■ ROZŠÍŘENÍ VÝKLADU ■

2.3 Reálný Schurův rozklad

Při studiu vlastních čísel obecné matice se nevyhneme práci s komplexními čísly a při praktickém počítání musíme využít komplexní aritmetiku. Tu je třeba na běžném počítači emulovat a tudíž je pomalá (součin dvou komplexních čísel odpovídá čtyřem součinům a dvěma součtům reálných čísel). Proto je při výpočtu vhodné přechod do komplexní aritmetiky co nejvíce oddálit. Nyní se budeme věnovat tzv. *reálnému Schurovu rozkladu*, tj. rozkladu reálné matice. Budeme se snažit přiblížit Schurovu rozkladu a zároveň dosáhnout toho, aby výsledné faktory zůstaly reálné. Nejprve dokážeme, že pro reálné symetrické matice lze přímo Schurův rozklad volit reálný. Pak se zaměříme na reálné normální matice a nakonec se budeme věnovat reálným rozkladům obecných reálných matic.

Definice 2.11 (Ortogonální matice a transformace)**.** *Řekneme, že čtvercová reálná matice U je* ortogonální*, jestliže*

$$U^T U = UU^T = I\,,$$

kde I je jednotková matice. Ortogonální transformace *jsou transformace realizované ortogonálními maticemi.*

Poznamenejme, že se pro ortogonální matice též používá přesnější název *ortonormální* matice (její sloupce jsou normalizované). Často však vyplývá přímo z kontextu, že daná ortogonální matice je i ortonormální. Poznamenejme také, že ortogonálními maticemi rozumíme vždy reálné matice, ale ortogonálními vektory rozumíme i komplexní vektory x, y ortogonální vzhledem k danému skalárnímu součinu, např. $y^*x = 0$.

Věta 2.12 (Spektrální rozklad pro reálné symetrické matice)**.** *Nechť A je reálná symetrická matice. Potom existují ortogonální matice U a reálná diagonální matice D takové, že platí $D = U^T AU$.*

Důkaz. Ze cvičení 2.5 víme, že reálná symetrická matice $A \in \mathbb{R}^{n\times n}$ je diagonalizovatelná a má reálná vlastní čísla. Vlastní vektor odpovídající vlastnímu číslu λ je nenulovým řešením homogenní soustavy $(A - \lambda I)x = 0$ s reálnou maticí, a je tedy také reálný. Uvažujeme-li dva vlastní vektory x a y, jež přísluší dvěma různým vlastním číslům λ a μ, platí

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \mu\langle x, y\rangle.$$

Z $\lambda \neq \mu$ plyne $\langle x, y\rangle = 0$, tj. vlastní vektory příslušné k různým vlastním číslům reálné symetrické matice jsou ortogonální. Je-li λ násobné vlastní číslo, je možné volit příslušné vlastní vektory jako ortonormální bázi nulového prostoru matice $A - \lambda I$. Z výše popsané konstrukce plyne, že vlastní vektory reálné symetrické matice lze volit tak, aby tvořily reálnou ortonormální bázi prostoru $\mathbb{R}^n$. Sestavíme-li z vlastních čísel reálnou diagonální matici D a z příslušných vlastních vektorů reálnou ortogonální matici U, platí $AU = UD$, a $A = UDU^T$ je hledaný reálný Schurův rozklad. Jiný důkaz této věty lze provést indukcí přes rozměr matice, analogicky důkazu Schurovy věty, viz [38] či [142, str. 340]. □

V dalším se zaměříme na *normální* reálné matice $A \in \mathbb{R}^{n\times n}$. Je-li $A \in \mathbb{R}^{n\times n}$ normální, pak existují dle Schurovy věty (obecně komplexní) unitární matice U a diagonální matice D takové, že $AU = UD$. Na diagonále matice D jsou vlastní čísla matice A, ve sloupcích matice U jsou příslušné vlastní vektory. Jelikož je A reálná, jsou její vlastní čísla buď reálná, nebo komplexně sdružená. Je-li vlastní číslo λ reálné, lze volit (podobně jako u reálné symetrické matice) příslušné vlastní vektory reálné a ortonormální. Uvažujme tedy takový Schurův rozklad, že vlastní vektory odpovídající reálným vlastním číslům jsou reálné. Vlastní vektory odpovídající komplexně sdruženým párům vlastních čísel jsou komplexně sdružené, neboť z $Az = \lambda z$ plyne pro reálnou matici A vztah $A\overline{z} = \overline{\lambda}\overline{z}$. Dále nechť jsou vlastní čísla na diagonále matice D seřazena tak, že komplexně sdružená čísla vždy bezprostředně následují.

