Lineární algebra I (pro fyziky)

Diagonalizace

Dalibor Šmíd

MFF UK

Diagonalizace $n \times n$ matice A je hledání diagonální matice $D = \begin{pmatrix} \lambda_1 & ... & 0 \\ \vdots & \ddots & \vdots \\ 0 & ... & \lambda_n \end{pmatrix} =: \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, splňující

$$A = RDR^{-1}$$

pro nějakou regulární matici R. Protože $D^k = \begin{pmatrix} \lambda_1^k & ... & 0 \\ \vdots & \ddots & \vdots \\ 0 & ... & \lambda_n^k \end{pmatrix}$ a

$$A^k = \underbrace{RDR^{-1}RDR^{-1} \ldots RDR^{-1}}_{k} = RD^kR^{-1},$$

stačí nám znalost D a R k nalezení libovolné mocniny matice A. Protože $A = [F_A]_K^K$, lze vztah $D = R^{-1}AR$ chápat jako transformační formuli

$$[F_A]_B^B = [\mathrm{Id}]_K^B \ [F_A]_K^K \ [\mathrm{Id}]_B^K,$$

pro nějakou vhodnou bázi B, zvanou *báze z vlastních vektorů*.

Definice

Nechť V je vektorový prostor nad $\mathbb{F}$. Číslo $\lambda \in \mathbb{F}$ je *vlastním číslem* endomorfismu $f \in \mathrm{End}(V)$, pakliže pro nějaký nenulový vektor $v \in V$ platí $f(v) = \lambda v$. Každý vektor, který toto splňuje, se nazývá *vlastním vektorem* f příslušným vlastnímu číslu λ.

Podmínku $f(v) = \lambda v$ lze přepsat jako $(f - \lambda\,\mathrm{Id})(v) = 0$. Vlastní vektory jsou tedy prvky $V_\lambda := \mathrm{Ker}(f - \lambda\,\mathrm{Id})$ neboli *vlastního podprostoru endomorfismu f příslušného vlastnímu číslu λ.*

Dle definice je λ vlastní číslo f, pakliže $\mathrm{Ker}(f - \lambda\,\mathrm{Id}) \neq 0$. Je-li $\dim V = n$ a C báze V, nastává to právě když $[f - \lambda\,\mathrm{Id}]_C^C$ je singulární matice. V označení $[f]_C^C = A$ to znamená, že

$$\det(A - \lambda E) \equiv \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & & \vdots \\ \vdots & & \ddots & \\ a_{n1} & \dots & & a_{nn} - \lambda \end{vmatrix} = 0$$

Pro $n = 2$ a $n = 3$ máme

$$\begin{aligned}\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} &= (-\lambda)^2 + (a_{11}+a_{22})(-\lambda) + (a_{11}a_{22} - a_{21}a_{12}) \\ &= \lambda^2 - \operatorname{Tr} A\ \lambda + \det A\end{aligned}$$

$$\det(A - \lambda E) = (-\lambda)^3 + (\operatorname{Tr} A)(-\lambda)^2 + (\operatorname{Tr}(\operatorname{adj}(A)))(-\lambda) + \det A$$

Charakteristický polynom $p_A(\lambda) := \det(A - \lambda E)$ *matice* $A \in \mathbb{F}^{n\times n}$ má tedy

- stupeň n, s koeficientem $(-1)^n$ u λ^n
- absolutní člen $p(0) = \det(A - 0E) = \det A$
- koeficient $(-1)^{n-1} \operatorname{Tr} A$ u λ^{n-1} a $-\operatorname{Tr}(\operatorname{adj}(A))$ u λ.
- obecný koeficient u λ^{n-k} roven $(-1)^{n-k} \sum_{|I|=k} \det(A_{II})$, kde A_{II} je $k \times k$ podmatice A vzniklá vynecháním všech řádků a sloupců, jejichž indexy nejsou v množině I ♣.

Charakteristický polynom endomorfismu f je $\det(f - \lambda \operatorname{Id})$, čili charakteristický polynom jeho libovolné matice $[f]_C^C$.

