

Kapitola 11

Vlastní čísla a vlastní vektory

Základní motivace pro studium vlastních čísel a vektorů pochází z teorie řešení diferenciálních rovnic. Tato teorie říká, že obecné řešení lineární diferenciální rovnice $u' = \lambda u$, kde $u(t)$ je neznámá reálná funkce reálné proměnné t , je tvaru $u(t) = \alpha e^{\lambda t}$, kde α je reálné číslo. Podobně soustava lineárních diferenciálních rovnic

$$\begin{aligned}u_1' &= 7u_1 - 4u_2 \\u_2' &= 5u_1 - 2u_2\end{aligned}$$

má řešení ve tvaru

$$u_1 = \alpha_1 e^{\lambda t}, \quad u_2 = \alpha_2 e^{\lambda t}$$

pro vhodná reálná čísla $\alpha_1, \alpha_2, \lambda$. Pokud tato vyjádření zderivujeme podle t a dosadíme do původní soustavy diferenciálních rovnic, dostaneme

$$\begin{aligned}\alpha_1 \lambda e^{\lambda t} &= 7\alpha_1 e^{\lambda t} - 4\alpha_2 e^{\lambda t} \\ \alpha_2 \lambda e^{\lambda t} &= 5\alpha_1 e^{\lambda t} - 2\alpha_2 e^{\lambda t},\end{aligned}$$

tj.

$$\begin{aligned}\alpha_1 \lambda &= 7\alpha_1 - 4\alpha_2 \\ \alpha_2 \lambda &= 5\alpha_1 - 2\alpha_2,\end{aligned}$$

neboli

$$\begin{pmatrix} 7 & -4 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

V maticovém tvaru můžeme tuto soustavu vyjádřit jako

$$\mathbf{Ax} = \lambda \mathbf{x}, \quad \text{kde } \mathbf{A} = \begin{pmatrix} 7 & -4 \\ 5 & -2 \end{pmatrix} \quad \text{a} \quad \mathbf{x} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

Tato soustava má triviální řešení $\mathbf{x} = \mathbf{0}$, které pro libovolné λ vede k nulovým funkcím u_1, u_2 . To není příliš zajímavé řešení. Zajímají nás spíš ty hodnoty skaláru λ , pro které existuje *nenulové* řešení soustavy $\mathbf{Ax} = \lambda \mathbf{x}$, tj. soustavy $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$. Nulové řešení existuje právě když je matice $\mathbf{A} - \lambda \mathbf{I}$ singularní, tj. právě když $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$. Nulové řešení proto existuje pouze pro hodnoty λ vyhovující kvadratické rovnici

$$(7 - \lambda)(-2 - \lambda) + 20 = 0,$$

tj.

$$\lambda^2 - 5\lambda + 6 = 0.$$

Tato kvadratická rovnice má řešení $\lambda = 2$ a $\lambda = 3$. Pro $\lambda = 2$ hledáme řešení homogenní soustavy s maticí

$$\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} 5 & -4 \\ 5 & -4 \end{pmatrix},$$

která má řešení tvaru $\mathbf{x} = a(4/5, 1)^T$. Hodnota $\lambda = 2$ tak vede k následujícímu řešení původní soustavy diferenciálních rovnic

$$\mathbf{u}_1 = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = e^{2t} \begin{pmatrix} 4/5 \\ 1 \end{pmatrix}.$$

Druhé řešení kvadratické rovnice $\lambda = 3$ vede k homogenní soustavě s maticí

$$\mathbf{A} - 3\mathbf{I} = \begin{pmatrix} 4 & -4 \\ 5 & -5 \end{pmatrix},$$

která má řešení $\mathbf{x} = b(1, 1)^T$. Hodnota $\lambda = 3$ pak vede k řešení

$$\mathbf{u}_2 = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = e^{3t} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Lze dokázat, že každé řešení původní soustavy diferenciálních rovnic $\mathbf{u}' = \mathbf{A}\mathbf{u}$ je lineární kombinací těchto dvou řešení \mathbf{u}_1 a \mathbf{u}_2 .

Uvedená úloha vede k následující definici.

Definice 11.1 Pro čtvercovou matici \mathbf{A} řádu n s reálnými (komplexními) prvky definujeme vlastní číslo matice \mathbf{A} jako komplexní číslo λ , pro které platí $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ pro nějaký nenulový vektor $\mathbf{x} \in \mathcal{R}^{n \times 1}(\mathcal{C}^{n \times 1})$. Je-li λ vlastní číslo matice \mathbf{A} , pak každé řešení soustavy lineárních rovnic $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$ nazýváme vlastní vektor matice \mathbf{A} příslušný vlastnímu číslu λ . Množinu všech vlastních čísel matice \mathbf{A} nazýváme spektrum matice \mathbf{A} a označujeme ji $\sigma(\mathbf{A})$.

Z definice tak přímo plyne, že pro komplexní matici \mathbf{A} řádu n a komplexní číslo λ platí $\lambda \in \sigma(\mathbf{A})$ právě když matice $\mathbf{A} - \lambda\mathbf{I}_n$ je singulární, což je právě když $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$. Poslední podmínka říká, jak najít vlastní čísla nějaké matice, pokud existují.

Definice 11.2 Je-li \mathbf{A} čtvercová matice řádu n s reálnými (komplexními) prvky, pak pro každé reálné (komplexní) číslo λ definujeme hodnotu $p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}_n)$. Funkci $p(\lambda)$ nazýváme charakteristický polynom matice \mathbf{A} . Rovnici $p(\lambda) = 0$ nazýváme charakteristická rovnice matice \mathbf{A} .

Pro důkaz Tvzení 11.4 budeme potřebovat následující obecnou větu, které se říká *základní věta algebry*. Její důkaz je hodně obtížný a překračuje rámec úvodního kurzu lineární algebry. Uvedeme si ji proto bez důkazu.

Věta 11.3 Je-li $f(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0$ polynom stupně $n \geq 1$ s komplexními koeficienty, pak jej lze vyjádřit ve tvaru

$$f(x) = b_n(x - \beta_1)^{l_1}(x - \beta_2)^{l_2} \dots (x - \beta_k)^{l_k},$$

pro nějaká komplexní čísla β_1, \dots, β_k a kladná celá čísla l_1, \dots, l_k , pro která platí $l_1 + l_2 + \dots + l_k = n$.

Čísla β_1, \dots, β_k se nazývají *kořeny* polynomu f , číslo l_i se nazývá násobnost kořenu β_i .

Tvrzení 11.4 Pro komplexní čtvercovou matici $\mathbf{A} = (a_{ij})$ řádu n platí

- charakteristický polynom matice \mathbf{A} řádu n je polynom stupně n s vedoucím koeficientem rovným $(-1)^n$,
- komplexní číslo λ je vlastním číslem matice \mathbf{A} právě když je kořenem charakteristického polynomu $p(\lambda)$ matice \mathbf{A} ,
- matice \mathbf{A} má n vlastních komplexních čísel, počítáme-li každé tolikrát, kolik je jeho násobnost jako kořene charakteristického polynomu,

- pokud je \mathbf{A} reálná matice, pak $\lambda \in \sigma(\mathbf{A})$ právě když $\bar{\lambda} \in \sigma(\mathbf{A})$.

Důkaz. První tvrzení vyplývá z definice determinantu. Platí

$$\mathbf{A} - \lambda \mathbf{I}_n = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}.$$

Vyjdeme-li z definice determinantu, je především je zřejmé, že determinant matice $\mathbf{A} - \lambda \mathbf{I}_n$ je polynom v proměnné λ . V součinu diagonálních prvků (odpovídajícímu identické permutaci)

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$$

dostaneme po roznásobení, že koeficient u λ^n je $(-1)^n$ a koeficient u λ^{n-1} je $(-1)^{n-1} \sum_{i=1}^n a_{ii}$. V jakémkoliv jiném součinu odpovídajícím neidentické permutaci $p \in S_n$ se musí objevit dva prvky mimo hlavní diagonálu, po případném roznásobení se tak může objevit neznámá λ s koeficientem nejvýše $n-2$. Proto je koeficient u λ^n v charakteristickém polynomu matice \mathbf{A} roven $(-1)^n$ a koeficient u λ^{n-1} se rovná $(-1)^{n-1} \sum_{i=1}^n a_{ii}$. Podobně můžete najít koeficienty u každé mocniny λ^{n-k} pro $k = 2, 3, \dots, n$.

Druhé tvrzení plyne přímo z definice vlastního čísla a charakteristického polynomu.

Třetí tvrzení plyne zase ze základní věty algebry.

Čtvrté tvrzení plyne z toho, že je-li $p(\lambda) = b_n \lambda^n + b_{n-1} \lambda^{n-1} + \cdots + b_1 \lambda + b_0 = 0$ a koeficienty b_0, b_1, \dots, b_n polynomu p jsou reálná čísla, pak rovněž

$$p(\bar{\lambda}) = b_n \bar{\lambda}^n + b_{n-1} \bar{\lambda}^{n-1} + \cdots + b_1 \bar{\lambda} + b_0 = 0.$$

Je-li tedy λ vlastní číslo reálné matice, pak také $\bar{\lambda}$ je vlastní číslo této matice.

□

Každé vlastní číslo λ matice \mathbf{A} je kořenem charakteristického polynomu $p(\lambda)$ matice \mathbf{A} a má tedy nějakou násobnost coby kořen polynomu $p(\lambda)$. Tuto násobnost budeme rovněž nazývat *algebraická násobnost* charakteristického čísla λ , případně *algebraická dimenze* charakteristického čísla λ .

