Bevezetés az algebrába 2 – Cayley–Hamilton-tétel, minimálpolinom

Wettl Ferenc
Algebra Tanszék

BUDAPESTI MŰSZAKI
ÉS GAZDASÁGTUDOMÁNYI
EGYETEM
MATEMATIKA
INTÉZET

2016. április 29.

Diagonalizálható mátrixok polinomjai

D Mátrix polinomja Ha $p(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ egy tetszőleges polinom, akkor értelmezhető e polinomnak egy tetszőleges négyzetes mátrixban fölvett értéke:

$$p(\mathbf{A}) = c_n \mathbf{A}^n + c_{n-1}\mathbf{A}^{n-1} + \cdots + c_1 \mathbf{A} + c_0 \mathbf{I}.$$

m $\mathbf{A} = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}^{-1} \rightsquigarrow \mathbf{A}^2 = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}^{-1}\mathbf{C}\boldsymbol{\Lambda}\mathbf{C}^{-1} = \mathbf{C}\boldsymbol{\Lambda}^2\mathbf{C}^{-1}$, általában tetszőleges nemnegatív k egészre $\mathbf{A}^k = \mathbf{C}\boldsymbol{\Lambda}^k\mathbf{C}^{-1}$.

$\rightsquigarrow$ bármely $p(x)$ polinomra $p(\mathbf{A}) = \mathbf{C}p(\boldsymbol{\Lambda})\mathbf{C}^{-1}$, ahol

$$p\left(\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\right) = \operatorname{diag}\left(p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)\right).$$

Á Diagonalizálható mátrix polinomja Legyen $\mathbf{A} = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}^{-1}$, ahol $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, és $p(x)$ egy tetszőleges polinom. Ekkor

$$p(\mathbf{A}) = \mathbf{C}\begin{bmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & p(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & p(\lambda_n) \end{bmatrix}\mathbf{C}^{-1}.$$

Példa

P Számítsuk ki az $\mathbf{A}^{10}$ mátrixot!

$$\mathbf{A} = \begin{bmatrix} -4 & 6 \\ -3 & 5 \end{bmatrix}$$

M Sajátfelbontása $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

Így $\mathbf{A}^{10} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^{10} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} =$

$$\begin{bmatrix} 2\cdot(-1)^{10} - 2^{10} & -2\cdot(-1)^{10} + 2\cdot 2^{10} \\ (-1)^{10} - 2^{10} & -(-1)^{10} + 2\cdot 2^{10} \end{bmatrix} = \begin{bmatrix} -1022 & 2046 \\ -1023 & 2047 \end{bmatrix}$$

Fibonacci-sorozat explicit alakja

T A Fibonacci-sorozatra ($F_0 = 0$, $F_1 = 1$, $F_{n+1} = F_n + F_{n-1}$, első néhány tagja: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987,...) igaz a következő:

$$F_n = \left(\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n\right)_{1,2} = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$$

B Az első egyenlőség teljes indukcióval bizonyítható. Jelölés:

$$\mathbf{F} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

$n = 1$-re trivi, $n \Rightarrow n+1$:

$$\mathbf{F}^{n+1} = \begin{bmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} F_n & F_{n-1} + F_n \\ F_{n+1} & F_n + F_{n+1} \end{bmatrix} = \begin{bmatrix} F_n & F_{n+1} \\ F_{n+1} & F_{n+2} \end{bmatrix}.$$

Fibonacci-sorozat (1. bizonyítás)

B $\chi_{\mathbf{F}}(\lambda) = \lambda^2 - \lambda - 1$, $\mathbf{F}$ sajátértékei $\lambda_{1,2} = \frac{1}{2}(1 \pm \sqrt{5})$, a hozzájuk tartozó sajátvektorok $\mathbf{x}_{1,2} = (1, \frac{1}{2}(1 \pm \sqrt{5}))$.

$\mathbf{F}^n$ sajátfelbontásával:

$$\mathbf{F}^n = \begin{bmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^n \begin{bmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{bmatrix}^{-1}$$

$$\begin{bmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{bmatrix}^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{\sqrt{5}-1}{2} & 1 \\ \frac{1+\sqrt{5}}{2} & -1 \end{bmatrix}$$

ahonnan

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right)$$

Fibonacci-sorozat (2. bizonyítás)

B Vegyük észre

$$\begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix} = \mathbf{F}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Az $\mathbf{x}_1$ és $\mathbf{x}_2$ sajátvektorok bázist alkotnak $\mathbb{R}^2$-ben, így a $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ előáll azok lineáris kombinációjaként:

$\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right] = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = 1/\sqrt{5}\mathbf{x}_1 - 1/\sqrt{5}\mathbf{x}_2$.