Příklady

1. Zobrazení $f : P^n(x, \mathbb{R}) \to P^n(x, \mathbb{R})$ definované jako derivace polynomu $[f(p)](x) := \frac{d}{dx}p(x)$ má jediné vlastní číslo 0, vlastní podprostor je tvořen konstantními polynomy.
2. Vlastními čísly $D = \mathrm{diag}(d_1, \ldots, d_n)$ jsou $d_1, \ldots, d_n$ a příslušné vlastní podprostory jsou (pro případ vzájemně různých d_i) lineární obaly prvků kanonické báze $\langle \mathbf{e}_1 \rangle, \ldots, \langle \mathbf{e}_n \rangle \in \mathbb{C}^n$.
3. Pokud $Ax = \lambda x$ a A je regulární, pak $A^{-1}x = \lambda^{-1}x$, čili inverzní matice má stejné vlastní podprostory, ale převrácená vlastní čísla.
4. Pokud $A = Q^{-1}BQ$, pak $Ax = \lambda x$ znamená $BQx = \lambda Qx$, tedy λ je vlastním číslem B s vlastním vektorem Qx. Podobné matice mají tedy stejná vlastní čísla.
5. $\det(A^T - \lambda E) = \det(A - \lambda E)^T = \det(A - \lambda E)$, tedy A^T má stejná vlastní čísla jako A.

Vlastní čísla lze tedy nalézt jako kořeny charakteristického polynomu. Teorie se zjednoduší, když budeme moci předpokládat, že A je komplexní matice a vlastní čísla také hledáme v $\mathbb{C}$. Opíráme se přitom o tzv. Základní větu algebry:

Věta

Každý nekonstantní polynom s komplexními koeficienty má alespoň jeden kořen v $\mathbb{C}$.

Důkaz je nad rámec kurzu. Můžeme ale odvodit jednoduchý důsledek:

Důsledek

Nechť $p(x)$ je polynom stupně $n > 0$ s komplexními koeficienty. Pak má n komplexních kořenů včetně násobností.

Důkaz indukcí. Pro $n = 1$ zjevně platí. Nechť $x_0 \in \mathbb{C}$ je kořen $p(x) = \sum_{i=0}^{n} a_i x^i$. Pak $p(x) = p(x) - p(x_0) = \sum_{i=0}^{n} a_i (x^i - x_0^i)$, což lze zapsat jako$(x - x_0)q(x)$, kde $q(x)$ je polynom stupně $n - 1$. Ten má z indukčního předpokladu $n - 1$ kořenů včetně násobností. Tedy $p(x)$ má n kořenů včetně násobností.

Množina všech vlastních čísel matice A se nazývá její *spektrum*, značí se $\sigma(A)$. Pro $\lambda_i \in \sigma(A)$ je jeho *algebraická násobnost* definována jako násobnost λ_i coby kořenu $p_A(\lambda)$, a jeho *geometrická násobnost* jako dimenze prostoru $\mathrm{Ker}(A - \lambda_i E)$. Analogicky jsou definovány tytéž pojmy pro endomorfismus.

LEMMA

Nechť V je vektorový prostor nad $\mathbb{F}$ dimenze n, $f \in \mathrm{End}(V)$ a pro všechny prvky množiny $M = \{\lambda_1, \ldots, \lambda_k\}$ je zvolena nějaká báze B_i prostoru V_{λ_i}. Pak je množina $B = B_1 \cup \ldots \cup B_k$ lineárně nezávislá.

Je-li V nad $\mathbb{C}$, $M = \sigma(f)$ a každé vlastní číslo f má stejnou algebraickou i geometrickou násobnost, pak má $|B| = n$ a je to tedy na základě lemmatu báze V. Speciálně to musí nastat v případě, že jsou všechny algebraické násobnosti všech $\lambda_i \in \sigma(f)$ rovny jedné. Protože $\dim \mathrm{Ker}(f - \lambda_i \,\mathrm{Id})$ je vždy alespoň 1, má B alespoň n prvků, a protože je LN, musí jich mít právě n.

Důkaz.

