Kapitola 9

Ortogonální a unitární matice

Také v této kapitole budeme pracovat pouze nad tělesy reálných a komplexních čísel.

Definice 9.1 *Čtvercová matice* $\mathbf{U}$ *řádu* n *nad* $\mathbf{C}$ *(nebo nad* $\mathbf{R}$*) se nazývá unitární (nebo ortogonální), tvoří-li její sloupce ortonormální množinu vektorů v* $\mathbf{C}^n$ *(nebo v* $\mathbf{R}^n$*).*

Věta 9.1 *Buď* $\mathbf{U}$ *čtvercová matice řádu* n *nad* $\mathbf{C}$ *(nebo* $\mathbf{R}$*). Následující podmínky jsou ekvivalentní:*

1. *matice* $\mathbf{U}$ *unitární (nebo ortogonální),*
2. *matice* $\mathbf{U}$ *má ortonormální řádky,*
3. $\mathbf{U}^*\mathbf{U} = \mathbf{I}_n$*, tj.* $\mathbf{U}^{-1} = \mathbf{U}^*$ *(*$\mathbf{U}^T\mathbf{U} = \mathbf{I}^n$ *v případě reálné matice),*
4. $\|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$ *pro každý vektor* $\mathbf{x} \in \mathbf{C}^n$ *(nebo* $\mathbf{x} \in \mathbf{R}^n$*).*

Důkaz. Dokážeme pouze pro komplexní matici $\mathbf{U}$. V případě reálné matice $\mathbf{U}$ jsou důkazy analogické, v případě implikace 4. $\Rightarrow$ 1. je důkaz dokonce mnohem snazší.

Zřejmě 1. je ekvivalentní s 3. Stejně tak 2. je ekvivalentní s rovností $\mathbf{U}\mathbf{U}^* = \mathbf{I}_n$, která je zase ekvivalentní s 3. v případě čtvercových matic. (Proč?)

1. $\Rightarrow$ 4. Platí $\|\mathbf{U}\mathbf{x}\|^2 = (\mathbf{U}\mathbf{x})^*\mathbf{U}\mathbf{x} = \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{x} = \mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2$.

4. $\Rightarrow$ 1. Pro každý vektor $\mathbf{e}_j$ standardní báze v $\mathbf{C}^n$ platí $\mathbf{U}_{*j} = \mathbf{U}\mathbf{e}_j$. Potom $\|\mathbf{U}_{*j}\|^2 = \|\mathbf{U}\mathbf{e}_j\|^2 = \|e_j\|^2 = 1$, tj. $\|\mathbf{U}_{*j}\| = 1$. Dále platí pro $j \neq k$, že $\mathbf{U}_{*j}^*\mathbf{U}_{*k} = \mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k$. Zbývá tedy dokázat, že $\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k = 0$ pro $j \neq k$.

Použitím podmínky 4. na vektor $\mathbf{e}_j + \mathbf{e}_k$ dostaneme

$$\begin{aligned} (\mathbf{e}_j + \mathbf{e}_k)^*\mathbf{U}^*\mathbf{U}(\mathbf{e}_j + \mathbf{e}_k) &= \|\mathbf{U}(\mathbf{e}_j + \mathbf{e}_k)\|^2 = \|\mathbf{e}_j + \mathbf{e}_k\|^2 \\ &= \|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 = 2, \end{aligned}$$

neboť vektory $\mathbf{e}_j, \mathbf{e}_k$ jsou kolmé a platí Pythagorova věta. Roznásobením a dvojím použitím předpokladu 4. dále dostaneme

$$\begin{aligned} (\mathbf{e}_j + \mathbf{e}_k)^*\mathbf{U}^*\mathbf{U}(\mathbf{e}_j + \mathbf{e}_k) &= \mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k + \mathbf{e}_k^*\mathbf{U}^*\mathbf{U}\mathbf{e}_j + \mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_j + \mathbf{e}_k^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k \\ &= \mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k + (\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k)^* + \|\mathbf{U}\mathbf{e}_j\|^* + \|\mathbf{U}\mathbf{e}_k\|^2 \\ &= 2\,\mathrm{Re}\,(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) + \|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 \\ &= 2\,\mathrm{Re}\,(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) + 2. \end{aligned}$$