Reálná matice nemusí mít žádné reálné vlastní číslo. Pokud je ale navíc symetrická, pak je každé vlastní číslo této matice reálné. Podobné tvrzení platí i pro hermitovské matice, tj. pro matice s komplexními prvky, pro které platí $\mathbf{B} = \mathbf{B}^*$.

Tvrzení 11.5 *Je-li \mathbf{A} reálná symetrická matice, pak každé vlastní číslo matice \mathbf{A} je reálné. Podobně, je-li \mathbf{B} hermitovská matice, pak každé vlastní číslo matice \mathbf{B} je reálné.*

Důkaz. Důkaz stačí provést pro hermitovské matice, neboť každá symetrická matice je hermitovská. Je-li $\lambda \in \sigma(\mathbf{B})$ a $\mathbf{x} \neq \mathbf{0}$ vlastní vektor příslušný vlastnímu číslu λ , pak platí $\mathbf{x}^* \mathbf{x} \neq 0$ a $\lambda \mathbf{x} = \mathbf{B} \mathbf{x}$. Z poslední rovnosti vyplývá pomocí přechodu ke komplexně konjugovaným maticím také rovnost $\bar{\lambda} \mathbf{x}^* = \mathbf{x}^* \mathbf{B}^*$, a tedy také

$$\mathbf{x}^* \mathbf{x} (\lambda - \bar{\lambda}) = \mathbf{x}^* (\lambda - \bar{\lambda}) \mathbf{x} = \mathbf{x}^* \mathbf{B} \mathbf{x} - \mathbf{x}^* \mathbf{B}^* \mathbf{x} = 0,$$

neboť $\mathbf{B} = \mathbf{B}^*$. Proto $\lambda = \bar{\lambda}$. \square

Cvičení 11.1 *Dokažte, že každé vlastní číslo kososymetrické matice nebo kosohermitovské matice je ryze imaginární.*

Tvrzení 11.6 *Je-li \mathbf{A} čtvercová reálná (komplexní) matice řádu n , \mathbf{P} reálná (komplexní) regulární matice stejného řádu a $\mathbf{B} = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}$, pak obě matice \mathbf{A} a \mathbf{B} mají stejný charakteristický polynom, a proto rovněž $\sigma(\mathbf{A}) = \sigma(\mathbf{B})$, tj. obě mají také stejné spektrum.*

Důkaz. Podle Věty 10.15 o součinu determinantů platí pro každý reálné (komplexní) číslo λ

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}_n) &= \det \mathbf{I}_n \det(\mathbf{A} - \lambda \mathbf{I}_n) = \det(\mathbf{P}^{-1} \mathbf{P}) \det(\mathbf{A} - \lambda \mathbf{I}_n) = \\ &= \det \mathbf{P}^{-1} \det(\mathbf{A} - \lambda \mathbf{I}_n) \det \mathbf{P} = \det(\mathbf{P}^{-1} (\mathbf{A} - \lambda \mathbf{I}_n) \mathbf{P}) = \\ &= \det(\mathbf{P}^{-1} \mathbf{A} \mathbf{P} - \mathbf{P}^{-1} \lambda \mathbf{I}_n \mathbf{P}) = \\ &= \det(\mathbf{B} - \lambda \mathbf{I}_n). \end{aligned}$$

\square

Důsledek 11.7 *Předpokládáme, že \mathcal{V} je vektorový prostor nad reálnými (komplexními) čísly a $F : \mathcal{V} \rightarrow \mathcal{V}$ je lineární zobrazení. Dále předpokládáme, že \mathcal{A} a \mathcal{B} jsou dvě báze prostoru \mathcal{V} . Označíme \mathbf{A} matici $[F]_{\mathcal{A}}$ lineárního zobrazení F vzhledem k bázi \mathcal{A} a \mathbf{B} matici $[F]_{\mathcal{B}}$ zobrazení F vzhledem k bázi \mathcal{B} . Potom jsou charakteristické polynomy obou matic \mathbf{A} a \mathbf{B} stejné a platí rovnost $\sigma(\mathbf{A}) = \sigma(\mathbf{B})$.*

Důkaz. Podle Tvzení 7.13 platí $[F]_{\mathcal{B}} = \mathbf{P}^{-1}[F]_{\mathcal{A}}\mathbf{P}$, kde $\mathbf{P} = [I]_{\mathcal{B}\mathcal{A}}$ je matice přechodu od báze \mathcal{A} k bázi \mathcal{B} . Protože matice přechodu $[I]_{\mathcal{B}\mathcal{A}}$ je regulární, plynou všechna tvrzení bezprostředně z předchozího Tvzení 11.6 \square

Tento důsledek ukazuje, že vlastní čísla a spektrum jsou spíše vlastnosti lineárních zobrazení než pouze vlastnosti matic. Každá reálná (komplexní) matice \mathbf{A} řádu n určuje lineární zobrazení $F : \mathcal{R}^n \rightarrow \mathcal{R}^n$ ($F : \mathcal{C}^n \rightarrow \mathcal{C}^n$) předpisem $F(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Matice \mathbf{A} má vlastní číslo λ právě když existuje nenulový vektor $\mathbf{x} \in \mathcal{R}^n(\mathcal{C}^n)$, pro který platí $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Pro lineární zobrazení F určené maticí \mathbf{A} pak platí

$$F(\mathbf{x}) = \mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Tyto úvahy vedou k následující definici.

Definice 11.8 *Je-li $F : \mathcal{V} \rightarrow \mathcal{V}$ lineární operátor na reálném (komplexním) vektorovém prostoru \mathcal{V} , pak skalár λ nazýváme vlastní číslo lineárního operátoru \mathbf{A} pokud existuje nenulový vektor $\mathbf{x} \in \mathcal{U}$, pro který platí $F(\mathbf{x}) = \lambda\mathbf{x}$. Je-li λ vlastní číslo operátoru F , pak každý vektor $\mathbf{x} \in \mathcal{V}$, pro který platí $F(\mathbf{x}) = \lambda\mathbf{x}$, nazýváme vlastní vektor lineárního operátoru F příslušný vlastnímu číslu λ . Množinu všech vlastních čísel operátoru F označujeme $\sigma(F)$ a nazýváme spektrum operátoru F .*

Bezprostředním důsledkem Tvzení 11.4 je následující tvrzení.

Tvrzení 11.9 *Je-li $F : \mathcal{U} \rightarrow \mathcal{U}$ lineární operátor na komplexním vektorovém prostoru \mathcal{U} , pak existuje vlastní číslo λ operátoru F .*

Pokusíme se vyjasnit, do jaké míry lze matici lineárního operátoru $F : \mathcal{V} \rightarrow \mathcal{V}$ zjednodušit vhodnou volbou báze \mathcal{B} prostoru \mathcal{V} .

Definice 11.10 *Lineární operátor $F : \mathcal{V} \rightarrow \mathcal{V}$ na reálném (komplexním) vektorovém prostoru \mathcal{V} se nazývá diagonalizovatelný, pokud existuje báze \mathcal{B} prostoru \mathcal{V} , pro kterou platí, že matice $[F]_{\mathcal{B}}$ operátoru F vzhledem k bázi \mathcal{B} je diagonální.*

Reálná (komplexní) matice \mathbf{A} řádu n se nazývá diagonalizovatelná, pokud existuje regulární reálná (komplexní) matice \mathbf{P} řádu n , pro kterou platí, že součin $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ je diagonální matice.

Platí tedy, že lineární operátor $F : \mathcal{V} \rightarrow \mathcal{V}$ je diagonalizovatelný právě když je diagonalizovatelná matice $[F]_{\mathcal{A}}$ tohoto operátoru vzhledem k libovolné bázi \mathcal{A} prostoru \mathcal{V} . Ne každá matice je ale diagonalizovatelná.

Úloha 11.1 *Dokažte, že matice*

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

není diagonalizovatelná.

Řešení. Pro matici \mathbf{A} platí $\mathbf{A}^2 = \mathbf{0}$. Pokud by byla diagonalizovatelná, existovaly by regulární matice \mathbf{P} a diagonální matice \mathbf{D} takové, že $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$. Potom by platilo

$$\mathbf{D}^2 = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}^2\mathbf{P} = \mathbf{0},$$

proto rovněž $\mathbf{D} = \mathbf{0}$, a tedy $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \mathbf{0}$, což je spor. Matice \mathbf{A} proto není diagonalizovatelná. \square

Následující tvrzení udává nutnou a postačující podmínku pro diagonalizovatelnost.

Tvrzení 11.11 *Lineární operátor $F : \mathcal{V} \rightarrow \mathcal{V}$ na reálném (komplexním) vektorovém prostoru \mathcal{V} je diagonalizovatelný právě když existuje báze prostoru \mathcal{V} složená z vlastních vektorů operátoru F .*

Čtvercová (komplexní) matice \mathbf{A} řádu n je diagonalizovatelná právě když existuje báze prostoru $\mathcal{R}^n(\mathbb{C}^n)$, která je složená z vlastních vektorů matice \mathbf{A} .