Uvažujme množinu $N \subset V$ nenulových vektorů, v níž každý prvek patří jinému vlastnímu prostoru V_{λ_i}. Je-li N LZ, lze z ní vybrat bázi $N' = \{v_{i_1}, \dots, v_{i_q}\}$ jejího lineárního obalu, kde $v_{i_j} \in V_{\lambda_{i_j}}$, a vyjádřit nějaký $v_j \in N \setminus N'$ jako $v_j = \sum_{p=1}^{q} r_p v_{i_p}$. Je-li $v_j \in V_{\lambda_j}$, pak $f(v_j) = \lambda_j v_j = \sum_{p=1}^{q} r_p \lambda_j v_{i_p}$, ale zároveň

$$f(v_j) = \sum_{p=1}^{q} r_p f(v_{i_p}) = \sum_{p=1}^{q} r_p \lambda_{i_p} v_{i_p}$$

Tedy $\sum_{p=1}^{q} r_p(\lambda_{i_p} - \lambda_j) v_{i_p} = 0$. Protože žádný rozdíl $\lambda_{i_p} - \lambda_j$ není 0 a i některé r_p musí být nenulové, je to ve sporu s lineární nezávislostí N'. Tedy N nemůže být LZ.

Označme nyní $B_i = (w_{i1}, \dots, w_{in_i})$ a uvažujme LK ve tvaru $\sum_{i_1=1}^{n_1} r_{1i_1} w_{1i_1} + \ldots + \sum_{i_k=1}^{n_k} r_{ki_k} w_{ki_k} = o$. Označme jednotlivé sumy jako $v_1, \dots, v_k$, pak $v_j \in V_{\lambda_j}$. Podle prvního odstavce musí být všechny v_j nulové, a protože B_j je LN, musí být pak nulové i všechny koeficienty r_{ji_j}. Tedy B je LN. $\square$

Předpokládejme nyní, že pro $f \in \mathrm{End}(V)$ v prostoru V máme bázi $B = (v_1, \dots, v_n)$ složenou z vlastních vektorů f. Pokud $v_i \in V_\lambda$ pro nějaké $\lambda_i \in \sigma(f)$, pak $f(v_i) = \lambda_i v_i$, neboli $[f(v_i)]^B = \lambda_i \mathbf{e}_i$. Pak ale

$$[f]_B^B = \big([f(v_1)]^B \big| \dots \big| [f(v_n)]^B\big) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

V tomto vztahu na rozdíl od předchozích dvou snímků nemusí být všechna λ_i vzájemně různá. Speciálně pro $f = F_A$ s bází z vlastních vektorů $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ máme

$$\begin{aligned} A \equiv [F_A]_K^K = [\mathrm{Id}]_B^K \; [F_A]_B^B \; [\mathrm{Id}]_K^B = \\ = (\mathbf{v}_1 | \dots | \mathbf{v}_n) \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} (\mathbf{v}_1 | \dots | \mathbf{v}_n)^{-1} \end{aligned}$$

Příklad

Diagonalizujme matici $A = \begin{pmatrix} 1 & 1 \\ -3 & 5 \end{pmatrix}$. Char. polynom je

$$p_A(\lambda) = \lambda^2 - \operatorname{Tr} A\ \lambda + \det A = \lambda^2 - 6\lambda + 8 = (\lambda - 2)(\lambda - 4)$$

Tedy $\sigma(A) = \{2, 4\}$. Vlastní podprostory pak jsou

$$V_4 = \operatorname{Ker} \begin{pmatrix} -3 & 1 \\ -3 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\rangle, \ V_2 = \operatorname{Ker} \begin{pmatrix} -1 & 1 \\ -3 & 3 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$$

Tedy

$$A = \begin{pmatrix} 1 & 1 \\ -3 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \frac{1}{(-2)} \begin{pmatrix} 1 & -1 \\ -3 & 1 \end{pmatrix}$$

Má-li A včetně násobností n vlastních čísel, pak je její charakteristický polynom rozložitelný, tj.

$$\begin{aligned} p_A(\lambda) &\equiv (-\lambda)^n + \operatorname{Tr} A(-\lambda)^{n-1} + \ldots + \det(A) \\ &= (\lambda_1 - \lambda)(\lambda_2 - \lambda)\ldots(\lambda_n - \lambda) \\ &= (-\lambda)^n + (\lambda_1 + \ldots + \lambda_n)(-\lambda)^{n-1} + \ldots + \lambda_1\lambda_2\ldots\lambda_n \end{aligned}$$

