Bevezetés az algebrába 2 – Mátrixfüggvények

Wettl Ferenc
Algebra Tanszék

BUDAPESTI MŰSZAKI
MATEMATIKA
ÉS GAZDASÁGTUDOMÁNYI
INTÉZET
EGYETEM

2016. április 28.

m Ha $f(x)=\sum_{k=0}^{\infty} a_k x^k$ és $\mathbf{D}$ diagonális, továbbá $\mathbf{D}$ főátlóbeli elemei benne vannak a hatványsor konvergenciatartományában, akkor

$$\begin{aligned} f(\mathbf{D}) = \sum_{k=0}^{\infty} a_k \mathbf{D}^k &= \operatorname{diag}\left(\sum_{k=0}^{\infty} a_k d_1^k, \dots, \sum_{k=0}^{\infty} a_k d_1^k\right) \\ &= \operatorname{diag}(f(d_1), \dots, f(d_n)). \end{aligned}$$

Eszerint például bármely diagonalizálható $\mathbf{A}$ mátrixra értelmezhető az $\mathrm{e}^{\mathbf{A}}$ hatvány, nevezetesen

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \dots + \frac{\mathbf{A}^n}{n!} + \dots$$

Hasonlóképp definiálható az $\ln(\mathbf{I}+\mathbf{A})$ mátrixfüggvény is. Fölhasználva a

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \qquad |x|<1$$

hatványsort kapjuk, hogy

$$\ln(\mathbf{I}+\mathbf{A}) = \mathbf{A} - \frac{\mathbf{A}^2}{2} + \frac{\mathbf{A}^3}{3} - \frac{\mathbf{A}^4}{4} + \dots,$$

- Egy hatványsorba fejthető függvénynek egy diagonális mátrixban – és így bármely diagonalizálható mátrixban – fölvett értékét a függvénynek csak a sajátértékekben való viselkedése befolyásolja.
- A Cayley–Hamilton-tétel szerint minden mátrix kielégíti saját karakterisztikus egyenletét, így egy n-edrendű mátrix minden hatványa legföljebb $n-1$-edik hatványok lineáris kombinációjával helyettesíthető, azaz a függvény értéke egy polinomba való helyettesítéssel is kiszámolható.

Á Legyen az $\mathbf{A} \in \mathbb{C}^{n\times n}$ mátrix Jordan-felbontása $\mathbf{A} = \mathbf{CJC}^{-1}$ és $p \in \mathbb{C}[x]$ egy tetszőleges polinom. Ekkor

$$p(\mathbf{A}) = \mathbf{C}p(\mathbf{J})\mathbf{C}^{-1} = \mathbf{C}\begin{bmatrix} p(\mathbf{J}_1) & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & p(\mathbf{J}_2) & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & p(\mathbf{J}_k) \end{bmatrix}\mathbf{C}^{-1},$$

m Tegyük fel, hogy az f függvény λ körül Taylor-sorba fejthető, azaz

$$f(x) = f(\lambda) + f'(\lambda)(x-\lambda) + \ldots + \frac{f^{(m)}(\lambda)}{m!}(x-\lambda)^m + \ldots$$

és legyen $\mathbf{J} \in \mathbb{C}^{n\times n}$ egy Jordan-blokk, azaz

$$\mathbf{J} = \lambda\mathbf{I} + \mathbf{N} = \begin{bmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & \ldots \\ 0 & 0 & 1 & \ldots \\ 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix} = \begin{bmatrix} \lambda & 1 & 0 & \ldots \\ 0 & \lambda & 1 & \ldots \\ 0 & 0 & \lambda & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

- Mivel $\mathbf{N}^n = \mathbf{O}$, fenn kell álljon az

$$f(\mathbf{J}) = f(\lambda\mathbf{I} + \mathbf{N}) = f(\lambda)\mathbf{I} + f'(\lambda)\mathbf{N} + \ldots + \frac{f^{(n-1)}(\lambda)}{(n-1)!}\mathbf{N}^{n-1} \qquad (1)$$

összefüggés – ha egyáltalán van értelme az $f(\mathbf{J})$ kifejezésnek.
Tehát az f függvénynek csak a Jordan-mátrix rendjénél kisebb rendű deriváltjai játszanak szerepet a függvényértékben.