Důkaz. Nejdříve dokážeme druhou část týkající se matic. Pokud je matice \mathbf{A} diagonalizovatelná, existuje regulární matice \mathbf{P} a diagonální matice \mathbf{D} , pro které platí $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, neboli $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D}$. Označme

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Potom platí

$$\mathbf{A}[\mathbf{P}_{*1}|\mathbf{P}_{*2}|\cdots|\mathbf{P}_{*n}] = [\mathbf{P}_{*1}|\mathbf{P}_{*2}|\cdots|\mathbf{P}_{*n}] \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Pro každé $k = 1, 2, \dots, n$ tak platí $\mathbf{A}\mathbf{P}_{*k} = \lambda_k \mathbf{P}_{*k}$. Každý sloupcový vektor \mathbf{P}_{*k} je tak vlastním vektorem matice \mathbf{A} příslušným vlastním číslu λ_k . Matice \mathbf{P} je regulární, její sloupcové vektory jsou proto lineárně nezávislé a tvoří tak bázi prostoru $\mathcal{R}^n(\mathcal{C}^n)$.

Pokud naopak existuje báze $\mathbf{u}_1, \dots, \mathbf{u}_n$ prostoru $\mathcal{R}^n(\mathcal{C}^n)$ složená ze stejných vlastních vektorů matice \mathbf{A} , platí pro každé $k = 1, 2, \dots, n$, že $\mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_k$ pro nějaké skaláry λ_k . Označíme $\mathbf{P} = [\mathbf{u}_1 | \mathbf{u}_2 | \dots | \mathbf{u}_n]$, tato matice je regulární, protože sloupce jsou lineárně nezávislé. Dále označíme

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Stejně jako v předchozí části důkazu ukážeme, že potom $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D}$, neboli $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, matice \mathbf{A} je proto diagonalizovatelná.

Nyní dokážeme první část tvrzení. Zvolíme bázi $\mathcal{A} : \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ prostoru \mathcal{V} libovolně. Označíme $\mathbf{A} = [F]_{\mathcal{A}}$ matici lineárního operátoru F vzhledem k bázi \mathcal{A} . Podle předchozí části důkazu existuje reálná (komplexní) regulární matice $\mathbf{P} = (c_{ij})$ řádu n taková, že $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ pro nějakou diagonální matici \mathbf{D} . Označíme

$$\mathbf{u}_j = \sum_{i=1}^n c_{ij} \mathbf{v}_i$$

pro $i = 1, 2, \dots, n$. Protože je matice \mathbf{P} regulární, je posloupnost vektorů $\mathcal{B} : \mathbf{u}_1, \dots, \mathbf{u}_n$ lineárně nezávislá a tedy báze prostoru \mathcal{V} . Matice \mathbf{P} se potom rovná matici $[I]_{\mathcal{B}\mathcal{A}}$ identického zobrazení vzhledem k bázím \mathcal{B} a \mathcal{A} a je tedy maticí přechodu od báze \mathcal{A} k bázi \mathcal{B} . Matice $[F]_{\mathcal{B}}$ lineárního zobrazení $F : \mathcal{V} \rightarrow \mathcal{V}$ vzhledem k bázi \mathcal{B} se potom rovná podle Tvrzení 7.13 rovná matici $[I]_{\mathcal{A}\mathcal{B}}[F]_{\mathcal{A}}[I]_{\mathcal{B}\mathcal{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$. \square

V první části posledního důkazu jsme dokázali rovněž následující jednoduchý důsledek.

Důsledek 11.12 *Je-li \mathbf{A} čtvercová matice řádu n a \mathbf{P} regulární matice taková, že $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, kde \mathbf{D} je diagonální matice, pak na hlavní diagonále matice \mathbf{D} jsou všechna vlastní čísla matice \mathbf{A} .*

Následující tvrzení ukazuje postačující podmínku, kdy jsou nějaká matice nebo lineární operátor diagonalizovatelné.

Tvrzení 11.13 Jsou-li $\lambda_1, \dots, \lambda_m$ navzájem různá vlastní čísla matice \mathbf{A} řádu n a $\mathbf{u}_i \neq \mathbf{0}$ je vlastní vektor matice \mathbf{A} příslušný vlastnímu číslu λ_i pro libovolné $i = 1, \dots, m$, pak je posloupnost vektorů $\mathbf{u}_1, \dots, \mathbf{u}_m$ lineárně nezávislá.

Má-li matice \mathbf{A} řádu n celkem n navzájem různých vlastních čísel, pak je diagonalizovatelná.

Má-li lineární operátor $F : \mathcal{V} \rightarrow \mathcal{V}$ celkem n navzájem různých vlastních čísel, pak je diagonalizovatelný.

Důkaz. K důkazu první části stačí dokázat, že pro každé $k = 1, \dots, m$, jsou vektory $\mathbf{u}_1, \dots, \mathbf{u}_k$ lineárně nezávislé. Toto tvrzení zřejmě platí pro $k = 1$, protože vektor $\mathbf{u}_1 \neq \mathbf{0}$. Pokud je $k = 2, \dots, m$, tak budeme předpokládat, že posloupnost $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$ je lineárně nezávislá a dokážeme, že také posloupnost $\mathbf{u}_1, \dots, \mathbf{u}_k$ je lineárně nezávislá.

Nechť tedy platí

$$a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_k \mathbf{u}_k = \mathbf{0}$$

pro nějaké skaláry a_1, \dots, a_k . Potom také

$$\mathbf{A}(a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_k \mathbf{u}_k) = a_1 \lambda_1 \mathbf{u}_1 + a_2 \lambda_2 \mathbf{u}_2 + \dots + a_k \lambda_k \mathbf{u}_k = \mathbf{0}.$$

Rovněž platí

$$\lambda_k a_1 \mathbf{u}_1 + \lambda_k a_2 \mathbf{u}_2 + \dots + \lambda_k a_k \mathbf{u}_k = \mathbf{0}.$$

Odečtením posledních dvou rovností dostaneme

$$(\lambda_1 - \lambda_k) a_1 \mathbf{u}_1 + (\lambda_2 - \lambda_k) a_2 \mathbf{u}_2 + \dots + (\lambda_k - \lambda_k) a_k \mathbf{u}_k = \mathbf{0},$$

tj.

$$(\lambda_1 - \lambda_k) a_1 \mathbf{u}_1 + (\lambda_2 - \lambda_k) a_2 \mathbf{u}_2 + \dots + (\lambda_{k-1} - \lambda_k) a_{k-1} \mathbf{u}_{k-1} = \mathbf{0}.$$

Z lineární nezávislosti vektorů $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$ a vzájemné různosti skalárů $\lambda_1, \dots, \lambda_k$ plyne $a_1 = a_2 = \dots = a_{k-1} = 0$, a proto musí platit $a_k \mathbf{u}_k = \mathbf{0}$. Protože $\mathbf{u}_k \neq \mathbf{0}$, dostáváme také $a_k = 0$.

Pokud má matice \mathbf{A} celkem n navzájem různých vlastních čísel a jsou-li $\mathbf{u}_i \neq \mathbf{0}$ vlastní vektory příslušné vlastním číslům λ_i pro $i = 1, 2, \dots, n$, pak je posloupnost vektorů $\mathbf{u}_1, \dots, \mathbf{u}_n$ lineárně nezávislá podle první části důkazu a tedy bází prostoru $\mathcal{R}^n(\mathcal{C}^n)$. Matice \mathbf{A} je proto diagonalizovatelná podle Tvrzení 11.11. \square

Tvrzení 11.14 Čtvercová reálná (komplexní) matice \mathbf{A} řádu n je diagonalizovatelná právě když pro každé vlastní číslo λ matice \mathbf{A} platí, že algebraická násobnost λ se rovná dimenzi nulového prostoru matice $\mathbf{A} - \lambda\mathbf{I}_n$, tj. číslu $\dim \mathcal{N}(\mathbf{A} - \lambda\mathbf{I}_n)$.

Důkaz. Je-li matice \mathbf{A} diagonalizovatelná, existuje regulární matice \mathbf{P} , pro kterou platí $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, kde \mathbf{D} je diagonální matice. Na hlavní diagonále matice \mathbf{D} jsou vlastní čísla $\lambda_1, \lambda_2, \dots, \lambda_n$ matice \mathbf{A} podle Důsledku 11.12. Budeme předpokládat, že přesně k z těchto vlastních čísel se rovná nějakému reálnému (komplexnímu) číslu λ , tj. že algebraická násobnost vlastního čísla λ se rovná k . Diagonální matice $\mathbf{D} - \lambda\mathbf{I}_n$ má potom právě k nul na hlavní diagonále, její hodnota je proto $n - k$. Protože

$$\mathbf{A} - \lambda\mathbf{I}_n = \mathbf{P}(\mathbf{D} - \lambda\mathbf{I}_n)\mathbf{P}^{-1},$$

má matice $\mathbf{A} - \lambda\mathbf{I}_n$ také hodnota $n - k$. Její nulový prostor má proto dimenzi k .

K důkazu obrácené implikace budeme předpokládat, že $\lambda_1, \dots, \lambda_t$ jsou všechna navzájem různá vlastní čísla matice \mathbf{A} . Algebraickou násobnost vlastního čísla λ_i označíme m_i pro $i = 1, \dots, t$. Podle Základní věty algebry 11.3 platí

$$m_1 + m_2 + \dots + m_t = n$$

a podle předpokladu o matici \mathbf{A} dostáváme $m_i = \dim \mathcal{N}(\mathbf{A} - \lambda_i\mathbf{I}_n)$ pro $i = 1, 2, \dots, t$.

V každém z podprostorů $\mathcal{N}(\mathbf{A} - \lambda_i\mathbf{I}_n)$ zvolíme bázi

$$\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{im_i}.$$

Dokážeme, že posloupnost vektorů

$$\mathbf{x}_{11}, \mathbf{x}_{12}, \dots, \mathbf{x}_{1m_1}, \mathbf{x}_{21}, \mathbf{x}_{22}, \dots, \mathbf{x}_{2m_2}, \dots, \mathbf{x}_{t1}, \mathbf{x}_{t2}, \dots, \mathbf{x}_{tm_t}$$

je lineárně nezávislá.