$\mathbf{F}^n\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right] = c_1\lambda_1^n\mathbf{x}_1 + c_2\lambda_2^n\mathbf{x}_2$, behelyettesítés után ezt kapjuk:

$$\mathbf{F}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n \begin{bmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{bmatrix} - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n \begin{bmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{bmatrix}.$$

Itt csak az első koordinátát kiszámolva, a tételbeli állítást igazoltuk.

Diagonalizálható mátrix karakterisztikus polinomja

Á (Cayley–Hamilton-tétel spec. eset) Ha $\chi_{\mathbf{A}}$ a diagonalizálható $\mathbf{A}$ karakterisztikus polinomja, akkor $\chi_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$.

B A λ sajátértékre $\chi_{\mathbf{A}}(\lambda) = 0$, így

$$\begin{aligned}
\chi_{\mathbf{A}}(\mathbf{A}) &= \mathbf{C}\,\chi_{\mathbf{A}}(\mathbf{\Lambda})\,\mathbf{C}^{-1} \\
&= \mathbf{C}\ \operatorname{diag}(\chi_{\mathbf{A}}(\lambda_1), \chi_{\mathbf{A}}(\lambda_2), \dots, \chi_{\mathbf{A}}(\lambda_n))\,\mathbf{C}^{-1} \\
&= \mathbf{O}.
\end{aligned}$$

m E tétel általánosan is igaz

Cayley–Hamilton-tétel

T Ha $\mathbf{A}$ egy tetszőleges négyzetes mátrix, melynek karakterisztikus polinomja $\chi_{\mathbf{A}}$, akkor $\chi_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$.

B L! $\mathbf{B} = \mathbf{A} - \lambda\mathbf{I}$. Az $\mathbf{A}$ karakterisztikus polinomja így $\det\mathbf{B} = \chi_{\mathbf{A}}(\lambda) = (-1)^n\lambda^n + p_{n-1}\lambda^{n-1} + \ldots + p_1\lambda + p_0$.
$\mathbf{B}$ bármely eleméhez tartozó előjeles aldeterminánás λ egy legföljebb $n-1$-edfokú polinomja, így léteznek olyan konstans elemű $\mathbf{C}_0$, $\mathbf{C}_1, \ldots$, $\mathbf{C}_{n-1}$ mátrixok, hogy

$$\operatorname{adj}\mathbf{B} = \lambda^{n-1}\mathbf{C}_{n-1} + \ldots + \lambda\mathbf{C}_1 + \mathbf{C}_0.$$

$\det(\mathbf{B})\mathbf{I} = \mathbf{B}\operatorname{adj}(\mathbf{B}) \rightsquigarrow$

$$\begin{aligned}\mathbf{B}\operatorname{adj}\mathbf{B} &= (\mathbf{A} - \lambda\mathbf{I})\left(\sum_{k=0}^{n-1}\lambda^k\mathbf{C}_k\right)\\ &= \mathbf{AC}_0 + \left(\sum_{k=1}^{n-1}\lambda^k(\mathbf{AC}_k - \mathbf{C}_{k-1})\right) - \lambda^n\mathbf{C}_{n-1}.\end{aligned}$$

Cayley–Hamilton-tétel

Együtthatóösszahasonlítás:

$$
\begin{array}{rcll}
\det(\mathbf{B})\mathbf{I} & & \mathbf{B}\,\mathrm{adj}\,\mathbf{B} & \\
\hline
(-1)^n\mathbf{I} &=& -\mathbf{C}_{n-1} & \cdot\,\mathbf{A}^n \\
p_{n-1}\mathbf{I} &=& \mathbf{A}\mathbf{C}_{n-1}-\mathbf{C}_{n-2} & \cdot\,\mathbf{A}^{n-1} \\
 & \vdots & & \vdots \\
p_2\mathbf{I} &=& \mathbf{A}\mathbf{C}_2-\mathbf{C}_1 & \cdot\,\mathbf{A}^2 \\
p_1\mathbf{I} &=& \mathbf{A}\mathbf{C}_1-\mathbf{C}_0 & \cdot\,\mathbf{A} \\
p_0\mathbf{I} &=& \mathbf{A}\mathbf{C}_0. &
\end{array}
$$

A beszorzás után kapott egyenlőségeket összeadva

$$(-1)^n\mathbf{A}^n+p_{n-1}\mathbf{A}^{n-1}+\ldots+p_1\mathbf{A}+p_0\mathbf{I}=\mathbf{O}.$$

Á Tetszőleges n-edrendű invertálható $\mathbf{A}$ mátrixhoz létezik olyan legföljebb $n-1$-edfokú q polinom, hogy $\mathbf{A}^{-1} = q(\mathbf{A})$.

