Lineární algebra II (pro fyziky)

Ortogonální diagonalizace

Dalibor Šmíd

MFF UK

V některých oblastech matematiky a fyziky (funkcionální analýza, kvantová teorie) se používá místo pojmů (lineární) zobrazení mezi vektorovými prostory pojem *(lineární) operátor*. Operátor mezi vektorovými prostory V a W se typicky značí velkým tučným či zdvojeným písmenem

$$\mathbb{A} : V \to W$$

a $\mathbb{A}v \in W$ označuje obraz $v \in V$ v operátoru $\mathbb{A}$. Budeme předpokládat, že V a W jsou unitární prostory nad $\mathbb{C}$, a pojmem operátor rozumět vždy lineární operátor. Pokud $V = W$, mluvíme o *operátoru na* V. Operátor $\mathbb{A}^* : W \to V$, který splňuje $\forall v \in V$, $w \in W$

$$\langle w, \mathbb{A}v \rangle_W = \langle \mathbb{A}^* w, v \rangle_V ,$$

nazýváme operátor *adjungovaný* (či *sdružený*) k $\mathbb{A}$. Indexy W, V rozlišují, ve kterém z prostorů skalární součin počítáme. To je obvykle zřejmé z kontextu, proto dále značíme oba skalární součiny jen $\langle , \rangle$.

Tvrzení

Nechť $B = (v_i)_1^n$, $C = (w_i)_1^m$ *jsou ON báze* V, *resp.* W, $A = [\mathbb{A}]_B^C \in \mathbb{C}^{m\times n}$. *Pak* $\mathbb{A}^*$ *existuje a* $[\mathbb{A}^*]_C^B = A^+ \in \mathbb{C}^{n\times m}$.

Důkaz.

Protože $\mathbb{A}v_j = \sum_{i=1}^m a_{ij}w_i$, je $a_{kj} = \langle w_k, \mathbb{A}v_j\rangle$. Definujeme-li sdružený operátor na bázi předpisem $\mathbb{A}^*w_k := \sum_{i=1}^n \bar{a}_{ki}v_i$, pak $\langle \mathbb{A}^*w_k, v_j\rangle = a_{kj}$ a pro $w = \sum_{k=1}^m r_k w_k$, $v = \sum_{j=1}^n s_j v_j$ platí

$$\langle w, \mathbb{A}v\rangle = \sum_{k,j} \bar{r}_k s_j \langle w_k, \mathbb{A}v_j\rangle = \sum_{k,j} \bar{r}_k s_j \langle \mathbb{A}^*w_k, v_j\rangle = \langle \mathbb{A}^*w, v\rangle$$

Tedy $\mathbb{A}^*$ je sdružený k $\mathbb{A}$ a z jeho definice je jk-tý element jeho matice $[\mathbb{A}^*]_C^B$ roven $\bar{a}_{kj} \equiv a_{jk}^+$. $\square$

Operace $*$ sdružení operátoru je tedy úzce spojená s operací $+$ hermitovského sdružení (v reálném případě transponování) matic.

Další vlastnosti sdružení operátorů lze odvodit z vlastností hermitovského sdružení, ale i přímo z definice ♣:

1. $\mathbb{A}, \mathbb{B} : V \to W$, $\alpha, \beta \in \mathbb{C}$: $(\alpha\mathbb{A} + \beta\mathbb{B})^* = \bar{\alpha}\mathbb{A}^* + \bar{\beta}\mathbb{B}^*$
2. $\mathbb{A} : V \to W$, $\mathbb{B} : U \to V$: $(\mathbb{A}\mathbb{B})^* = \mathbb{B}^*\mathbb{A}^*$
3. $(\mathbb{A}^*)^* = \mathbb{A}$
4. $\operatorname{Ker} \mathbb{A} = (\operatorname{Im} \mathbb{A}^*)^\perp, \operatorname{Ker} \mathbb{A}^* = (\operatorname{Im} \mathbb{A})^\perp$
5. $\operatorname{rank}(\mathbb{A}) = \operatorname{rank}(\mathbb{A}^*)$
6. $\operatorname{Ker} \mathbb{A} = \operatorname{Ker} \mathbb{A}^*\mathbb{A}$, $\operatorname{rank}(\mathbb{A}) = \operatorname{rank}(\mathbb{A}^*\mathbb{A})$