Z porovnání absolutních členů vidíme, že $\det A$ je roven součinu všech n vlastních čísel a $\operatorname{Tr} A$ je rovna součtu všech vlastních čísel. Charakteristický polynom, vlastní čísla a jejich algebraické i geometrické násobnosti, stopa, determinant i ostatní koeficienty charakteristického polynomu se nemění při podobnostní transformaci $A \to R^{-1}AR$. Jsou to tedy invarianty, lze je zavést i pro endomorfismy, se zachováním vztahů mezi nimi. Jeden takový je (pro 2×2 matice) ♣,

$$\det A = \frac{1}{2}(\operatorname{Tr}(A)^2 - \operatorname{Tr}(A^2))$$

Koeficienty $p_A(\lambda)$ lze vždy vyjádřit pomocí $\operatorname{Tr}(A^k)$, $k \leq n$ ♣.

Některé endomorfismy a matice diagonalizovatelné nejsou, protože LN množina $B = B_1 \cup \ldots \cup B_k$ neobsahuje dost vektorů, aby to byla báze. Příkladem je třeba matice

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \;\Rightarrow\; \sigma(A) = \{0\}, \text{ tedy } \; V_0 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle,$$

nebo endomorfismus D prostoru $V = P^n(x, \mathbb{C})$, který přiřazuje polynomu $p(x)$ jeho první derivaci podle x. Pokud definujeme ve V bázi $C = (1, x, \frac{x^2}{2}, \ldots, \frac{x^n}{n!})$, spojuje ji D do řetízku

$$0 \xleftarrow{\;D\;} 1 \xleftarrow{\;D\;} x \xleftarrow{\;D\;} \frac{x^2}{2} \xleftarrow{\;D\;} \ldots \xleftarrow{\;D\;} \frac{x^{n-1}}{(n-1)!} \xleftarrow{\;D\;} \frac{x^n}{n!} \;,$$

z něhož lze vyčíst, že $[D]^C_C = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & & & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$. Matice tohoto typu se objevují v teorii Jordanova tvaru, která řeší otázku matic, které diagonalizovatelné nejsou. Více v letním semestru.

Příklad

Rotace v rovině $\mathbb{R}^2$ o pravý úhel má matici vzhledem ke K

$$[F_A]_K^K \equiv A = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

Ta má charakteristický polynom λ^2+1, tedy nemá reálná vlastní čísla, ani nemá v $\mathbb{R}^2$ žádné vlastní vektory. Chápeme-li ale F_A jako endomorfismus prostoru $\mathbb{C}^2$, pak $\sigma(A)=\{i,-i\}$ a příslušné vlastní podprostory jsou $\langle(1,-i)\rangle$, resp. $\langle(1,i)\rangle$. Diagonalizace pak má tvar

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}$$

Pro obecnou matici rotace se analogicky spočítá ♣, že

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \frac{1}{2i} \begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}$$

Nechť $A \in \mathbb{C}^{m \times n}$, pak $\bar{A}$ označuje matici stejného typu, jejíž každý element $\bar{a}_{ij}$ je komplexně sdružený k odpovídajícímu elementu a_{ij} matice A. Matice $A^+ := \bar{A}^T$ se nazývá matice *hermitovsky sdružená* k matici A. Protože $\overline{AB} = \bar{A}\bar{B}$, platí u hermitovského sdružení součinu analogická vlastnost jako u transponování součinu, a sice $(AB)^+ = B^+A^+$ ♣.
Pokud $A^+ = A$, říkáme, že A je *hermitovská matice.*

Věta
Nechť $A \in \mathbb{C}^{n \times n}$ je hermitovská matice. Pak každé její vlastní číslo je reálné.

Důkaz.
Pokud $\lambda \in \mathbb{C}$, $\mathbf{v} \in \mathbb{C}^n$ splňují $A\mathbf{v} = \lambda\mathbf{v}$, máme i $\mathbf{v}^+A^+ = \bar{\lambda}\mathbf{v}^+$ a

$$\lambda\mathbf{v}^+\mathbf{v} = \mathbf{v}^+A\mathbf{v} = \mathbf{v}^+A^+\mathbf{v} = \bar{\lambda}\mathbf{v}^+\mathbf{v}$$

Protože $\mathbf{v}^+\mathbf{v} = \sum_{i=1}^n |v_i|^2 \neq 0$ pro $\mathbf{v} \neq 0$, musí být $\lambda = \bar{\lambda}$. $\square$

Speciálním případem hermitovské matice je reálná symetrická matice, tedy $A = A^T$. Pokud $\lambda \in \mathbb{R}$ je její vlastní číslo, pak i $A - \lambda E$ je reálná matice a báze jejího jádra sestává z vektorů z $\mathbb{R}^n$.