D L! $\mathbf{A} \in \mathbb{F}^{n\times n}$. Minimálpolinomnak nevezünk egy olyan minimális fokszámú $\mu_{\mathbf{A}}$ főpolinomot (1 főegyütthatójú), melyre $\mu_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$.

D Azt mondjuk, hogy a p polinom az $\mathbf{A}$ mátrix annullátora, vagy hogy annullálja azt, ha $p(\mathbf{A}) = \mathbf{O}$. A minimálpolinom tehát egy legkisebb fokú annullátor főpolinom.

m A nullpolinom nem lehet minimálpolinom, mert nem 1 a főegyütthatója. Az 1 nem lehet minimálpolinom, mert bármely mátrix behelyettesítése után $\mathbf{I}$ lesz (nem $\mathbf{O}$).

m Az $\mathbf{I}$ mátrixnak (az identikus transzformációnak) $\mu(x) = x - 1$ a minimálpolinomja.

m $\mathbf{A} \sim \mathbf{B} \Rightarrow \mu_{\mathbf{A}} = \mu_{\mathbf{B}}$ (ugyanis $p(\mathbf{B}) = \mathbf{C}^{-1}p(\mathbf{A})\mathbf{C}$, így minden p polinomra $p(\mathbf{A})$ és $p(\mathbf{B})$ egyszerre $\mathbf{O}$, illetve egyszerre nem.)

m A minimálpolinom invariáns a mátrixok hasonlóságára, így egy lineáris transzformáció minimálpolinomja megegyezik bármely bázisban fölírt mátrixának minimálpolinomjával. Definálható egy $A : \mathcal{V} \to \mathcal{V}$ lineáris transzformáció minimálpolinomja, ahol $\mathcal{V}$ egy $\mathbb{F}$ test fölötti véges dimenziós vektortér.

Minimálpolinom tulajdonságai

T L! $\mathbf{A} \in \mathbb{F}^{n\times n}$.

1. $\mathbf{A}$-nak pontosan egy $\mu_{\mathbf{A}}$ minimálpolinomja van.
2. Bármely p polinomra $p(\mathbf{A}) = \mathbf{O} \iff \mu_{\mathbf{A}} \mid p$.
3. $\mu_{\mathbf{A}} \mid \chi_{\mathbf{A}}$
4. $\mathbf{A}$ minden sajátértéke gyöke $\mu_{\mathbf{A}}$-nak.

B 1 A C–H-tétel $\forall$ $\mathbf{A}$: $\chi_{\mathbf{A}}(\mathbf{A}) = \mathbf{O} \rightsquigarrow \exists$ anullátor $\rightsquigarrow \exists$ főpolinom

Tfh p és q két különböző minimális fokszámú főpolinom, melyekre $p(\mathbf{A}) = q(\mathbf{A}) = \mathbf{O} \rightsquigarrow (p-q)(\mathbf{A}) = p(\mathbf{A}) - q(\mathbf{A}) = \mathbf{O} \rightsquigarrow p = q$.

2 $\mu_{\mathbf{A}} \mid p \rightsquigarrow p = \mu_{\mathbf{A}} q$ vmilyen q pol.-ra $\rightsquigarrow p(\mathbf{A}) = \mu_{\mathbf{A}}(\mathbf{A})q(\mathbf{A}) = \mathbf{O}$.

$p(\mathbf{A}) = \mathbf{O}$ és $p = \mu_{\mathbf{A}} q + r$, ahol r foka kisebb, mint $\mu_{\mathbf{A}}$ foka, másrészt $\mathbf{O} = p(\mathbf{A}) = \mu_{\mathbf{A}}(\mathbf{A})q(\mathbf{A}) + r(\mathbf{A}) \rightsquigarrow r(\mathbf{A}) = \mathbf{O} \rightsquigarrow r = 0$.

3 az előzőekből

4 $(\lambda, \mathbf{x})$ sajátpár $\rightsquigarrow \mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}$ $(k \in \mathbb{N}^+)$ $\rightsquigarrow \forall$ p-re $p(\mathbf{A})\mathbf{x} = p(\lambda)\mathbf{x}$.

$\mu_{\mathbf{A}}(\mathbf{A})\mathbf{x} = \mu_{\mathbf{A}}(\lambda)\mathbf{x}$, de $\mu_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$, és $\mathbf{x} \neq \mathbf{0}$, ezért $\mu_{\mathbf{A}}(\lambda) = 0$.