Pokud

$$\sum_{i=1}^t \sum_{j=1}^{m_i} a_{ij} \mathbf{x}_{ij} = \mathbf{0}$$

pro nějaké skaláry a_{ij} , označíme

$$\mathbf{y}_i = \sum_{j=1}^{m_i} a_{ij} \mathbf{x}_{ij}$$

pro $i = 1, \dots, t$.

Potom $\mathbf{y}_i \in \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I}_n)$, tj. \mathbf{y}_i je vlastní vektor matice \mathbf{A} odpovídající vlastnímu číslu λ_i , pro každé $i = 1, \dots, t$. Kromě toho

$$\mathbf{y}_1 + \mathbf{y}_2 + \dots + \mathbf{y}_t = \mathbf{0}.$$

Pokud by byl některý z vektorů $\mathbf{y}_i \neq \mathbf{0}$, označili bychom symbolem J množinu $\{i = 1, 2, \dots, t : \mathbf{y}_i \neq \mathbf{0}\} \neq \emptyset$. Z poslední rovnosti bychom pak dostali

$$\sum_{i \in J} \mathbf{y}_i = \mathbf{0},$$

tj. vektory \mathbf{y}_i pro $i \in J$ by byly lineárně závislé. Protože jsou to vlastní vektory odpovídající různým vlastním číslům matice \mathbf{A} , dostali bychom spor s Tvzením 11.13. Proto $\mathbf{y}_i = \mathbf{0}$ pro každé $i = 1, 2, \dots, t$.

Protože

$$\mathbf{0} = \mathbf{y}_i = \sum_{j=1}^{m_i} a_{ij} \mathbf{x}_{ij}$$

a vektory $\mathbf{x}_{i1}, \dots, \mathbf{x}_{im_i}$ tvoří bázi podprostoru $\mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I}_n)$, platí $a_{i1} = a_{i2} = \dots = a_{im_i} = 0$ pro každé $i = 1, \dots, t$. Posloupnost vektorů

$$\mathbf{x}_{11}, \mathbf{x}_{12}, \dots, \mathbf{x}_{1m_1}, \mathbf{x}_{21}, \mathbf{x}_{22}, \dots, \mathbf{x}_{2m_2}, \dots, \mathbf{x}_{t1}, \mathbf{x}_{t2}, \dots, \mathbf{x}_{tm_t}$$

je proto skutečně lineárně nezávislá, a protože obsahuje n prvků, tvoří bázi prostoru $\mathcal{R}^{n \times 1}(\mathcal{C}^{n \times 1})$. Prostor $\mathcal{R}^{n \times 1}(\mathcal{C}^{n \times 1})$ tak má bázi složenou z vlastních vektorů matice \mathbf{A} , matice \mathbf{A} je proto diagonalizovatelná podle Tvzení 11.11. \square

Následující věta je často nazývána *spektrální věta pro diagonalizovatelné matice*.

Věta 11.15 Čtvercová matice \mathbf{A} řádu n se spektrem $\sigma(\mathbf{A}) = \{\lambda_1, \dots, \lambda_t\}$ je diagonalizovatelná právě když existují matice $\mathbf{E}_1, \dots, \mathbf{E}_t$ řádu n , pro které platí

1. $\mathbf{A} = \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \dots + \lambda_t \mathbf{E}_t$,
2. $\mathbf{E}_i^2 = \mathbf{E}_i$ pro každé $i = 1, 2, \dots, t$,
3. $\mathbf{E}_i \mathbf{E}_j = \mathbf{0}$ pro libovolné dva různé indexy $i, j = 1, 2, \dots, t$,
4. $\mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_t = \mathbf{I}_n$.

Dále pro diagonalizovatelnou matici \mathbf{A} platí, že

5. matice \mathbf{E}_i jsou jednoznačně určeny maticí \mathbf{A} a vlastnostmi 1, 2, 3, 4,
 6. hodnota každé z matic \mathbf{E}_i se rovná algebraické násobnosti charakteristického čísla λ_i ,
 7. je-li $f(x) = c_0 + c_1x + \dots + c_kx^k$ libovolný polynom s komplexními koeficienty, pak platí

$$f(\mathbf{A}) = c_0\mathbf{I}_n + c_1\mathbf{A} + \dots + c_k\mathbf{A}^k = f(\lambda_1)\mathbf{E}_1 + f(\lambda_2)\mathbf{E}_2 + \dots + f(\lambda_k)\mathbf{E}_k,$$

8. nějaká matice \mathbf{B} komutuje s maticí \mathbf{A} právě tehdy, když komutuje s každou z matic \mathbf{E}_i pro $i = 1, 2, \dots, t$.

Důkaz. Označíme m_i algebraickou násobnost vlastního čísla λ_i pro $i = 1, 2, \dots, t$. Z diagonalizovatelnosti matice \mathbf{A} pak vyplývá existence regulární matice \mathbf{P} řádu n , pro kterou platí

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1\mathbf{I}_{m_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \lambda_2\mathbf{I}_{m_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \lambda_t\mathbf{I}_{m_t} \end{pmatrix}.$$

Označíme pro $i = 1, 2, \dots, t$ symbolem \mathbf{D}_i matici, kterou dostaneme z blokové matice na pravé straně poslední rovnosti tak, že nahradíme všechny výskyty vlastního čísla λ_i číslem 1 a výskyty ostatních vlastních čísel λ_j pro $j \neq i$ číslem 0. Například

$$\mathbf{D}_2 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}.$$

Potom platí

$$\begin{aligned} \mathbf{I}_n &= \mathbf{D}_1 + \mathbf{D}_2 + \cdots + \mathbf{D}_t, \\ \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \lambda_1\mathbf{D}_1 + \lambda_2\mathbf{D}_2 + \cdots + \lambda_t\mathbf{D}_t, \\ \mathbf{A} &= \lambda_1\mathbf{P}\mathbf{D}_1\mathbf{P}^{-1} + \lambda_2\mathbf{P}\mathbf{D}_2\mathbf{P}^{-1} + \cdots + \lambda_t\mathbf{P}\mathbf{D}_t\mathbf{P}^{-1}. \end{aligned}$$

Položíme $\mathbf{E}_i = \mathbf{P}\mathbf{D}_i\mathbf{P}^{-1}$ pro $i = 1, 2, \dots, t$ a dostaneme tak ze třetí rovnosti vlastnost 1. Protože $\mathbf{D}_i^2 = \mathbf{D}_i$ a $\mathbf{D}_i\mathbf{D}_j = \mathbf{0}$ pro libovolné různé

indexy $i, j = 1, 2, \dots, t$, dostáváme

$$\begin{aligned}\mathbf{E}_i^2 &= \mathbf{P}\mathbf{D}_i\mathbf{P}^{-1}\mathbf{P}\mathbf{D}_i\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}_i\mathbf{P}^{-1} = \mathbf{E}_i, \\ \mathbf{E}_i\mathbf{E}_j &= \mathbf{P}\mathbf{D}_i\mathbf{P}^{-1}\mathbf{P}\mathbf{D}_j\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}_i\mathbf{D}_j\mathbf{P}^{-1} = \mathbf{0}, \\ \mathbf{E}_1 + \dots + \mathbf{E}_t &= \mathbf{P}(\mathbf{D}_1 + \dots + \mathbf{D}_t)\mathbf{P}^{-1} = \mathbf{P}\mathbf{I}_n\mathbf{P}^{-1} = \mathbf{I}_n,\end{aligned}$$

což dokazuje vlastnosti 2, 3, 4.

Abychom dokázali opačnou implikaci, budeme předpokládat, že pro matici \mathbf{A} existují matice $\mathbf{E}_1, \dots, \mathbf{E}_t$ splňující podmínky 1, 2, 3, 4. Jako první krok dokážeme, že platí

$$\mathcal{S}(\mathbf{E}_1 + \dots + \mathbf{E}_j) = \mathcal{S}(\mathbf{E}_1) + \dots + \mathcal{S}(\mathbf{E}_j)$$

pro každé $j = 1, \dots, t$. Je-li $\mathbf{x} \in \mathcal{S}(\mathbf{E}_1 + \dots + \mathbf{E}_j)$, existuje vektor $\mathbf{y} \in \mathcal{C}^{n \times 1}$ takový, že

$$\mathbf{x} = (\mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_j)\mathbf{y} = \mathbf{E}_1\mathbf{y} + \dots + \mathbf{E}_j\mathbf{y}.$$

Jelikož $\mathbf{E}_i\mathbf{y} \in \mathcal{S}(\mathbf{E}_i)$ pro každé $i = 1, \dots, j$, platí

$$\mathbf{x} = \mathbf{E}_1\mathbf{y} + \dots + \mathbf{E}_j\mathbf{y} \in \mathcal{S}(\mathbf{E}_1) + \dots + \mathcal{S}(\mathbf{E}_j),$$

proto $\mathcal{S}(\mathbf{E}_1 + \dots + \mathbf{E}_j) \subseteq \mathcal{S}(\mathbf{E}_1) + \dots + \mathcal{S}(\mathbf{E}_j)$.