Spektrumon definiált függvény

D Legyen az **A** mátrix spektruma $\{\lambda_1, \dots, \lambda_k\}$, a λ_i sajátértékhez tartozó legnagyobb Jordan-blokk rendjét jelölje m_i. Azt mondjuk, hogy f definiálva van az **A** spektrumán, ha az

$$f^{(j)}(\lambda_i), \quad j = 0, 1, \dots, m_i - 1, \; i = 1, \dots, k$$

értékek léteznek. Azt mondjuk, hogy ezek az értékek az f értékei az **A** spektrumán.

m Minden függvény, mely **C** minden pontjában akárhányszor differenciálható, tetszőleges mátrixra értelmezve van annak spektrumán. Így minden polinom értelmezve van minden mátrix spektrumán, ami összhangban lesz azzal, hogy minden négyzetes mátrixnak bármely polinomfüggvénye értelmezve van.

m Ha μ az **A** mátrix minimálpolinomja, akkor μ értelmezve van **A** spektrumán, és μ értékei **A** spektrumán mind nullák. Ez egyből következik a minimálpolinom előállításából és a fenti definícióból.

Mátrixfüggvény a Jordan-alakból

D Legyen $\mathbf{A} \in \mathbb{C}^{n\times n}$ Jordan-felbontása $\mathbf{A} = \mathbf{CJC}^{-1}$, ahol $\mathbf{J} = \operatorname{diag}(\mathbf{J}_1, \dots, \mathbf{J}_k)$ a Jordan-féle normálalakja, és n_i jelöli a $\mathbf{J}_i$ blokk rendjét. Ekkor

$$f(\mathbf{A}) = \mathbf{C}f(\mathbf{J})\mathbf{C}^{-1} = \mathbf{C}\operatorname{diag}(f(\mathbf{J}_1), \dots, f(\mathbf{J}_k))\mathbf{C}^{-1},$$

ahol

$$f(\mathbf{J}_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{f''(\lambda_i)}{2!} & \dots & \frac{f^{(n_i-2)}(\lambda_i)}{(n_i-2)!} & \frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!} \\ 0 & f(\lambda_i) & f'(\lambda_i) & \dots & \dots & \frac{f^{(n_i-2)}(\lambda_i)}{(n_i-2)!} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & f'(\lambda_i) & \frac{f''(\lambda_i)}{2!} \\ 0 & 0 & 0 & \dots & f(\lambda_i) & f'(\lambda_i) \\ 0 & 0 & 0 & \dots & 0 & f(\lambda_i) \end{bmatrix} \tag{2}$$

- Egyszerű képletbehelyettesítéssel $f(x) = x^3$ esetén

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}^3 = \begin{bmatrix} f(2) & f'(2) \\ 0 & f(2) \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 0 & 8 \end{bmatrix}.$$

- Az $f(x) = \mathrm{e}^x$ függvény esetén, ha

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad \text{akkor} \quad \mathrm{e}^{\mathbf{A}} = \begin{bmatrix} e^2 & e^2 & \frac{e^2}{2} \\ 0 & e^2 & e^2 \\ 0 & 0 & e^2 \end{bmatrix} = e^2 \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

- Általában a λ-hoz tartozó Jordan-blokkra

$$\mathbf{J} = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix} \quad \text{esetén } e^{\mathbf{J}} = e^{\lambda} \begin{bmatrix} 1 & 1 & \frac{1}{2!} & \dots & \frac{1}{(n-1)!} \\ 0 & 1 & 1 & \dots & \frac{1}{(n-2)!} \\ 0 & 0 & 1 & \dots & \frac{1}{(n-3)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Mátrix exponenciális függvénye

P Legyen

$$\mathbf{A} = \begin{bmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & -2 & -5 \end{bmatrix}$$

Határozzuk meg az $e^{\mathbf{A}}$ mátrixot!