- **m** Ha $\chi_{\mathbf{A}}(x) = \prod_i (x - \lambda_i)^{a_i}$, akkor a 3-beli oszthatóság miatt $\mu_{\mathbf{A}}(x) = \prod_i (x - \lambda_i)^{m_i}$, ahol $1 \leqslant m_i \leqslant a_i$, és a_i a λ_i algebrai multiplicitása.
- **m** Ha $\mathbf{A}$ nilpotens, ahol $\mathbf{A}^k = \mathbf{O}$, de $\mathbf{A}^{k-1} \neq \mathbf{O}$, akkor $\mu_{\mathbf{A}}(x) = x^k$, ugyanis x^k annullátor, így a minimálpolinom csak valamely osztója lehet. Az osztói viszont mind x^m alakúak, ahol $m \leqslant k$, de azok $m < k$ esetén nem annullátorok.
- **Á** $\mathbf{A}$ pontosan akkor diagonalizálható, ha minimálpolinomja különböző lineáris tényezők szorzata (egyelőre az egyik irányt tudjuk bizonyítani).

Példák minimálpolinomra

P Határozzuk meg a $\chi_{\mathbf{A}}$ és $\mu_{\mathbf{A}}$ polinomokat!

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

M $\chi_{\mathbf{A}}(x) = \mu_{\mathbf{A}}(x) = x^6$.

Példák minimálpolinomra

P

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

M $\chi_{\mathbf{A}}(x) = \chi_{\mathbf{B}}(x) = (x-1)^4$ Minimálpolinom lehet $(x-1)^k$, ahol $k \leqslant 4$. Mivel

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{B}^2 = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

és $\mathbf{A}^2 - 2\mathbf{A} + \mathbf{I} = \mathbf{O}$, de $\mathbf{A} - \mathbf{I} \neq \mathbf{O}$, ezért $\mu_{\mathbf{A}}(x) = (x-1)^2$.
$\mathbf{B}^2 - 2\mathbf{B} + \mathbf{I} = \mathbf{O}$, de $\mathbf{B} - \mathbf{I} \neq \mathbf{O}$, ezért $\mu_{\mathbf{B}}(x) = (x-1)^2$.

Példák minimálpolinomra

P Írjuk föl a következő mátrix minimálpolinomját!

$$\mathbf{M} = \begin{bmatrix} 4 & 3 & 5 & 6 \\ 2 & -1 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 1 \end{bmatrix}$$

M Legyen

$$\mathbf{A} = \begin{bmatrix} 4 & 3 \\ 2 & -1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$$

Ekkor $\chi_{\mathbf{A}}(x) = x^2 - 3x - 10 = (x-5)(x+2)$,
$\chi_{\mathbf{B}}(x) = x^2 - 3x - 4 = (x-4)(x+1)$, így
$\chi_{\mathbf{M}}(x) = (x-4)(x+1)(x-5)(x+2)$, de mivel nincsenek többszörös gyökök, ez egyúttal a minimálpolinom is.

Frobenius kísérő mátrix

Á Bármely $\chi(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n$ polinomhoz létezik olyan mátrix, melynek χ a karakterisztikus polinomja. Ezt a polinom kísérő mátrixának nevezzük. Alakja

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_n \\ 1 & 0 & \dots & 0 & -a_{n-1} \\ 0 & 1 & \dots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix}$$

B A karakterisztikus polinomot adó determinánsban alulról minden sor x-szeresét a fölötte lévőhöz adva kapjuk, hogy

$$\chi_{\mathbf{C}}(x) = \begin{vmatrix} -x & 0 & \dots & 0 & -a_n \\ 1 & -x & \dots & 0 & -a_{n-1} \\ 0 & 1 & \dots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_1 - x \end{vmatrix} = \begin{vmatrix} 0 & 0 & \dots & 0 & \chi(x) \\ 1 & 0 & \dots & 0 & ? \\ 0 & 1 & \dots & 0 & ? \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_1 - x \end{vmatrix}$$

Így $\chi_{\mathbf{C}}(x)$ előjelszorzót nem számítva megegyezik a megadott $\chi(x)$ polinommal.

Á A kísérő mátrix minimálpolinomja megegyezik karakterisztikus polinomjával (egy -1 szorzó erejéig).

B Tetszőleges, de nem csupa zérus c_j konstansokra

$$\left(\sum_{j=0}^{n-1} c_j \mathbf{C}^j\right)\mathbf{e}_1 = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix},$$

azaz nincs n-nél alacsonyabb fokú annullátor, tehát
$\mu(\mathbf{C}) = (-1)^n \chi(\mathbf{C})$.