K důkazu opačné inkluze předpokládejme, že

$$\mathbf{x} \in \mathcal{S}(\mathbf{E}_1) + \dots + \mathcal{S}(\mathbf{E}_j).$$

Existují tedy vektory $\mathbf{x}_i \in \mathcal{S}(\mathbf{E}_i)$ pro $i = 1, \dots, j$ takové, že $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_j$. Z $\mathbf{x}_i \in \mathcal{S}(\mathbf{E}_i)$ plyne existence vektoru $\mathbf{y}_i \in \mathcal{C}^{n \times 1}$, pro který platí $\mathbf{x}_i = \mathbf{E}_i\mathbf{y}_i$. Potom platí

$$\begin{aligned}\left(\sum_{i=1}^j \mathbf{E}_i\right)\mathbf{x} &= \left(\sum_{i=1}^j \mathbf{E}_i\right)\left(\sum_{k=1}^j \mathbf{x}_k\right) = \sum_{i=1}^j \sum_{k=1}^j \mathbf{E}_i\mathbf{x}_k = \sum_{i=1}^j \sum_{k=1}^j \mathbf{E}_i\mathbf{E}_k\mathbf{y}_k = \\ &= \sum_{i=1}^j \mathbf{E}_i\mathbf{y}_i = \sum_{i=1}^j \mathbf{x}_i = \mathbf{x}.\end{aligned}$$

Tím je dokázána opačná inkluze $\mathcal{S}(\mathbf{E}_1 + \dots + \mathbf{E}_j) \supseteq \mathcal{S}(\mathbf{E}_1) + \dots + \mathcal{S}(\mathbf{E}_j)$.

Dále dokážeme, že pro každé $j = 1, \dots, t-1$ platí

$$\mathcal{S}(\mathbf{E}_1 + \dots + \mathbf{E}_j) \cap \mathcal{S}(\mathbf{E}_{j+1}) = \{\mathbf{0}\}.$$

Pokud $\mathbf{x} \in \mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_j) \cap \mathcal{S}(\mathbf{E}_{j+1})$, existují vektory $\mathbf{y}, \mathbf{y}_{j+1} \in \mathcal{C}^{n \times 1}$ takové, že

$$\mathbf{x} = (\mathbf{E}_1 + \cdots + \mathbf{E}_j)\mathbf{y} = \mathbf{E}_{j+1}\mathbf{y}_{j+1}.$$

Z této rovnosti a z podmínek 2, 3 dostáváme

$$\mathbf{x} = \mathbf{E}_{j+1}\mathbf{y}_{j+1} = \mathbf{E}_{j+1}\mathbf{E}_{j+1}\mathbf{y}_{j+1} = \mathbf{E}_{j+1}(\mathbf{E}_1 + \cdots + \mathbf{E}_j)\mathbf{y} = \sum_{i=1}^j \mathbf{E}_{j+1}\mathbf{E}_i\mathbf{y} = \mathbf{0},$$

čímž je rovnost $\mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_j) \cap \mathcal{S}(\mathbf{E}_{j+1}) = \{\mathbf{0}\}$ dokázána.

Z této rovnosti, z dříve dokázané rovnosti $\mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_t) = \mathcal{S}(\mathbf{E}_1) + \cdots + \mathcal{S}(\mathbf{E}_t)$, a opakovaným použitím Tvrzení 6.22 pak dostaneme

$$\begin{aligned} n &= \dim \mathbf{I}_n = \dim \mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_{t-1} + \mathbf{E}_t) = \\ &= \dim(\mathcal{S}(\mathbf{E}_1) + \cdots + \mathcal{S}(\mathbf{E}_{t-1}) + \mathcal{S}(\mathbf{E}_t)) = \\ &= \dim(\mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_{t-1}) + \mathcal{S}(\mathbf{E}_t)) = \\ &= \dim \mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_{t-1}) + \dim \mathcal{S}(\mathbf{E}_t) - \\ &\quad - \dim(\mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_j) \cap \mathcal{S}(\mathbf{E}_{j+1})) = \\ &= \dim \mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_{t-2} + \mathbf{E}_{t-1}) + \dim \mathcal{S}(\mathbf{E}_t) = \\ &\quad \vdots \\ &= \dim \mathcal{S}(\mathbf{E}_1) + \cdots + \dim \mathcal{S}(\mathbf{E}_{t-1}) + \dim \mathcal{S}(\mathbf{E}_t). \end{aligned}$$

Označíme $n_i = \dim \mathcal{S}(\mathbf{E}_i)$ pro $i = 1, \dots, t$. V každém z podprostorů $\mathcal{S}(\mathbf{E}_i)$ zvolíme bázi $\mathbf{x}_{i1}, \dots, \mathbf{x}_{in_i}$. Posloupnost vektorů

$$\mathbf{x}_{11}, \dots, \mathbf{x}_{1n_1}, \mathbf{x}_{21}, \dots, \mathbf{x}_{2n_2}, \dots, \mathbf{x}_{t1}, \dots, \mathbf{x}_{tn_t}$$

generuje podprostor $\mathcal{C}^{n \times 1}$, který obsahuje každý z podprostorů $\mathcal{S}(\mathbf{E}_i)$ pro $i = 1, 2, \dots, t$, a tedy také jejich součet $\mathcal{S}(\mathbf{E}_1) + \cdots + \mathcal{S}(\mathbf{E}_t) = \mathcal{S}(\mathbf{E}_1 + \cdots + \mathbf{E}_t) = \mathcal{S}(\mathbf{I}_n) = \mathcal{C}^{n \times 1}$. V posloupnosti je celkem $n_1 + \cdots + n_t = \dim \mathcal{S}(\mathbf{E}_1) + \cdots + \dim \mathcal{S}(\mathbf{E}_t) = n$ prvků, a protože generuje prostor $\mathcal{C}^{n \times 1}$, který má dimenzi n , musí být rovněž lineárně nezávislá podle Tvrzení 6.9, a tedy báze prostoru $\mathcal{C}^{n \times 1}$.

Slupcový prostor matice

$$\mathbf{Q} = [\mathbf{x}_{11} | \cdots | \mathbf{x}_{1n_1} | \mathbf{x}_{21} | \cdots | \mathbf{x}_{2n_2} | \cdots | \mathbf{x}_{t1} | \cdots | \mathbf{x}_{tn_t}]$$

má proto dimenzi n a matice \mathbf{Q} je tak regulární. Protože $\mathbf{I}_n = \mathbf{Q}^{-1}\mathbf{Q}$, platí podle druhé části Tvrzení 3.7

$$\mathbf{e}_k = \mathbf{I}_{*k} = \mathbf{Q}^{-1}\mathbf{Q}_{*k}$$

pro každé $k = 1, \dots, n$. Připomeňme, že \mathbf{e}_k označuje k -tý vektor standardní báze prostoru $\mathcal{C}^{n \times 1}$.

Protože $\mathbf{x}_{il} \in \mathcal{S}(\mathbf{E}_i)$, existuje vektor $\mathbf{y}_{il} \in \mathcal{C}^{n \times 1}$, pro který platí $\mathbf{x}_{il} = \mathbf{E}_i \mathbf{y}_{il}$. Potom platí

$$\mathbf{E}_i \mathbf{x}_{il} = \mathbf{E}_i \mathbf{E}_i \mathbf{y}_{il} = \mathbf{E}_i \mathbf{y}_{il} = \mathbf{x}_{il},$$

zatímco pro každé $j \neq i$, $j \in \{1, 2, \dots, t\}$, platí podle podmínky 3

$$\mathbf{E}_j \mathbf{x}_{il} = \mathbf{E}_j \mathbf{E}_i \mathbf{x}_{il} = \mathbf{0} \mathbf{x}_{il} = \mathbf{0}.$$

Dokážeme nyní, že součin $\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q}$ je diagonální matice. Rovná-li se k -tý sloupcový vektor matice \mathbf{Q} vektoru \mathbf{x}_{il} , dostáváme s využitím podmínky 1, že

$$\begin{aligned} (\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q})_{*k} &= \mathbf{Q}^{-1} \mathbf{A} \mathbf{Q}_{*k} = \mathbf{Q}^{-1} \left(\sum_{j=1}^t \lambda_j \mathbf{E}_j \right) \mathbf{x}_{il} = \mathbf{Q}^{-1} \sum_{j=1}^t \lambda_j \mathbf{E}_j \mathbf{x}_{il} = \\ &= \mathbf{Q}^{-1} \lambda_i \mathbf{E}_i \mathbf{x}_{il} = \mathbf{Q}^{-1} \lambda_i \mathbf{x}_{il} = \lambda_i \mathbf{Q}^{-1} \mathbf{Q}_{*k} = \lambda_i \mathbf{e}_k. \end{aligned}$$

Pro každé $k = 1, \dots, n$ tak platí, že v k -tém sloupci součinu $\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q}$ je jediný případně nenulový prvek na hlavní diagonále (a rovná se jednomu z vlastních čísel λ_i , $i = 1, \dots, t$). Matice $\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q}$ je proto diagonální a matice \mathbf{A} je diagonalizovatelná.

Zbývá dokázat, že pro každou diagonalizovatelnou matici \mathbf{A} platí vlastnosti 5, 6, 7, 8. Protože matice \mathbf{D}_i má hodnotu m_i , má tutéž hodnotu i matice $\mathbf{E}_i = \mathbf{P} \mathbf{D}_i \mathbf{P}^{-1}$, což dokazuje 6.