M **A** karakterisztikus polinomja

$$x^3 + 10x^2 + 32x + 32 = (x+2)(x+4)^2,$$

$$\mathbf{J} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -2 \end{bmatrix},$$

$$e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1} = \frac{1}{2e^4}\begin{bmatrix} e^2+1 & 2e^2-2 & e^2-1 \\ e^2-1 & 2e^2 & e^2-1 \\ 1-e^2 & 2-2e^2 & 3-e^2 \end{bmatrix}$$

P Határozzuk meg az $\mathbf{A} = \begin{bmatrix} 2 & 3 & 9 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$ mátrix Jordan-féle normálalakját, $\mathbf{J}$-t, és az $\mathbf{A}^{100}$, $e^{\mathbf{J}}$, $e^{3\mathbf{A}}$ mátixokat.

M $\chi(\lambda) = (\lambda - 2)^2(\lambda + 5)$. A 2-höz tartozó s.v.: $(1, 0, 0)$, a -5-höz tartozó $(-9/7, 0, 1) \rightsquigarrow \mathbf{J} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$.

A 2-höz tartozó másik általánosított sajátvektor:

$$(\mathbf{A} - 2\mathbf{I})\mathbf{x}_2 = \mathbf{x}_1, \quad \text{azaz} \quad \begin{bmatrix} 0 & 3 & 9 \\ 0 & 0 & 0 \\ 0 & 0 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Ennek egy megoldása $\mathbf{x}_2 = (0, \frac{1}{3}, 0) \rightsquigarrow$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & -9/7 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{C}^{-1} = \begin{bmatrix} 1 & 0 & 9/7 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Mivel $(x^{100})' = 100x^{99}$, $(e^x)' = e^x$, $(e^{3x})' = 3e^{3x}$, ezért

$$\mathbf{J}^{100} = \begin{bmatrix} 2^{100} & 100 \cdot 2^{99} & 0 \\ 0 & 2^{100} & 0 \\ 0 & 0 & 5^{100} \end{bmatrix}, \quad e^{\mathbf{J}} = \begin{bmatrix} e^2 & e^2 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^{-5} \end{bmatrix},$$

$$e^{3\mathbf{J}} = \begin{bmatrix} e^6 & 3e^6 & 0 \\ 0 & e^6 & 0 \\ 0 & 0 & e^{-15} \end{bmatrix}.$$

Innen az $\mathbf{A}^{100} = \mathbf{C}\mathbf{J}^{100}\mathbf{C}^{-1}$ és $e^{3\mathbf{A}} = \mathbf{C}e^{3\mathbf{J}}\mathbf{C}^{-1}$ felhasználásával

$$\mathbf{A}^{100} = \begin{bmatrix} 2^{100} & 100 \cdot 3 \cdot 2^{99} & \frac{9}{7}(2^{100} - 5^{100}) \\ 0 & 2^{100} & 0 \\ 0 & 0 & 5^{100} \end{bmatrix},$$

$$e^{3\mathbf{A}} = \begin{bmatrix} e^6 & 9e^6 & \frac{9}{7}(e^6 - e^{-15}) \\ 0 & e^6 & 0 \\ 0 & 0 & e^{-15} \end{bmatrix}.$$

Spektrumon azonos értékeket adó polinomok

Á Tetszőleges p és q polinomokra és $\mathbf{A} \in \mathbb{C}^{n \times n}$ mátrixra $p(\mathbf{A}) = q(\mathbf{A})$, pontosan akkor teljesül, ha p és q értékei $\mathbf{A}$ spektrumán azonosak.

B Ha $p(\mathbf{A}) = q(\mathbf{A})$, akkor $h = p - q$ annullálja $\mathbf{A}$-t, így h osztható a minimálpolinommal, így a minimálpolinommal együtt h értékei is nullák az $\mathbf{A}$ spektrumán.

Ha p és q értékei $\mathbf{A}$ spektrumán azonosak, akkor a $h = p - q$ polinom értékei mind nullák. Az ilyen polinomok alakja $\prod_{i=1}^{s}(x - \lambda_i)^{m_i} g(x)$, azaz $h = \mu g$, tehát h annullálja $\mathbf{A}$-t, így $p(\mathbf{A}) = q(\mathbf{A})$.