Tvrzení

Nechť $A \in \mathbb{R}^{n \times n}$ je symetrická matice, $\lambda, \mu \in \sigma(A)$, $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\mathbf{v} \in V_\lambda$, $\mathbf{w} \in V_\mu$. Pak $\mathbf{v} \perp \mathbf{w}$.

Důkaz.

Z $A\mathbf{v} = \lambda\mathbf{v}$ a $A\mathbf{w} = \mu\mathbf{w}$ plyne, že

$$\mu\mathbf{v}^T\mathbf{w} = \mathbf{v}^T A\mathbf{w} = \mathbf{v}^T A^T\mathbf{w} = (A\mathbf{v})^T\mathbf{w} = \lambda\mathbf{v}^T\mathbf{w}$$

Protože $\lambda \neq \mu$, musí být $\mathbf{v}^T\mathbf{w} = 0$, neboli $\mathbf{v} \perp \mathbf{w}$. □

Má-li matice A n vzájemně různých vlastních čísel, je v bázi z vlastních vektorů každý vektor kolmý na každý jiný. Takové bázi se říká *ortogonální*. V letním semestru ukážeme, že ortogonální bázi z vlastních vektorů lze najít pro každou symetrickou matici.

Z každé ortogonální báze $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ v $\mathbb{R}^n$ lze vytvořit bázi $(\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|})$, která je stále ortogonální a navíc má každý vektor normu 1. Taková báze se nazývá *ortonormální*. Protože $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v}$, plyne odtud

Tvrzení
Nechť $B = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ je ortonormální báze $\mathbb{R}^n$, $U = [\mathrm{Id}]_B^K$. Pak $U^T = U^{-1}$.

Důkaz.
Platí $U^T U = (\mathbf{u}_1|\ldots|\mathbf{u}_n)^T(\mathbf{u}_1|\ldots|\mathbf{u}_n) = E$. $\square$

Reálná čtvercová matice U s vlastností $U^T U = E$ se nazývá *ortogonální matice*. Pokud tedy A je reálná $n \times n$ symetrická matice, která má n vzájemně různých vlastních čísel, víme, že musí existovat ortogonální matice U taková, že $U^T AU$ je diagonální.

Několik pozorování k ortogonálním maticím:

1. Endomorfismus $F_R \in \mathrm{End}(\mathbb{R}^n)$, kde R je ortogonální, zachovává normu, neboť $\|R\mathbf{v}\|^2 = \mathbf{v}^T R^T R \mathbf{v} = \|\mathbf{v}\|^2$.
2. Zachovává i skalární součin, protože ten lze vyjádřit pomocí norem $\mathbf{v}^T\mathbf{w} = \frac{1}{4}(\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2)$ $\clubsuit$.
3. Matice rotace vzhledem ke kanonické bázi je tedy ortogonální matice.
4. Množina všech $n \times n$ ortogonálních matic tvoří grupu, tedy speciálně i matice $U^T R U$, kde U, R jsou ortogonální, je ortogonální. $\clubsuit$. Pokud tedy B je ortonormální báze a $U = [\mathrm{Id}]_B^K$, pak $[F_R]_B^B$ je ortogonální matice.
5. Chápeme-li F_R jako endomorfismus $\mathbb{C}^n$ a $\mathbf{v} \in \mathbb{C}$ je jeho vlastní vektor příslušný vlastnímu číslu $\lambda \in \mathbb{C}$, pak

$$\bar{\lambda}\lambda\mathbf{v}^+\mathbf{v} = (R\mathbf{v})^+(R\mathbf{v}) = \mathbf{v}^+R^+R\mathbf{v} = \mathbf{v}^+R^TR\mathbf{v} = \mathbf{v}^+\mathbf{v},$$

tedy vlastní čísla F_R leží na jednotkové kružnici v $\mathbb{C}$.