Pro dvojici vlastních čísel $\alpha \pm \mathbf{i}\beta$ a jim odpovídajícím (ortonormálním) vlastním vektorům $x \pm \mathbf{i}y$, kde α, β, x a y jsou reálné, platí

$$A\,[x + \mathbf{i}y, x - \mathbf{i}y] = [x + \mathbf{i}y, x - \mathbf{i}y] \begin{bmatrix} \alpha + \mathbf{i}\beta & 0 \\ 0 & \alpha - \mathbf{i}\beta \end{bmatrix}. \tag{2.5}$$

Uvažujeme-li libovolnou unitární matici $W \in \mathbb{C}^{2\times 2}$, lze (2.5) ekvivalentně přepsat ve tvaru

$$A\,[x + \mathbf{i}y, x - \mathbf{i}y]\,W = [x + \mathbf{i}y, x - \mathbf{i}y]\,WW^* \begin{bmatrix} \alpha + \mathbf{i}\beta & 0 \\ 0 & \alpha - \mathbf{i}\beta \end{bmatrix} W. \tag{2.6}$$

Volíme-li konkrétně

$$W = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -\mathbf{i} \\ 1 & \mathbf{i} \end{bmatrix}, \tag{2.7}$$

což odpovídá pro první sloupec operaci

$$\frac{\sqrt{2}}{2}(x + \mathbf{i}y + x - \mathbf{i}y) = \sqrt{2}\,x$$

a pro druhý sloupec operaci

$$\frac{\sqrt{2}}{2}\,\mathbf{i}\,(-x - \mathbf{i}y + x - \mathbf{i}y) = \sqrt{2}\,y,$$

je výsledkem unitární podobnostní transformace

$$W^* \begin{bmatrix} \alpha + \mathbf{i}\beta & 0 \\ 0 & \alpha - \mathbf{i}\beta \end{bmatrix} W = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

reálná matice, a odpovídající transformace vlastních vektorů je dána vztahem $[x + \mathbf{i}y, x - \mathbf{i}y]\,W = \sqrt{2}\,[x, y]$. Jinak řečeno, volíme-li matici W podle (2.7), pak lze (2.6) přepsat ve tvaru

$$A\left[\sqrt{2}x, \sqrt{2}y\right] = \left[\sqrt{2}x, \sqrt{2}y\right] \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}. \tag{2.8}$$

Z transformace $[x + \mathbf{i}y, x - \mathbf{i}y]\,W = \sqrt{2}\,[x, y]$ a z ortogonality vektorů $x + \mathbf{i}y$ a $x - \mathbf{i}y$ plyne, že vektory $\sqrt{2}x$ a $\sqrt{2}y$ tvoří reálnou ortonormální bázi invariantního podprostoru span$\{x + \mathbf{i}y, x - \mathbf{i}y\}$ nad tělesem komplexních čísel (viz cvičení 2.13). V Schurově rozkladu tak lze každý blok odpovídající dvojici komplexně sdružených vlastních čísel (diagonální blok 2×2 v matici D a jemu odpovídající dvojice sloupců matice U) popsaný rovnicí (2.5) nahradit reálnými bloky stejných rozměrů, tj. rovnicí (2.8). Dostáváme rozklad $A = \tilde{U}\tilde{D}\tilde{U}^T$, jež budeme nazývat *reálný Schurův rozklad*. Reálná matice $\tilde{D}$ je blokově diagonální s bloky 1×1 a 2×2 odpovídajícími reálným vlastním číslům, resp. dvojicím komplexně sdružených vlastních čísel. Sloupce ortogonální matice $\tilde{U}$ odpovídající reálným vlastním číslům i nadále představují příslušné vlastní vektory, dvojice sloupců odpovídající párům komplexně sdružených vlastních čísel představují reálné báze příslušných invariantních podprostorů.

Nyní se dostáváme k reálnému Schurovu rozkladu obecné reálné matice $A \in \mathbb{R}^{n \times n}$. V následujícím textu ukážeme, že každou reálnou matici můžeme ortogonálně transformovat na tzv. *kvazi-trojúhelníkovou* matici.