S použitím vlastností 2, 3, 4 můžeme spočítat

$$\begin{aligned} \mathbf{A}^2 &= (\lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \dots + \lambda_t \mathbf{E}_t) (\lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \dots + \lambda_t \mathbf{E}_t) = \\ &= \sum_{i,j=1}^t \lambda_i \mathbf{E}_i \lambda_j \mathbf{E}_j = \sum_{i=1}^t \lambda_i^2 \mathbf{E}_i^2 = \sum_{i=1}^t \lambda_i^2 \mathbf{E}_i. \end{aligned}$$

Jestliže nyní předpokládáme, že

$$\mathbf{A}^l = \sum_{i=1}^t \lambda_i^l \mathbf{E}_i$$

pro nějaké $l \geq 2$, pak dostáváme, že

$$\begin{aligned} \mathbf{A}^{l+1} &= (\lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \dots + \lambda_t \mathbf{E}_t) (\lambda_1^l \mathbf{E}_1 + \lambda_2^l \mathbf{E}_2 + \dots + \lambda_t^l \mathbf{E}_t) = \\ &= \sum_{i=1}^t \sum_{j=1}^t \lambda_i \mathbf{E}_i \lambda_j^l \mathbf{E}_j = \sum_{i=1}^t \lambda_i^{l+1} \mathbf{E}_i^2 = \sum_{i=1}^t \lambda_i^{l+1} \mathbf{E}_i. \end{aligned}$$

Protože rovněž

$$\mathbf{A}^0 = \mathbf{I}_n = \mathbf{E}_1 + \cdots + \mathbf{E}_t = \lambda_1^0 \mathbf{E}_1 + \cdots + \lambda_t^0 \mathbf{E}_t,$$

rovnost

$$\mathbf{A}^l = \sum_{i=1}^t \lambda_i^l \mathbf{E}_i$$

platí pro každé nezáporné celé číslo l . Pro každé číslo $j = 0, \dots, k$ tak dostáváme

$$c_j \mathbf{A}^j = c_j \sum_{i=1}^t \lambda_i^j \mathbf{E}_i$$

a tedy

$$f(\mathbf{A}) = \sum_{j=0}^k c_j \mathbf{A}^j = \sum_{j=0}^k c_j \left(\sum_{i=1}^t \lambda_i^j \mathbf{E}_i \right) = \sum_{i=1}^t \left(\sum_{j=0}^k c_j \lambda_i^j \right) \mathbf{E}_i = \sum_{i=1}^t f(\lambda_i) \mathbf{E}_i.$$

Abychom dokázali jednoznačnost matic $\mathbf{E}_1, \dots, \mathbf{E}_t$ splňujících podmínky 1, 2, 3, 4, budeme předpokládat, že

$$\mathbf{A} = \lambda_1 \mathbf{F}_1 + \lambda_2 \mathbf{F}_2 + \cdots + \lambda_t \mathbf{F}_t,$$

kde $\mathbf{F}_1 + \cdots + \mathbf{F}_t = \mathbf{I}_n$, $\mathbf{F}_i \mathbf{F}_i = \mathbf{F}_i$ a $\mathbf{F}_i \mathbf{F}_j = \mathbf{0}$ pro libovolné $i \neq j$, $i, j \in \{1, 2, \dots, t\}$. Potom

$$\mathbf{E}_i \mathbf{A} = \mathbf{A} \mathbf{E}_i = \lambda_i \mathbf{E}_i, \quad \mathbf{F}_j \mathbf{A} = \mathbf{A} \mathbf{F}_j = \lambda_j \mathbf{F}_j.$$

Proto pro $i \neq j$ dostáváme

$$\lambda_j \mathbf{E}_i \mathbf{F}_j = \mathbf{E}_i (\mathbf{A} \mathbf{F}_j) = (\mathbf{E}_i \mathbf{A}) \mathbf{F}_j = \lambda_i \mathbf{E}_i \mathbf{F}_j,$$

proto $(\lambda_i - \lambda_j) \mathbf{E}_i \mathbf{F}_j = \mathbf{0}$, neboli $\mathbf{E}_i \mathbf{F}_j = \mathbf{0}$. Můžeme tak spočítat, že platí

$$\mathbf{E}_i = \mathbf{E}_i \mathbf{I}_n = \mathbf{E}_i \left(\sum_{j=1}^t \mathbf{F}_j \right) = \mathbf{E}_i \mathbf{F}_i = \left(\sum_{j=1}^t \mathbf{E}_j \right) \mathbf{F}_i = \mathbf{I}_n \mathbf{F}_i = \mathbf{F}_i$$

pro každé $i = 1, \dots, t$. Tím je dokázána jednoznačnost matic \mathbf{E}_i , tj. podmínka 5.

Zbývá dokázat podmínku 8. Z podmínky 1 okamžitě vyplývá, že pokud matice \mathbf{B} komutuje se všemi maticemi \mathbf{E}_i pro $i = 1, \dots, t$, komutuje rovněž s maticí \mathbf{A} . Abychom dokázali opačnou implikaci, vyjádříme každou matici \mathbf{E}_i jako polynom v matici \mathbf{A} . Označme

$$g(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_t)$$

a

$$g_i(x) = \frac{g(x)}{x - \lambda_i}$$

pro $i = 1, \dots, t$. Polynom $p_i(x)$ má stupeň $t - 1$ a kořeny λ_j pro $j \neq i$. Podle vlastnosti 7 platí

$$g_i(\mathbf{A}) = g_i(\lambda_1)\mathbf{E}_1 + \dots + g_i(\lambda_t)\mathbf{E}_t = g_i(\lambda_i)\mathbf{E}_i.$$

Označíme-li $c_i = g_i(\lambda_i) \neq 0$, dostaneme že

$$\mathbf{E}_i = c_i^{-1}g_i(\mathbf{A}).$$

Pokud tedy matice \mathbf{B} komutuje s maticí \mathbf{A} , komutuje také s každou maticí $c_i^{-1}g_i(\mathbf{A}) = \mathbf{E}_i$ pro $i = 1, \dots, t$. \square

Cvičení 11.2 Zjistěte pro několik matic \mathbf{A} , jsou-li diagonalizovatelné, pokud ano, najděte regulární matici \mathbf{P} , pro kterou platí, že $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ je diagonální matice, a najděte spektrální rozklad matice \mathbf{A} .

Unitární podobnost

Dále se budeme zabývat otázkou, kdy pro diagonalizovatelnou matici \mathbf{A} řádu n existuje ortogonální matice \mathbf{P} , pro kterou platí, že $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ je diagonální matice. Nebo jinak řečeno, kdy existuje ortonormální báze prostoru $\mathcal{C}^{n \times 1}$ tvořená vlastními vektory matice \mathbf{A} .

Definice 11.16 Říkáme, že komplexní čtvercová matice \mathbf{A} řádu n je unitárně diagonalizovatelná, pokud existují unitární matice \mathbf{P} a diagonální matice \mathbf{D} , pro které platí $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$.

Připomeňme si, že čtvercová matice \mathbf{P} je unitární právě když $\mathbf{P}^{-1} = \mathbf{P}^*$.

Tvrzení 11.17 Je-li komplexní matice \mathbf{A} řádu n unitárně diagonalizovatelná, pak platí $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$.

Důkaz. Jestliže existují unitární matice \mathbf{P} a diagonální matice \mathbf{D} , pro které platí $\mathbf{P}^*\mathbf{A}\mathbf{P} = \mathbf{D}$, pak přechodem ke komplexně konjugovaným maticím dostaneme rovněž $\mathbf{P}^*\mathbf{A}^*\mathbf{P} = \mathbf{D}^*$. Protože pro diagonální matici \mathbf{D} platí $\mathbf{D}^*\mathbf{D} = \mathbf{D}\mathbf{D}^*$, dostáváme postupně

$$\begin{aligned} \mathbf{P}^*\mathbf{A}\mathbf{P} \cdot \mathbf{P}^*\mathbf{A}^*\mathbf{P} &= \mathbf{D}\mathbf{D}^* = \mathbf{D}^*\mathbf{D} = \mathbf{P}^*\mathbf{A}^*\mathbf{P} \cdot \mathbf{P}^*\mathbf{A}\mathbf{P} \\ \mathbf{P}^*\mathbf{A}\mathbf{A}^*\mathbf{P} &= \mathbf{P}^*\mathbf{A}^*\mathbf{A}\mathbf{P} \\ \mathbf{A}\mathbf{A}^* &= \mathbf{A}^*\mathbf{A}. \end{aligned}$$

\square

Definice 11.18 *Komplexní čtvercová matice \mathbf{A} se nazývá normální, pokud platí $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$.*

Třída normálních matic je velmi široká, jak se můžete přesvědčit v následujícím cvičení.

Cvičení 11.3 *Dokažte, že všechny matice následujících typů jsou normální: hermitovské matice, kosohermitovské matice, reálné symetrické matice, reálné kososymetrické matice, unitární matice, ortogonální matice, diagonální matice, všechny matice tvaru $\mathbf{P}^* \mathbf{A} \mathbf{P}$, kde \mathbf{A} je normální matice a \mathbf{P} je unitární matice.*

Lemma 11.19 *Je-li \mathbf{A} normální matice a \mathbf{x} vlastní vektor matice \mathbf{A} příslušný vlastnímu číslu λ , pak je \mathbf{x} také vlastní vektor matice \mathbf{A}^* příslušný vlastnímu číslu $\bar{\lambda}$.*

Důkaz. Především si všimněme, že je-li \mathbf{A} normální matice, pak pro každé komplexní číslo λ platí

$$\begin{aligned} (\mathbf{A} - \lambda \mathbf{I}_n)^* (\mathbf{A} - \lambda \mathbf{I}_n) &= (\mathbf{A}^* - \bar{\lambda} \mathbf{I}_n^*) (\mathbf{A} - \lambda \mathbf{I}_n) = \\ &= \mathbf{A}^* \mathbf{A} - \lambda \mathbf{A}^* - \bar{\lambda} \mathbf{A} + \lambda \bar{\lambda} \mathbf{I}_n = \\ &= \mathbf{A} \mathbf{A}^* - \bar{\lambda} \mathbf{A} - \lambda \mathbf{A}^* + \lambda \bar{\lambda} \mathbf{I}_n = \\ &= (\mathbf{A} - \lambda \mathbf{I}_n) (\mathbf{A} - \lambda \mathbf{I}_n)^*. \end{aligned}$$

Matice $\mathbf{A} - \lambda \mathbf{I}_n$ je proto také normální.