Mátrixfüggvény kiszámítása polinominterpolációval

D Legyen $\mathbf{A}$ minimálpolinomja $\mu_{\mathbf{A}}$, és tegyük fel, hogy az f függvény definiálva van $\mathbf{A}$ spektrumán. Ekkor $f(\mathbf{A}) := p(\mathbf{A})$, ahol p az a polinom, melynek foka kisebb $\mu_{\mathbf{A}}$ fokánál, és amely eleget tesz a

$$p^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \quad j = 0, 1, \dots, m_i - 1,\ i = 1, \dots, k \tag{3}$$

feltételeknek, ahol m_i a λ_i sajátértékhez tartozó legnagyobb Jordan-blokk rendjét jelöli.

m A definícióban megadott polinom egyértelműen létezik, ezt nevezzük Hermite-polinomnak, mely explicit módon is megadható:

$$p(x) = \sum_{i=1}^{s} \left(\left(\sum_{j=0}^{m_i-1} \left(\frac{f(y)}{\prod\limits_{k\neq i} (y-\lambda_k)} \right)^{(j)} (\lambda_i) \frac{(x-\lambda_i)^j}{j!} \right) \prod_{j\neq i} (x-\lambda_j)^{m_j} \right).$$

m Ha **A**-nak minden sajátértéke egyszeres algebrai multiplicitású, azaz $s = n$ és $m_i = 1$ minden i-re, akkor az előző formula az ismert Lagrange-féle interpolációs polinomot adja:

$$p(x) = \sum_{i=1}^{n} \left(f(\lambda_i) \prod_{j\neq i} \frac{x-\lambda_j}{\lambda_i-\lambda_j} \right). \tag{4}$$

Ha pedig **A**-nak csak egyetlen sajátértéke λ, melynek n az algebrai multiplicitása, azaz $s = 1$, $m_1 = n$, akkor f Taylor-polinomját kapjuk:

$$p(x) = \sum_{j=0}^{n-1} f^{(j)}(\lambda) \frac{(x-\lambda)^j}{j!}.$$

- Az Hermite-polinom ugyan egyértelmű, de nem mindig tudjuk könnyen meghatározni, például ha a mátrixnak csak a sajátértékeit ismerjük, de a legnagyobb Jordan-blokk méretét nem. A korábbiak szerint bármely más polinom is megfelel, mely kielégíti a (3) feltételeket.
- Nézzük az $f(x) = x^3$ függvény helyettesítési értékét a $\mathbf{A} = \left[\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right]$ mátrixban. Ugyan f polinom, de mivel $\mu_{\mathbf{A}}(x) = (x-2)^2$, azaz a minimálpolinom kisebb fokú, ezért van olyan elsőfokú polinom is, mely $\mathbf{A}$-ban azonos értéket ad. E polinom az $x^3 : \mu_{\mathbf{A}}(x)$ osztás maradéka. Mivel $x^3 = (x-2)^2(x+4) + (12x - 16)$, azaz a maradék $12x - 16$, ezért

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}^3 = 12 \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} - 16 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 0 & 8 \end{bmatrix}.$$

m Keressük meg az $f(x) = x^3$ Hermite-féle interpolációs polinomját

$$\begin{aligned} f(2) &= 8 = p(2) = 2a + b \\ f'(2) &= 12 = p'(2) = a. \end{aligned} \quad \rightsquigarrow \quad a = 12, b = -16.$$

- $\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, $e^{\mathbf{A}} =?$

$\mu_{\mathbf{A}}(x) = (x-2)^3 \rightsquigarrow$ Hermite-polinom: $p(x) = ax^2 + bx + c$

$$\begin{aligned} e^x|_2 &= e^2 = p(2) = 4a + 2b + c \\ (e^x)'|_2 &= e^2 = p'(2) = 4a + b \\ (e^x)''|_2 &= e^2 = p''(2) = 2a. \end{aligned} \quad \rightsquigarrow \quad a = e^2/2, b = -e^2, c = e^2, \rightsquigarrow$$

$$\begin{aligned} e^{\mathbf{A}} &= p(\mathbf{A}) = \frac{e^2}{2}\mathbf{A}^2 - e^2\mathbf{A} + e^2\mathbf{I} \\ &= \frac{e^2}{2}\begin{bmatrix} 4 & 4 & 1 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix} - e^2\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} + e^2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = e^2\begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Exponenciális függvény Hermite-polinommal

