Poznamenejme, že izometrické lineární zobrazení zachovává vzdálenost. Pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ (resp. $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$) je totiž vzdálenost bodů $\boldsymbol{x}, \boldsymbol{y}$ rovna $\|\boldsymbol{x} - \boldsymbol{y}\|$ a je-li A izometrie, pak pro vzdálenost obrazů $A(\boldsymbol{x}), A(\boldsymbol{y})$ dostáváme

$$\big\|A(\boldsymbol{x}) - A(\boldsymbol{y})\big\| = \big\|A(\boldsymbol{x} - \boldsymbol{y})\big\| = \|\boldsymbol{x} - \boldsymbol{y}\|$$

Poznamenejme, že pojem izometrického zobrazení by bylo možné definovat o něco obecněji, zde však vystačíme s výše uvedenou definicí. Pro jistotu zde izometrii můžeme podrobněji říkat *lineární izometrie.* S reálnými lineárními izometrickými zobrazeními se můžeme setkat, kromě jiného, např. v počítačové grafice.

Má-li lineární zobrazení tvar (1.55), pak podle věty 1.5.12 můžeme snadno poznat, zda A je izometrie. Z toho, že ve větě 1.5.12 jsou podmínky (i) a (v) ekvivalentní, vidíme, že A je izometrie *právě když* matice $\mathbf{A}$ je unitární (v reálném případě ortonormální).

V této souvislosti je také zajímavá ekvivalence bodů (iv) a (v) ve větě 1.5.12, zvláště v reálném případě. Připomeňme, že pro $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{x}, \boldsymbol{y} \neq \mathbf{0}$, úhel α mezi těmito vektory je určen skalárním součinem $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ — je

$$\cos\alpha = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}.$$

Je-li A ve tvaru (1.55) izometrie na $\mathbb{R}^n$, pak podle předchozího je $\mathbf{A}$ reálná a ortonormální a podle věty 1.5.12 je $\langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$, $\|\mathbf{A}\boldsymbol{x}\| = \|\boldsymbol{x}\|$, $\|\mathbf{A}\boldsymbol{y}\| = \|\boldsymbol{y}\|$. Pro úhel β mezi obrazy $A(\boldsymbol{x}), A(\boldsymbol{y})$ tedy platí

$$\cos\beta = \frac{\langle \mathbf{A}\boldsymbol{x}, \mathbf{A}\boldsymbol{y} \rangle}{\|\mathbf{A}\boldsymbol{x}\|\|\mathbf{A}\boldsymbol{y}\|} = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = \cos\alpha.$$

Vidíme, že (reálná) *lineární izometrie zachovává úhly* mezi vektory.

1.5.18 Poznámka. V odstavci 1.3.6 jsme mluvili o matici přechodu a o vyjádření lineárního operátoru v různých bázích. Nechť

$$\mathcal{A} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$$

je standardní báze v $\mathbb{C}^n$ (resp. v $\mathbb{R}^n$) a

$$\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n\}$$

nějaká ortonormální báze v $\mathbb{C}^n$ (resp. v $\mathbb{R}^n$). Jelikož souřadnice vektoru v $\mathbb{C}^n$ vzhledem ke standardní bázi jsou rovny přímo příslušným složkám tohoto vektoru, vidíme, že matice přechodu ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$ od báze $\mathcal{B}$ k $\mathcal{A}$ má tvar

$${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} = \left(\ (\boldsymbol{b}_1)_{\mathcal{A}} \mid (\boldsymbol{b}_2)_{\mathcal{A}} \mid \ldots \mid (\boldsymbol{b}_n)_{\mathcal{A}}\ \right) = \left(\ \boldsymbol{b}_1 \mid \boldsymbol{b}_2 \mid \ldots \mid \boldsymbol{b}_n\ \right).$$

Jelikož báze $\mathcal{B}$ byla ortonormální, je matice ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$ unitární a, kromě jiného, $({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}})^{-1} = ({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}})^*$. Pro matici přechodu ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}$ dostáváme (viz odstavec 1.3.6)

$${}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} = \left({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}\right)^{-1} = \left(\ \boldsymbol{b}_1 \mid \boldsymbol{b}_2 \mid \ldots \mid \boldsymbol{b}_n\ \right)^* = \begin{pmatrix} \boldsymbol{b}_1^* \\ \boldsymbol{b}_2^* \\ \vdots \\ \boldsymbol{b}_n^* \end{pmatrix}$$

(v reálném případě uvažujeme pouze transpozici).