Srovnáním výsledků obou posledních výpočtů dostáváme

$$\mathrm{Re}\,(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) = 0$$

Tytéž výpočty provedeme s vektorem $\mathbf{e}_j + \mathrm{i}\mathbf{e}_k$.

$$\begin{aligned} (\mathbf{e}_j + \mathrm{i}\mathbf{e}_k)^*\mathbf{U}^*\mathbf{U}(\mathbf{e}_j + \mathrm{i}\mathbf{e}_k) &= \|\mathbf{U}(\mathbf{e}_j + \mathrm{i}\mathbf{e}_k)\|^2 = \|\mathbf{e}_j + \mathrm{i}\mathbf{e}_k\|^2 \\ &= \|\mathbf{e}_j\|^2 + \|\mathrm{i}\mathbf{e}_k\|^2 = 2, \end{aligned}$$

a

$$\begin{aligned} & (\mathbf{e}_j + \mathrm{i}\mathbf{e}_k)^*\mathbf{U}^*\mathbf{U}(\mathbf{e}_j + \mathrm{i}\mathbf{e}_k) \\ = {} & \mathbf{e}_j^*\mathbf{U}^*\mathbf{U}(\mathrm{i}\mathbf{e}_k) + (\mathrm{i}\mathbf{e}_k)^*\mathbf{U}^*\mathbf{U}\mathbf{e}_j + \mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_j + (\mathrm{i}\mathbf{e}_k)^*\mathbf{U}^*\mathbf{U}(\mathrm{i}\mathbf{e}_k) \\ = {} & \mathrm{i}(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) - \mathrm{i}(\mathbf{e}_k^*\mathbf{U}^*\mathbf{U}\mathbf{e}_j) + \|\mathbf{U}\mathbf{e}_j\|^* + \|\mathbf{U}(\mathrm{i}\mathbf{e}_k)\|^2 \\ = {} & 2\,\mathrm{i}\,\mathrm{Im}\,(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) + \|\mathbf{e}_j\|^2 + \|\mathbf{e}_k\|^2 \\ = {} & 2\,\mathrm{i}\,\mathrm{Im}\,(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) + 2, \end{aligned}$$

tj.

$$\mathrm{Im}\,(\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k) = 0.$$

Proto $\mathbf{e}_j^*\mathbf{U}^*\mathbf{U}\mathbf{e}_k = \mathbf{U}_{*j}^*\mathbf{U}_{*k} = 0$ pro libovolné $j \neq k$. $\square$

Tvrzení 9.2 *Součin unitárních (ortogonálních) matic je unitární (ortogonální) matice.*

Věta 9.3 *Je-li* $\mathbf{A}$ *regulární matice a* $\mathbf{A} = \mathbf{QR}$ *její QR-rozklad, pak jsou matice* $\mathbf{Q}$ *a* $\mathbf{R}$ *určené jednoznačně.*

Definice 9.2 *Je-li* $X \subseteq \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*) pak definujeme* ortogonální doplněk *množiny* X *jako množinu*

$$X^\perp = \{\mathbf{y} :< \mathbf{x}, \mathbf{y} >= 0 \text{ pro každývektor } \mathbf{x} \in X\}.$$

V případě jednoprvkové množiny $X = \{\mathbf{u}\}$ budeme ortogonální doplněk $\{\mathbf{u}\}$ označovat $\mathbf{u}^\perp$.