P Számítsuk ki az $e^{\mathbf{A}}$ mátrixot ha

$$\mathbf{A} = \begin{bmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & -2 & -5 \end{bmatrix}$$

M $\chi_{\mathbf{A}}(x) = (x+2)(x+4)^2$, Jordan-alakja $\operatorname{diag}(-2,-4,-4)$, a minimálpolinom $\mu_{\mathbf{A}}(x) = (x+2)(x+4) = x^2+6x+8$. olyan elsőfokú $p(x) = ax+b$ alakú polinomot keresünk, melyre

$$\mathrm{e}^{-2} = p(-2) = -2a+b$$
$$\mathrm{e}^{-4} = p(-4) = -4a+b.$$

Innen $a = \frac{1}{2}(\mathrm{e}^{-2}-\mathrm{e}^{-4})$, $b = 2\mathrm{e}^{-2}-\mathrm{e}^{-4}$, így

$$\mathrm{e}^{\mathbf{A}} = a\mathbf{A} + b\mathbf{I} = \frac{1}{2\mathrm{e}^4}\begin{bmatrix} \mathrm{e}^2+1 & 2\mathrm{e}^2-2 & \mathrm{e}^2-1 \\ \mathrm{e}^2-1 & 2\mathrm{e}^2 & \mathrm{e}^2-1 \\ 1-\mathrm{e}^2 & 2-2\mathrm{e}^2 & 3-\mathrm{e}^2 \end{bmatrix}.$$

P $\mathbf{A} = \begin{bmatrix} 2 & -3 & 2 \\ 4 & 10 & -4 \\ 4 & 6 & 0 \end{bmatrix}$, $\mathrm{e}^{\mathbf{A}} = ?$

M $\chi(x) = (4-x)^3$, $\mu(x) = (4-x)^2$ (korábbi feladatból)

Keresünk egy $p(x) = ax + b$ polinomot, melyre

$$\begin{aligned} \exp(4) = \mathrm{e}^4 = p(4) = 4a + b \\ \exp'(4) = \mathrm{e}^4 = p'(4) = a \end{aligned} \quad \rightsquigarrow \quad a = \mathrm{e}^4,\ b = -3\mathrm{e}^4$$

- Tehát

$$\mathrm{e}^{\mathbf{A}} = \mathrm{e}^4\mathbf{A} - 3\mathrm{e}^4\mathbf{I} = \mathrm{e}^4 \begin{bmatrix} -1 & -3 & 2 \\ 4 & 7 & -4 \\ 4 & 6 & -3 \end{bmatrix}$$

A definíciók ekvivalenciája

T A mátrixfüggvény kiszámítására adott fenti két definíció ekvivalens.

B A „Hermite”-definíció szerint $f(\mathbf{A}) = p(\mathbf{A})$ és bármely más polinom is megadja $f(\mathbf{A})$-t, ha kielégíti a (3) feltételeket.

Ha $\mathbf{A} \sim \mathbf{J}$, akkor $f(\mathbf{A}) = p(\mathbf{A}) = p(\mathbf{CJC}^{-1}) = \mathbf{C}p(\mathbf{J})\mathbf{C}^{-1}$, ezért elég a Jordan-alakokra ellenőrizni az ekvivalenciát.

A „Jordan”-definícióban f-nek épp azok a deriváltjai szerepelnek azokban a sajátértékekben kiértékelve, amelyek a (3) feltételekben is szerepelnek. Így elég csak azt ellenőrizni, hogy egy Jordan-blokk Hermite-polinomja megegyezik-e a „Jordan”-definícióban szereplővel. Ezt a polinom Taylor-polinomjára fölírt (1) képlet igazolja.