Je-li nyní $\mathbf{x} \neq \mathbf{0}$ vlastní vektor matice \mathbf{A} odpovídající vlastnímu číslu λ , pak označíme $\mathbf{B} = \mathbf{A} - \lambda \mathbf{I}_n$. Platí tedy $\mathbf{B}^* \mathbf{B} = \mathbf{B} \mathbf{B}^*$ a $\mathbf{B} \mathbf{x} = \mathbf{0}$. Označíme dále $\mathbf{y} = \mathbf{B}^* \mathbf{x}$. Potom platí

$$\mathbf{0} = (\mathbf{B} \mathbf{x})^* (\mathbf{B} \mathbf{x}) = \mathbf{x}^* \mathbf{B}^* \mathbf{B} \mathbf{x} = \mathbf{x}^* \mathbf{B} \mathbf{B}^* \mathbf{x} = \mathbf{y}^* \mathbf{y}.$$

Odtud plyne, že $\mathbf{y} = \mathbf{0}$, tj. $\mathbf{0} = \mathbf{B}^* \mathbf{x} = (\mathbf{A}^* - \bar{\lambda} \mathbf{I}_n) \mathbf{x}$, tj. $\mathbf{A}^* \mathbf{x} = \bar{\lambda} \mathbf{x}$. Vektor \mathbf{x} je tak vlastní číslem matice \mathbf{A}^* odpovídajícím vlastnímu číslu $\bar{\lambda}$ této matice. \square

Tvrzení 11.20 *Je-li \mathbf{A} normální matice, pak pak vlastní vektory odpovídající různým vlastním číslům matice \mathbf{A} jsou ortogonální.*

Důkaz. Předpokládejme nyní, že λ_1, λ_2 jsou dvě různá vlastní čísla matice \mathbf{A} a $\mathbf{x}_1, \mathbf{x}_2$ jsou jim odpovídající vlastní vektory. Právě jsme dokázali, že pak rovněž $\mathbf{A}^* \mathbf{x}_1 = \bar{\lambda}_1 \mathbf{x}_1$. Spočítáme dále, že

$$\lambda_1 \mathbf{x}_1^* \mathbf{x}_2 = (\bar{\lambda}_1 \mathbf{x}_1)^* \mathbf{x}_2 = (\mathbf{A}^* \mathbf{x}_1)^* \mathbf{x}_2 = (\mathbf{x}_1^* \mathbf{A}) \mathbf{x}_2 = \mathbf{x}_1^* (\mathbf{A} \mathbf{x}_2) = \lambda_2 \mathbf{x}_1^* \mathbf{x}_2.$$

Protože $\lambda_1 \neq \lambda_2$, plyne odtud $\mathbf{x}_1^* \mathbf{x}_2 = 0$, tj. vektory \mathbf{x}_1 a \mathbf{x}_2 jsou ortogonální. \square

Věta 11.21 *Komplexní matice \mathbf{A} řádu n je unitárně diagonalizovatelná právě když je normální.*

Důkaz. Každá unitárně diagonalizovatelná matice \mathbf{A} je normální podle Tvzení 11.17.

Je-li naopak matice \mathbf{A} normální, pak každý vlastní vektor \mathbf{x} matice \mathbf{A} je současně vlastním vektorem matice \mathbf{A}^* . Zvolíme tedy vlastní vektor \mathbf{x}_1 jednotkové délky matice \mathbf{A} příslušný vlastnímu číslu $\lambda \in \sigma(\mathbf{A})$. Podle Lemma 11.19 je \mathbf{x}_1 rovněž vlastní vektor matice \mathbf{A}^* příslušný vlastnímu číslu $\bar{\lambda}$. Označíme $\mathcal{U}_1 = \mathcal{L}(\mathbf{x}_1)^\perp = \{\mathbf{y} \in \mathcal{C}^{n \times 1} : \mathbf{y}^* \mathbf{x}_1 = 0\}$, ortogonální doplněk vektoru \mathbf{x}_1 v prostoru $\mathcal{C}^{n \times 1}$. Dokážeme, že pro každý vektor $\mathbf{y} \in \mathcal{U}_1$ platí, že také $\mathbf{A}\mathbf{y} \in \mathcal{U}_1$. Skutečně,

$$(\mathbf{A}\mathbf{y})^* \mathbf{x}_1 = (\mathbf{y}^* \mathbf{A}^*) \mathbf{x}_1 = \mathbf{y}^* (\mathbf{A}^* \mathbf{x}_1) = \mathbf{y}^* \bar{\lambda} \mathbf{x}_1 = 0.$$

Podobně dokážeme, že také $\mathbf{A}^* \mathbf{y} \in \mathcal{U}_1$.

Lineární zobrazení $F(\mathbf{y}) = \mathbf{A}\mathbf{y}$ je tedy lineární operátor na prostoru \mathcal{U}_1 . Podle Tvzení 11.9 existuje (jednotkový) vlastní vektor \mathbf{x}_2 lineárního operátoru F , tj.

$$\lambda_2 \mathbf{x}_2 = F(\mathbf{x}_2) = \mathbf{A}\mathbf{x}_2.$$

Vektor \mathbf{x}_2 je tak vlastním vektorem matice \mathbf{A} a tedy i matice \mathbf{A}^* . Označíme $\mathcal{U}_2 = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2)^\perp$. Stejně jako v případě prostoru \mathcal{U}_1 dokážeme, že pro každý vektor $\mathbf{y} \in \mathcal{U}_2$ platí jak $\mathbf{A}\mathbf{y} \in \mathcal{U}_2$ tak i $\mathbf{A}^* \mathbf{y} \in \mathcal{U}_2$.

Prostor \mathcal{U}_2 je tak invariantním podprostorem operátoru F , existuje tedy jednotkový vlastní vektor $\mathbf{x}_3 \in \mathcal{U}_2$ operátoru F odpovídající vlastnímu číslu λ_3 , atd.

Takto postupně sestrojíme ortonormální posloupnost $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ tvořenou vlastními vektory matice \mathbf{A} . Pro unitární matici

$$\mathbf{P} = [\mathbf{x}_1 | \mathbf{x}_2 | \dots | \mathbf{x}_n]$$

pak platí $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D}$, kde \mathbf{D} je diagonální matice obsahující na hlavní diagonále postupně vlastní čísla $\lambda_1, \lambda_2, \dots, \lambda_n$. Matice \mathbf{A} je proto unitárně diagonalizovatelná. \square

Důsledek 11.22 *Je-li \mathbf{A} normální matice a $\mathbf{A} = \lambda_1 \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 + \dots + \lambda_t \mathbf{E}_t$ spektrální rozklad matice \mathbf{A} podle Věty 11.15, pak všechny matice \mathbf{E}_i jsou*

hermitovské pro $i = 1, \dots, t$. Obráceně, je-li \mathbf{A} diagonalizovatelná komplexní matice a všechny matice \mathbf{E}_i ve spektrálním rozkladu $\mathbf{A} = \lambda_1 \mathbf{E}_1 + \dots + \lambda_t \mathbf{E}_t$ jsou hermitovské, pak je matice \mathbf{A} normální.

Důkaz. Je-li \mathbf{A} normální, pak $\mathbf{A} = \mathbf{PDP}^{-1}$ pro nějakou unitární matici \mathbf{P} a diagonální matici \mathbf{D} podle Věty 11.21, pak matice $\mathbf{E}_i = \mathbf{PD}_i \mathbf{P}^{-1} = \mathbf{PDP}^*$ sestavené v důkazu Věty 11.15 jsou hermitovské.

Obráceně, jsou-li všechny matice \mathbf{E}_i ve spektrálním rozkladu $\mathbf{A} = \lambda_1 \mathbf{E}_1 + \dots + \lambda_t \mathbf{E}_t$ hermitovské, pak rovněž

$$\mathbf{A}^* = \overline{\lambda_1} \mathbf{E}_1 + \dots + \overline{\lambda_t} \lambda_t \mathbf{E}_t$$

je spektrální rozklad matice \mathbf{A}^* . Pak

$$\mathbf{AA}^* = \lambda_1 \overline{\lambda_1} \mathbf{E}_1 + \dots + \lambda_t \overline{\lambda_t} \mathbf{E}_t = \mathbf{A}^* \mathbf{A},$$

což dokazuje, že \mathbf{A} je normální. \square

Závěrem části o vlastních číslech a vlastních vektorech si ještě ukážeme, jak lze zesílit Větu 9.25 o URV-faktorizaci.