Tvrzení 9.4 *Pro libovolné dvě podmnožiny* $X, Y \subseteq \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*) platí*

1. $X^\perp$ *je podprostor* $\mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*),*
2. *z* $X \subseteq Y$ *plyne* $Y^\perp \subseteq X^\perp$,
3. $X^\perp = \mathbf{L}(X)^\perp$,
4. *je-li* $\mathbf{P}$ *podprostor* $\mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*), pak* $(\mathbf{P}^\perp)^\perp = \mathbf{P}$,
5. *pro podprostor* $\mathbf{P}$ *platí* $\mathbf{P} \cap \mathbf{P}^\perp = \{\mathbf{0}\}$,
6. *každý prvek* $\mathbf{x} \in \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*) lze jednoznačně vyjádřit ve tvaru* $\mathbf{x} = \mathbf{p} + \mathbf{q}$, *kde* $\mathbf{p} \in \mathbf{P}$ *a* $\mathbf{q} \in \mathbf{P}^\perp$,
7. $(X^\perp)^\perp = \mathbf{L}(X)$.

Definice 9.3 *Je-li* $\mathbf{u} \in \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*),* $\|\mathbf{u}\| = 1$, *pak matici* $\mathbf{I}_n - \mathbf{uu}^*$ *(nebo* $\mathbf{I}_n - \mathbf{uu}^T$*) nazýváme elementární projektor určený vektorem* $\mathbf{u}$.

Tvrzení 9.5 *Buď* $\mathbf{Q}$ *elementární projektor určený vektorem* $\mathbf{u} \in \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*). Pak platí*

1. $\mathbf{Q}^2 = \mathbf{Q}$,
2. $\mathbf{Qx} \in \mathbf{u}^\perp$ *pro každý vektor* $\mathbf{x} \in \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*),*
3. $(\mathbf{I}_n - \mathbf{Q})\mathbf{x} \in \mathbf{L}(\mathbf{u})$.

Definice 9.4 *Buď* $\mathbf{u}$ *nenulový vektor z* $\mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*). Matici*

$$\mathbf{R} = \mathbf{I}_n - \frac{2\mathbf{uu}^*}{\mathbf{u}^*\mathbf{u}} \text{ nebo } \mathbf{R} = \mathbf{I}_n - \frac{2\mathbf{uu}^T}{\mathbf{u}^T\mathbf{u}}$$

nazýváme elementární (Householderův) reflektor *určený vektorem* $\mathbf{u}$.

Tvrzení 9.6 *Buď* $\mathbf{R}$ *elementární reflektor určený vektorem* $\mathbf{u} \in \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*), pak platí*

1. $\mathbf{R}^2 = \mathbf{I}_n$, *tj.* $\mathbf{R}^{-1} = \mathbf{R}$,
2. $\mathbf{R}^* = \mathbf{R}$ *(nebo* $\mathbf{R}^T = \mathbf{R}$*), tj.* $\mathbf{R}$ *je unitární (nebo ortogonální) matice.*

Příklad 9.1 Matice rotací kolem souřadných os v $\mathbf{R}^3$.

Definice 9.5 Rovinná (Givensova) rotační matice *je matice řádu* n

$$\mathbf{P}_{ij} = (a_{kl}) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & c & \cdots & s & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & -s & \cdots & c & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix},$$

kde c, s *jsou komplexní (nebo reálná) čísla taková, že* $c^2+s^2 = 1$. *Konkrétněji,* $a_{ii} = a_{jj} = c$, $a_{ji} = -s$, $a_{ij} = s$, *všechny ostatní prvky na hlavní diagonále se rovnají* 1 *a všechny ostatní prvky mimo hlavní diagonálu se rovnají* 0.

Jednoduchým výpočtem zjistíme, že je-li $\mathbf{x} = (x_1, \ldots, x_n)^T$, pak

$$\mathbf{P}_{ij}\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ cx_i + sx_j \\ \vdots \\ -sx_i + cx_j \\ \vdots \\ x_n \end{pmatrix}.$$

Přímým výpočtem se dokáže

Lemma 9.7 *Rovinná Givensova rotační matice je unitární (ortogonální) matice.*

Tvrzení 9.8 *Pro každý nenulový vektor* $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbf{C}^n$ *(nebo* $\mathbf{R}^n$*) platí*

$$\mathbf{P}_{in} \cdots \mathbf{P}_{i2}\mathbf{P}_{i1} = \|\mathbf{x}\|\, \mathbf{e}_i.$$