Definice 2.13 (Kvazi-trojúhelníková matice)**.** *Řekneme, že čtvercová matice T je horní kvazi-trojúhelníková, pokud je blokově horní trojúhelníková,*

$$T = \begin{bmatrix} T_{1,1} & T_{1,2} & \dots & T_{1,m} \\ 0 & T_{2,2} & & T_{2,m} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T_{m,m} \end{bmatrix},$$

kde každý blok na hlavní diagonále je buď 1×1 *nebo* 2×2*.*

Platí následující věta, viz [142, str. 341].

Věta 2.14 (Winter, Murnaghan (Schurova věta pro reálné matice))**.** *Nechť A je reálná čtvercová matice. Potom existují ortogonální matice U a reálná kvazi-trojúhelníková matice T takové, že $T = U^T AU$. Navíc vlastní čísla každého* 2×2 *diagonálního bloku matice T tvoří komplexně sdružený pár.*

Důkaz. Pro reálné symetrické matice a reálné normální matice jsme větu dokázali v předchozím textu. Pro obecnou reálnou matici budeme v důkazu Winter-Murnaghanovy věty postupovat analogicky jako v důkazu Schurovy věty, tj. indukcí přes rozměr matice. Pro $n = 1$ a $n = 2$ tvrzení věty platí. Předpokládejme, že tvrzení platí pro matice řádu $n - 1$ a $n - 2$, kde $n \geq 3$, a ukážeme, že platí i pro matice řádu n.

Nechť je dáno uspořádání vlastních čísel matice A a nechť λ je první vlastní číslo v daném uspořádání. V indukci budeme rozlišovat dva případy. Je-li λ reálné vlastní číslo a v odpovídající (reálný) vlastní vektor, pak se krok indukce redukuje na krok použitý při důkazu Schurovy věty. Z existence rozkladu pro matice řádu $n - 1$ plyne existence rozkladu i pro matice řádu n.

Uvažujme nyní případ, kdy je λ komplexní vlastní číslo (imaginární složka je nenulová). Protože je A reálná, komplexní vlastí čísla se vyskytují v komplexně sdružených párech (včetně násobnosti). Nechť $\alpha \pm \mathbf{i}\beta$, $\beta \neq 0$ je pár komplexně sdružených vlastních čísel a $x \pm \mathbf{i}y$ jim odpovídající pár komplexně sdružených vlastních vektorů. Použijeme-li unitární transformaci (2.6) s unitární maticí W

určenou podle (2.7), dostáváme

$$A\,[x,y] = [x,y] \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}. \tag{2.9}$$

Jelikož jsou vlastní vektory příslušné k různým vlastním číslům lineárně nezávislé, musí být i vektory x a y, $[x,y] = 1/\sqrt{2}\,[x + \mathbf{i}y, x - \mathbf{i}y]\,W$, lineárně nezávislé.

Uvažujme nyní libovolnou ortonormální bázi $\{v_1, v_2\}$ prostoru span$\{x,y\}$ a položme $V = [v_1, v_2]$. Jelikož jsou x a y lineárně nezávislé, existuje reálná regulární matice $M \in \mathbb{R}^{2\times 2}$ taková, že $[x,y] = VM$. Dosazením za $[x,y]$ do (2.9) dostáváme

$$AV = VC, \qquad \text{kde} \qquad C \equiv M \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} M^{-1},$$

a podobnostní transformace nám zajišťuje, že C má vlastní čísla $\alpha \pm \mathbf{i}\beta$. Doplňme matici V tak, aby výsledná matice $H \equiv [V, X] \in \mathbb{R}^{n\times n}$ byla ortogonální. Potom platí

$$H^T A H = \begin{bmatrix} V^T \\ X^T \end{bmatrix} [AV, AX] = \begin{bmatrix} C & V^T A X \\ 0 & X^T A X \end{bmatrix},$$

kde jsme použili $X^T A V = X^T V C = 0$. Na matici $X^T A X \in \mathbb{R}^{(n-2)\times(n-2)}$ aplikujeme indukční předpoklad, tj. existuje ortogonální matice Q a kvazi-trojúhelníková matice $\tilde{T}$ tak, že $Q^T X^T A X Q = \tilde{T}$. Položíme-li $U \equiv [V, XQ]$, potom platí

$$U^T A U = \begin{bmatrix} C & V^T A X Q \\ 0 & \tilde{T} \end{bmatrix} \equiv T,$$

kde U je reálná ortogonální matice a T je reálná kvazi-trojúhelníková matice. □

2.4 Funkce matic

Pojem funkce matice může být chápán různými způsoby. Můžeme uvažovat např. funkci, která dané matici přiřadí její determinant, číslo podmíněnosti apod. Nyní se zaměříme na otázku, jak vhodně zobecnit skalární funkci $f : \mathbb{C} \longrightarrow \mathbb{C}$ na funkci, která dané matici $A \in \mathbb{C}^{n\times n}$ přiřadí matici $f(A)$ stejné dimenze.