Buď dále A nějaká lineární izometrie v $\mathbb{C}^n$ (resp. v $\mathbb{R}^n$). Je-li ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$ matice zobrazení A vzhledem ke standardní bázi $\mathcal{A}$, pak podle předchozí poznámky je ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$ unitární. Je-li ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}}$ matice zobrazení A vzhledem k bázi $\mathcal{B}$, pak podle (1.28) a podle předchozího je

$$(1.57)\qquad {}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} = \left({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}\right)^{-1} \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} = \left({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}\right)^* \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}.$$

Jelikož matice ${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}}$ je unitární a matice ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}$ a $({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}})^*$ jsou unitární (rozmyslete si, že z věty 1.5.12 okamžitě plyne, že čtvercová matice $\mathbf{A}$ je unitární právě když $\mathbf{A}^*$ je unitární), podle důsledku 1.5.16 je také matice ${}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}}$ unitární. Vidíme, že matice lineární izometrie v $\mathbb{C}^n$ vzhledem ke každé ortonormální bázi je unitární (v reálném případě ortonormální).

Jak jsme viděli, matice přechodu od standardní báze k nějaké jiné ortonormální bázi (a obráceně) je unitární. Všimněme si ještě, že matice přechodu mezi dvěma libovolnými ortonormálními bázemi je unitární. Buď tedy $\mathcal{C}$ nějaká ortonormální báze $\mathbb{C}^n$ (resp. $\mathbb{R}^n$), ${}^{\mathcal{C}}\mathbf{A}^{\mathcal{C}}$ matice zobrazení A vzhledem k bázi $\mathcal{C}$. Podle předchozího je matice přechodu ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}$ od báze $\mathcal{C}$ k bázi $\mathcal{A}$ unitární (stejně jako v předchozím $\mathcal{A}$ značí standardní bázi, $\mathcal{B}$ nějakou bázi ortonormální). Rovnost (1.57) můžeme psát ve tvaru

$${}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} = {}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}} \cdot {}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} \cdot \left({}^{\mathcal{A}}\mathbf{I}^{\mathcal{B}}\right)^* = \left({}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}\right)^* \cdot {}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} \cdot {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}.$$

Odtud dostáváme

$${}^{\mathcal{C}}\mathbf{A}^{\mathcal{C}} = \left({}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}\right)^* \cdot {}^{\mathcal{A}}\mathbf{A}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}} = \left({}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}\right)^* \cdot \left({}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}\right)^* \cdot {}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} \cdot {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}} = \left({}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}\right)^* \cdot {}^{\mathcal{B}}\mathbf{A}^{\mathcal{B}} \cdot \left({}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}\right).$$

Tato rovnost znamená, že matice přechodu ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{C}}$ od báze $\mathcal{C}$ k bázi $\mathcal{B}$ má tvar

$${}^{\mathcal{B}}\mathbf{I}^{\mathcal{C}} = {}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}} \cdot {}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}.$$

Jelikož ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{A}}$ a ${}^{\mathcal{A}}\mathbf{I}^{\mathcal{C}}$ jsou unitární, podle důsledku 1.5.16 je opravdu matice ${}^{\mathcal{B}}\mathbf{I}^{\mathcal{C}}$ také unitární.

1.5.19 Poznámka. V odstavci 1.3 jsme definovali pojem podobnosti matic. Můžeme si všimnout, že rovnost (1.57) je speciální případ podobnostní transformace, kde transformující matice je unitární (zde se navíc jednalo o případ podobnosti mezi unitárními maticemi). Řekneme, že matice $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ jsou *unitárně podobné*, jestliže existuje unitární matice $\mathbf{X} \in \mathbb{C}^{n \times n}$ taková, že

$$\mathbf{B} = \mathbf{X}^* \mathbf{A} \mathbf{X}.$$

Podle předchozího jsou matice dané lineární izometrie vzhledem k různým ortonormálním bázím unitárně podobné.