Tvrzení 11.23 *Je-li $\mathbf{A} \in \mathcal{C}^{m \times n}(\mathcal{R}^{m \times n})$ komplexní (reálná) matice, pak každé vlastní číslo matice $\mathbf{A}^* \mathbf{A}$ ($\mathbf{A}^T \mathbf{A}$) je nezáporné reálné číslo.*

Důkaz. Je-li \mathbf{A} komplexní matice, pak součin $\mathbf{A}^* \mathbf{A}$ je hermitovská matice řádu n . Je-li to reálná matice, pak součin $\mathbf{A}^T \mathbf{A}$ je symetrická matice řádu n . V obou případech jsou všechna vlastní čísla matice $\mathbf{A}^* \mathbf{A}$ ($\mathbf{A}^T \mathbf{A}$) reálná podle Tvrzení 11.5.

Je-li λ nějaké vlastní číslo matice $\mathbf{A}^* \mathbf{A}$ a $\mathbf{x} \neq \mathbf{0}$ vlastní vektor odpovídající λ , pak platí $\mathbf{A}^* \mathbf{A} \mathbf{x} = \lambda \mathbf{x}$. Vynásobíme tuto rovnost zleva vektorem \mathbf{x}^* a dostaneme

$$\|\mathbf{Ax}\|^2 = \mathbf{x}^* \mathbf{A}^* \mathbf{A} \mathbf{x} = \lambda \mathbf{x}^* \mathbf{x} = \lambda \|\mathbf{x}\|^2,$$

odkud vyplývá

$$\lambda = \frac{\|\mathbf{Ax}\|^2}{\|\mathbf{x}\|^2} \geq 0.$$

V případě reálné matice \mathbf{A} platí $\mathbf{A}^* = \mathbf{A}^T$, proto rovněž $\lambda \geq 0$ pro každé vlastní číslo λ matice $\mathbf{A}^T \mathbf{A}$. \square

Dále si ještě ukážeme následující tvrzení.

Tvrzení 11.24 *Pro každou reálnou (komplexní) matici \mathbf{A} tvaru $m \times n$ platí*

$$\bullet \quad r(\mathbf{A}^T \mathbf{A}) = r(\mathbf{A}) = r(\mathbf{AA}^T), \quad (r(\mathbf{A}^* \mathbf{A}) = r(\mathbf{A}) = r(\mathbf{AA}^*)),$$

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

Důkaz. Dokážeme pouze “reálnou” část tvrzení, komplexní část se dokáže analogicky, pouze všude nahradíme transponovanou matici \mathbf{A}^T maticí \mathbf{A}^* .

Nejdříve dokážeme, že $\mathcal{N}(\mathbf{A}^T) \cap \mathcal{S}(\mathbf{A}) = \{\mathbf{0}\}$. Je-li $\mathbf{x} \in \mathcal{N}(\mathbf{A}^T) \cap \mathcal{S}(\mathbf{A})$, pak platí $\mathbf{A}^T \mathbf{x} = \mathbf{0}$ a současně $\mathbf{x} = \mathbf{A} \mathbf{y}$ pro nějaký vektor $\mathbf{y} \in \mathcal{R}^{n \times 1}$. Odtud plyne $\mathbf{x}^T \mathbf{x} = \mathbf{y}^T \mathbf{A}^T \mathbf{x} = 0$ a proto $\mathbf{x} = \mathbf{0}$. Z Tvrzení 6.19 pak plyne

$$r(\mathbf{A}^T \mathbf{A}) = r(\mathbf{A}) - \dim \mathcal{N}(\mathbf{A}^T \cap \mathcal{S}(\mathbf{A})) = r(\mathbf{A}).$$

Zaměníme-li role \mathbf{A}^T a \mathbf{A} , dostaneme rovněž $r(\mathbf{A} \mathbf{A}^T) = r(\mathbf{A}^T) = r(\mathbf{A})$.

K důkazu druhé části tvrzení si napřed všimněme, že $\mathcal{S}(\mathbf{A}^T \mathbf{A}) \subseteq \mathcal{S}(\mathbf{A}^T)$. Protože dále

$$\dim \mathcal{S}(\mathbf{A}^T \mathbf{A}) = r(\mathbf{A}^T \mathbf{A}) = r(\mathbf{A}) = r(\mathbf{A}^T) = \dim \mathcal{S}(\mathbf{A}^T),$$

vyplývá odtud rovnost $\mathcal{S}(\mathbf{A}^T \mathbf{A}) = \mathcal{S}(\mathbf{A}^T)$. Třetí část tvrzení ihned vyplývá z druhé, zaměníme-li role \mathbf{A} a \mathbf{A}^T .

K důkazu čtvrté části opět stačí uvědomit si, že $\mathcal{N}(\mathbf{A}) \subseteq \mathcal{N}(\mathbf{A}^T \mathbf{A})$, a dále

$$\dim \mathcal{N}(\mathbf{A}) = n - r(\mathbf{A}) = n - r(\mathbf{A}^T \mathbf{A}) = \dim \mathcal{N}(\mathbf{A}^T \mathbf{A}).$$

Proto $\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^T \mathbf{A})$. Poslední část tvrzení opět ihned vyplývá z předchozí, pokud zaměníme role \mathbf{A} a \mathbf{A}^T . \square

Nyní můžeme dokázat slíbené zesílení Věty 9.25.

Věta 11.25 *Předpokládáme, že \mathbf{A} je reálná (komplexní) matice tvaru $m \times n$ a hodnosti r . Potom existují reálná (komplexní) diagonální matice $\mathbf{D}_{r \times r}$ řádu r a ortogonální (unitární) matice \mathbf{U} řádu n a \mathbf{V} řádu m takové, že platí*

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

Důkaz. Dokážeme pouze část týkající se komplexních matic. Pro reálné matice se věta dokáže analogicky.

Matice $\mathbf{A}^*\mathbf{A}$ je normální, neboť

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

Podle Věty 11.21 je matice $\mathbf{A}^*\mathbf{A}$ unitárně diagonalizovatelná, existuje tedy podle Věty 11.21 unitární matice \mathbf{V} řádu n taková, že

$$\mathbf{V}^*\mathbf{A}^*\mathbf{A}\mathbf{V} = \mathbf{D},$$

kde \mathbf{D} je diagonální matice s nezápornými reálnými čísly λ_i (vlastní čísla matice $\mathbf{A}^*\mathbf{A}$) na hlavní diagonále. Protože $r(\mathbf{A}^*\mathbf{A}) = r(\mathbf{A}) = r$, můžeme předpokládat, že $\lambda_i > 0$ pro $i = 1, \dots, r$ a $\lambda_i = 0$ pro $i = r + 1, \dots, n$.

Označíme $\sigma_i = \sqrt{\lambda_i}$ pro $i = 1, \dots, n$ a

$$\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i}$$

pro $i = 1, \dots, k$. Potom platí pro každé dva indexy $i, j \in \{1, 2, \dots, r\}$

$$\mathbf{u}_j^*\mathbf{u}_i = \frac{\mathbf{v}_j^*\mathbf{A}^*}{\sigma_i} \cdot \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i} = \frac{\mathbf{v}_j^*\mathbf{A}^*\mathbf{A}\mathbf{v}_i}{\lambda_i} = \frac{\lambda_i\mathbf{v}_j^*\mathbf{v}_i}{\lambda_i} = \mathbf{v}_j^*\mathbf{v}_i = \delta_{ij}.$$

Posloupnost vektorů $\mathbf{u}_1, \dots, \mathbf{u}_k$ je tak ortonormální, protože ortonormální je posloupnost vektorů $\mathbf{v}_1, \dots, \mathbf{v}_n$. Můžeme ji proto doplnit do ortonormální báze $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ prostoru $\mathcal{C}^{n \times 1}$.

Označíme $\mathbf{U} = [\mathbf{u}_1 | \dots | \mathbf{u}_n]$. Potom pro i -tý sloupec součinu $\mathbf{U}^*\mathbf{A}\mathbf{V}$ platí

$$[\mathbf{U}^*\mathbf{A}\mathbf{V}]_{*i} = \mathbf{U}^*\mathbf{A}\mathbf{v}_i = \mathbf{U}^*\sigma_i\mathbf{u}_i.$$

Prvek na místě (j, i) v součinu $\mathbf{U}^*\mathbf{A}\mathbf{V}$ se proto rovná

$$[\mathbf{U}^*\mathbf{A}\mathbf{V}]_{ji} = [\mathbf{U}^*]_{j*}\sigma_i\mathbf{u}_i = \mathbf{u}_j^*\sigma_i\mathbf{u}_i = \sigma_i\delta_{ji}.$$

Součin $\mathbf{U}^*\mathbf{A}\mathbf{V}$ je proto diagonální matice \mathbf{D} , prvky na hlavní diagonále \mathbf{D} se rovnají $\sigma_i = \sqrt{\lambda_i}$, kde λ_i jsou vlastní čísla součinu $\mathbf{A}^*\mathbf{A}$. Odtud hned dostaneme (protože matice \mathbf{U} a \mathbf{V} jsou unitární), že $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^*$. \square

Číslům $\sigma_i > 0$ se říká *singulární hodnoty* matice \mathbf{A} . Jsou určeny jednoznačně maticí \mathbf{A} , neboť z rovnosti $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^*$ plyne

$$\mathbf{V}^*\mathbf{A}^*\mathbf{A}\mathbf{V} = \mathbf{V}^*\mathbf{A}^*\mathbf{U}\mathbf{U}^*\mathbf{A}\mathbf{V} = \mathbf{D}^*\mathbf{D} = \mathbf{D}^2.$$

Druhé mocniny singulárních hodnot σ_i^2 jsou proto vlastní hodnoty součinu $\mathbf{A}^*\mathbf{A}$. Geometrický význam singulárních hodnot si ukážeme v kapitole o kvadratických formách.