Operátor $\mathbb{A} : V \to V$ (resp. matice $A \in \mathbb{C}^{n \times n}$) se nazývá

- *samosdružený/á*, pokud $\mathbb{A} = \mathbb{A}^*$. (resp. $A = A^+$)
- *unitární*, pokud $\mathbb{A}\mathbb{A}^* = \mathbb{A}^*\mathbb{A} = \mathrm{Id}$. (resp. $A^{-1} = A^+$)
- *normální*, pokud $\mathbb{A}\mathbb{A}^* = \mathbb{A}^*\mathbb{A}$. (resp. $AA^+ = A^+A$)

Zjevně ss/u/n. operátor má vzhledem k ON bázi ss/u/n. matici. Samosdružený nebo unitární operátor je nutně také normální.

Věta (O OG diagonalizaci normálního operátoru)
Nechť V je unitární prostor konečné dimenze. Pro operátor $\mathbb{A}: V \to V$ existuje ortonormální báze prostoru V, vůči níž je jeho matice diagonální, právě když je $\mathbb{A}$ normální.

Důsledek
Pro každou čtvercovou matici A, která je hermitovská nebo unitární, existuje unitární matice U taková, že U^+AU je diagonální.

V důkazu věty použijeme fakt, že každou matici $A \in \mathbb{C}^{n\times n}$ lze zapsat jako $A = URU^+$, kde U je unitární a R horní trojúhelníková (*Schurův rozklad*). Plyne to ♣ z věty

Věta (Schurova reprezentace operátoru)
Nechť V je unitární prostor konečné dimenze, $\mathbb{A}$ operátor na něm. Pak existuje ON báze B ve V taková, že matice $[\mathbb{A}]_B^B$ je horní trojúhelníková.

V důkazu využijeme lemma, které plyne ♣ z GS-ortogonalizace:

Lemma

Nechť V je unitární prostor dimenze n a $(u_1, \dots, u_k)$ ON posloupnost ve V. Pak lze tuto posloupnost doplnit na ON bázi $(u_1, \dots, u_n)$ prostoru V.

Operátor $\mathbb{A}$ má alespoň jedno vlastní číslo $\lambda \in \mathbb{C}$ a k němu vlastní vektor v, $\|v\| = 1$. Z lemmatu máme ON bázi $C = (v, C')$ ve V, a tedy existují $\mathbf{x} \in \mathbb{C}^{n-1}$ a $A' \in \mathbb{C}^{n-1 \times n-1}$ takové, že

$$[\mathbb{A}]_C^C = \begin{pmatrix} \lambda & \mathbf{x}^T \\ \mathbf{o} & A' \end{pmatrix},$$

přičemž $A' = [\mathbb{A}']_{C'}^{C'}$ definuje operátor na podprostoru $v^\perp$. Z indukčního předpokladu má A' Schurův rozklad $U'R'U'^+$, kde $R' = [\mathbb{A}']_{B'}^{B'}$ je horní trojúhelníková pro ON bázi B' ve $v^\perp$. Báze $B := (v, B')$ je ON a

$$[\mathbb{A}]_B^B = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & U'^+ \end{pmatrix} \begin{pmatrix} \lambda & \mathbf{x}^T \\ \mathbf{o} & A' \end{pmatrix} \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & U' \end{pmatrix} = \begin{pmatrix} \lambda & \mathbf{x}^T U' \\ \mathbf{o} & R' \end{pmatrix} \quad \square$$