Tvrzení 9.9 *Buď* $\mathbf{x} \in \mathbf{C}^n$, $\mathbf{u} = \mathbf{x} \pm \mu\, \|\mathbf{x}\|\, \mathbf{e}_1$, *kde*

$$\mu = \begin{cases} 1, & \text{pokud} \quad x_1 \in \mathbf{R}, \\ \frac{x_1}{|x_1|}, & \text{pokud} \quad x_1 \in \mathbf{C} - \mathbf{R}. \end{cases}$$

Věta 9.10 *Pro každou matici* $\mathbf{A}$ *typu* $m \times n$ *nad* $\mathbf{C}^n$ *(nebo nad* $\mathbf{R}^n$*) existuje unitární (ortogonální) matice* $\mathbf{P}$ *řádu* m *taková, že* $\mathbf{PA} = \mathbf{T} = (t_{ij})$, *kde* $t_{ij} = 0$ *kdykoliv* $i > j$.

Důkaz. Dva důkazy, jeden pomocí Householderových reflektorů, druhý pomocí Givensových rotačních matic. □

Příklad 9.2 Výpočetní náročnost algoritmů použitých na regulární matici řádu n:

1. Gaussova eliminace $\approx \frac{1}{3}n^3$,
2. Gram-Schmidtova ortogonalizace $\approx n^3$,
3. Householderova redukce $\approx \frac{2}{3}n^3$,
4. Givensova redukce $\approx \frac{4}{3}n^3$.

Z uvedených algoritmů nejsou první dva numericky stabilní (i když Gaussovu eliminaci lze modifikovat tak, aby numericky stabilní byla pro většinu praktických problémů), zatímco poslední dva numericky stabilní jsou.

Algoritmus je považován za numericky stabilní, pokud dává přesné řešení blízké úlohy. Co se tím myslí, si ukážeme pomocí následujícího pojmu.

Definice 9.6 *Je-li* $\mathbf{A} = (a_{ij})$ *matice typu* $m \times n$ *nad* $\mathbf{C}$ *(*$\mathbf{R}$*), pak definujeme její* Frobeniovu normu *jako číslo*

$$\|\mathbf{A}\| = \sqrt{\sum_i \sum_j |a_{ij}|^2}.$$

Tvrzení 9.11 *Je-li* $\mathbf{A} = (a_{ij})$ *matice typu* $m \times n$ *nad* $\mathbf{C}$ *(*$\mathbf{R}$*) a* $\mathbf{Q}$ *unitární (ortogonální) matice řádu* m, *pak* $\|\mathbf{A}\| = \|\mathbf{QA}\|$.

Budeme počítat QR-rozklad matice $\mathbf{A}$, tj. vyjádření $\mathbf{A} = \mathbf{QR}$, pomocí ortogonální redukce ma základě Věty 9.10. Kvůli zaokrouhlovacím chybám nedostaneme přesně matice $\mathbf{Q}$ a $\mathbf{R}$, výsledky budou zatížené nějakou chybou, dostaneme $\mathbf{Q} + \mathbf{E}$ a $\mathbf{R} + \mathbf{F}$, kde $\mathbf{E}, \mathbf{F}$ jsou chybové matice zaokrouhlovacích chyb, které mají malou normu. Tyto matice jsou přesným QR-rozkladem matice

$$\tilde{\mathbf{A}} = (\mathbf{Q} + \mathbf{E})(\mathbf{R} + \mathbf{F}) = \mathbf{A} + \mathbf{QF} + \mathbf{RF} + \mathbf{EF}$$

Z těchto matic má $\mathbf{EF}$ zanedbatelnou normu, dále $\|\mathbf{QF}\| = \|\mathbf{F}\|$ podle Tvrzení 9.11 a $\|\mathbf{A}\| = \|\mathbf{QR}\| = \|\mathbf{R}\|$. Matice $\mathbf{QF}$ a $\mathbf{RE}$ proto neobsahují prvky s velkou hodnotou ve srovnání s prvky matice $\mathbf{A}$. Norma rozdílu $\tilde{\mathbf{A}} - \mathbf{A}$ je proto také malá ve srovnání s normou $\|\mathbf{A}\|$.