Uvažujme nejprve případ, kdy matice A je normální, tedy $A = UDU^*$, $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, a f je analytická. Pak je možné definovat funkci matice pomocí Taylorovy řady

$$f(z) = \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!}\, z^i,$$

kde za proměnnou z dosadíme matici A, tj.

$$\begin{aligned} f(A) &\equiv \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!}\, A^i = U \left(\sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!}\, D^i \right) U^* = U f(D) U^* \\ &= U \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix} U^*. \end{aligned}$$

Zdá se tedy přirozené definovat (analytickou) funkci obecné matice $A \in \mathbb{C}^{n\times n}$ pomocí jejího Jordanova kanonického tvaru s r bloky velikosti $n_1, \dots, n_r$ jako

$$f(A) \equiv Sf(J)S^{-1} = S\,\mathrm{diag}(f(J_{n_1}(\lambda_1)), \dots, f(J_{n_r}(\lambda_r)))\,S^{-1}.$$

V případě, že A není diagonalizovatelná, zbývá definovat vhodně funkci Jordanova bloku $f(J_{n_j}(\lambda_j))$ pro $j = 1, \dots, r$. Nechť jsou derivace $f^{(i)}(\lambda_j)$ definovány pro libovolné nezáporné celé číslo i a nechť lze f rozvinout v Taylorovu řadu

$$f(z) = \sum_{i=0}^{\infty} \frac{f^{(i)}(\lambda_j)}{i!}\,(z-\lambda_j)^i.$$

Pokud za $z - \lambda_j$ dosadíme $J_{n_j}(\lambda_j) - \lambda_j I = J_{n_j}(0)$, dostáváme

$$f(J_{n_j}(\lambda_j)) = \sum_{i=0}^{n_j-1} \frac{f^{(i)}(\lambda_j)}{i!}\,J_{n_j}(0)^i \tag{2.10}$$

$$= \begin{bmatrix} f(\lambda_j) & f'(\lambda_j) & \dots & \frac{f^{(n_j-1)}(\lambda_j)}{(n_j-1)!} \\ & f(\lambda_j) & \ddots & \vdots \\ & & \ddots & f'(\lambda_j) \\ & & & f(\lambda_j) \end{bmatrix}, \tag{2.11}$$

neboť pro $k \geq n_j$, $J_{n_j}(0)^k = 0$. Matice $f(J)$ je tedy zcela určena funkčními hodnotami funkce f a hodnotami jejich derivací na spektru matice A. Předchozí úvahy ukazují výhodnost následující definice maticové funkce.

Definice 2.15 (Funkce matice definovaná pomocí Jordanova kanonického tvaru)**.** *Nechť $A \in \mathbb{C}^{n\times n}$ a funkce f je taková, že $f^{(i)}(\lambda_j)$, $i = 0, \dots, n_j - 1$, existuje pro $j = 1, \dots, r$. Pak definujeme*

$$f(A) \equiv Sf(J)S^{-1} = S\,\mathrm{diag}(f(J_{n_1}(\lambda_1)), \dots, f(J_{n_r}(\lambda_r)))\,S^{-1},$$

kde $f(J_{n_j}(\lambda_j))$ je pro $j = 1, \dots, r$ definováno v (2.11).

Je třeba zdůraznit, že pro definici uvedenou výše není třeba předpokládat analytičnost funkce f.

Ze vztahu (2.10) vyplývá, že maticová funkce $f(J_{n_j}(\lambda_j))$ je ve skutečnosti dána polynomem

$$p(z) \equiv f(\lambda_j) + f'(\lambda_j)\,(z-\lambda_j) + \dots + \frac{f^{(n_j-1)}(\lambda_j)}{(n_j-1)!}\,(z-\lambda_j)^{n_j-1},$$

kde za z dosadíme Jordanův blok $J_{n_j}(\lambda_j)$ a za $z - \lambda_j$ dosadíme $J_{n_j}(0)$. Snadno ověříme, že polynom $p(z)$ je (jednoznačně určeným) Hermiteovým interpolačním polynomem splňujícím interpolační podmínky

$$p^{(i)}(\lambda_j) = f^{(i)}(\lambda_j), \quad i = 0, \dots, n_j - 1. \tag{2.12}$$