Netriviální implikace ve větě o OG diagonalizaci je $\Leftarrow$ $\clubsuit$. Dokážeme indukcí, že normální horní trojúhelníková matice je diagonální. Nechť B je ON báze V ze Schurovy reprezentace $A := [\mathbb{A}]_B^B$, čili matice A' v

$$A = \begin{pmatrix} \lambda & \mathbf{x}^T \\ \mathbf{o} & A' \end{pmatrix}, \quad A^+ = \begin{pmatrix} \bar{\lambda} & \mathbf{o}^T \\ \bar{\mathbf{x}} & A'^+ \end{pmatrix}$$

je horní trojúhelníková. Podmínka $A^+A = AA^+$ dává blokově

$$\begin{pmatrix} |\lambda|^2 & \bar{\lambda}\mathbf{x}^T \\ \lambda\bar{\mathbf{x}} & \bar{\mathbf{x}}\mathbf{x}^T + A'^+A' \end{pmatrix} = \begin{pmatrix} |\lambda|^2 + \mathbf{x}^T\bar{\mathbf{x}} & \mathbf{x}^T A'^+ \\ A'\bar{\mathbf{x}} & AA'^+ \end{pmatrix}$$

Protože $\mathbf{x}^T\bar{\mathbf{x}} = \sum_{i=1}^{n-1} |x_i|^2$, dostáváme z pozice 11, že $\mathbf{x} = 0$, tedy A' je i normální matice. Z IP je A' diagonální a tedy je i A.

Proveďme OG diagonalizaci matice $\begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$. Vlastní čísla jsou 4 a -2, druhé z nich je dvojnásobné. Příslušné vlastní podprostory jsou $\langle (1,1,1) \rangle$ a $\langle (1,-1,0),(0,1,-1) \rangle$, vidíme, že jsou skutečně navzájem kolmé. Ve druhém z nich pomocí Gramovy-Schmidtovy metody najdeme ortogonální bázi $((1,-1,0),(1,1,-2))$. V předpisu pro diagonalizaci $A = UDU^+$ je matice U přechodu od kanonické báze do báze z vlastních vektorů unitární, její sloupce tedy tvoří normalizované vlastní vektory. V tomto případě, kdy bylo možné vzít všechny vlastní vektory reálné, je $U^+ = U^T$, tedy U je ortogonální matice. Celkově máme A zapsáno jako součin

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

Věta (Spektrální rozklad normálního operátoru)

Nechť $\mathbb{A} : V \to V$ *je normální operátor,* $\mathbb{P}_\lambda : V \to V$ *operátor ortogonální projekce na vlastní podprostor příslušný vlastnímu číslu* λ*. Pak* $\mathbb{A} = \sum_{\lambda \in \sigma(\mathbb{A})} \lambda \mathbb{P}_\lambda$.

Důkaz.

Nechť $(u_i)_1^n$ je ortonormální báze, vzhledem k níž má $\mathbb{A}$ diagonální matici. Vektor u_i je vlastním vektorem $\mathbb{A}$ s vlastním číslem $\lambda \in \sigma(\mathbb{A})$, takže $\mathbb{P}_\lambda u_i = u_i$ a $\mathbb{P}_\mu u_i = 0$, pokud $\mu \neq \lambda$. Pravá i levá strana rovnosti tedy mají na u_i hodnotu λu_i, rovnají se proto na bázi a tudíž i jako operátory. $\square$

Pro matici $A = U \operatorname{diag}(4, -2, -2) U^T$ z předchozího příkladu, kde $U = (\mathbf{u}_1 | \mathbf{u}_2 | \mathbf{u}_3)$, vypadá ♣ spektrální rozklad jako

$$A = 4\mathbf{u}_1\mathbf{u}_1^T - 2(\mathbf{u}_2\mathbf{u}_2^T + \mathbf{u}_3\mathbf{u}_3^T)$$

Vlastní čísla samosdruženého operátoru jsou reálná:
Unitární operátor $\mathbb{U} : V \to V$ je možné charakterizovat ekvivalentními ♣ podmínkami:

- $\forall v \in V : \|\mathbb{U}v\| = \|v\|$
- $\forall v, w \in V,\ \langle \mathbb{U}v, \mathbb{U}w \rangle = \langle v, w \rangle$
- $U := [\mathbb{U}]_B^B$ je unitární matice pro B ON bázi V.
- Je-li $(v_i)_1^n$ ON báze V, pak je jí i $(\mathbb{U}v_i)_1^n$.
- $\mathbb{U}$ je normální operátor, jehož vlastní čísla leží na jednotkové kružnici.

Příkladem spektrálního rozkladu unitární matice může být třeba

$$\begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^+ + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^+ =$$

$$= i \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \end{pmatrix} + (-i) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \end{pmatrix}$$

Z tvaru matic $\mathbf{u}_1\mathbf{u}_1^T$, $\mathbf{u}_2\mathbf{u}_2^T + \mathbf{u}_3\mathbf{u}_3^T$, $\mathbf{u}_2\mathbf{u}_2^+$ vidíme, jak zapsat matici ortogonální projekce podprostoru $W \leq \mathbb{F}^n$, v němž známe ON bázi. Konkrétně OG projekce do sloupcového prostoru $\mathrm{Im}(A)$ matice A s LN sloupci má matici

$$[P_{\mathrm{Im}\,A}]_K^K = QQ^+,$$

kde Q je matice z QR-rozkladu matice A. Lze to i bez ON báze?

Definice

Je-li $(u_1, \ldots, u_k)$ posloupnost vektorů unitárního prostoru $(V, \langle,\rangle)$ nad $\mathbb{F}$, pak *Gramovou maticí* této posloupnosti rozumíme matici $G \in \mathbb{F}^{k\times k}$, pro niž $G_{ij} = \langle u_i, u_j\rangle$.

- Gramova matice je hermitovská.
- Je-li B ON báze V a $A = ([u_1]^B|\ldots|[u_k]^B)$, pak $G = A^+A$.
- G je regulární, právě když je posloupnost$(u_1, \ldots, u_k)$ LN $\clubsuit$.

Pokud $v \in V$, pak vektor $w = \sum_{i=1}^{k} x_i u_i \in \langle u_1, \ldots, u_k \rangle =: W$ je roven OG projekci $P_W(v)$, právě když $\forall i : u_i \perp v - w$, neboli

$$\langle u_i, v \rangle = x_1 \langle u_i, u_1 \rangle + \ldots + x_k \langle u_i, u_k \rangle$$

To je soustava rovnic $G\mathbf{x} = \mathbf{y} \in \mathbb{F}^k$, kde $y_i := \langle u_i, v \rangle$. Označme $\mathbf{b} = [v]^B \in \mathbb{F}^n$, pak lze tuto soustavu přepsat jako

$$A^+ A\mathbf{x} = A^+ \mathbf{b}$$

To je tzv. *normální soustava* příslušející soustavě $A\mathbf{x} = \mathbf{b}$, která v souřadnicích odpovídá hledání $\mathbf{x}$, aby $\sum_{i=1}^{k} x_i u_i = v$. Není-li $v \in W$, tedy nemá-li přeurčená soustava $A\mathbf{x} = \mathbf{b}$ řešení, nalezne její normální soustava aproximativní řešení minimalizující $\| \sum_{i=1}^{k} x_i u_i - v \|$ nebo též $\|A\mathbf{x} - \mathbf{b}\|$. Má-li A LN sloupce, je $A^+ A$ regulární a $\mathbf{x} = (A^+ A)^{-1} A^+ \mathbf{b}$. Matice P_W je pak vidět z

$$[P_W]_B^B [v]^B = [w]^B = \sum_{i=1}^{k} x_i [u_i]^B = A\mathbf{x} = A(A^+ A)^{-1} A^+ [v]^B$$

Klasický příklad je (lineární) regrese, tj. fitování funkce zadané v n bodech tzv. metodou nejmenších čtverců. Minimalizace výrazu

$$\sum_{i=1}^{n} |ax_i + b - y_i|^2$$

vzhledem k $a, b \in \mathbb{R}$ se zadanými $(x_i, y_i), i \in \{1, \ldots, n\}$ vede na aproximativní řešení přeurčené soustavy lineárních rovnic

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

jejíž normální soustava je

$$\begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{pmatrix}$$

Tu je možné explicitně vyřešit třeba Cramerovým pravidlem.