II Matriisieksponenttiaali

4. Eksponenttiaali ja logaritmi

Eksponenttifunktio $e^{\cdot}: \mathbb{R}\to\mathbb{R}$ ja logaritmi $\log: \mathbb{R}_+\to\mathbb{R}$ yleistetään matriiseille sarjakehitelmiensä kautta. Tässä täytyy kuitenkin olla varovainen sarjojen suppenemisen kanssa. Täsmennetään tämän takia kurssilla käytettävän matriisinormin käsite ja ominaisuuksia.

Määr 4.1 Matriisien operaattorinormi on normi

$$\|\cdot\|: M_{n\times m}(\mathbb{K})\to\mathbb{R} \qquad \|A\| = \sup_{x\in\mathbb{K}^m\setminus\{0\}} \frac{\|Ax\|_{\mathbb{K}^n}}{\|x\|_{\mathbb{K}^m}} = \sup_{\substack{x\in\mathbb{K}^n \\ \|x\|_{\mathbb{K}^n}=1}} \|Ax\|_{\mathbb{K}^n}$$

Lause 4.2

(i) $\|tA\| = |t|\cdot\|A\|$ $\forall t\in\mathbb{K}$ ja $A\in M_{n\times m}(\mathbb{K})$

(ii) $\|A+B\| \le \|A\|+\|B\|$ $\forall A\in M_{n\times m}(\mathbb{K})$ ja $B\in M_{n\times m}(\mathbb{K})$

(iii) $\|AB\| \le \|A\|\cdot\|B\|$ $\forall A\in M_{n\times m}(\mathbb{K})$ ja $B\in M_{m\times p}(\mathbb{K})$

(iv) $\|A_k - A\|\to 0 \iff A_k\to A$ komponenteittain

Tod

HT

TO 25.1
2

Reaaliselle eksponenttifunktiolla on (0:ssa) sarjakehitelmä

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

joka suppenee kaikilla $x \in \mathbb{R}$ itseisesti.

Logaritmilla taas on (1:ssä) sarjakehitelmä

$$\log(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x-1)^k$$

joka suppenee itseisesti kun $|x-1|<1$.

Tarkastellaan vastaavia sarjoja matriiseille $A \in M_n(\mathbb{K})$.

Lemma 4.3

(i) $\sum_{k=0}^{\infty} \frac{A^k}{k!}$ suppenee $\forall A \in M_n(\mathbb{K})$

(ii) $\left(\sum_{k=0}^{\infty} \frac{(tA)^k}{k!}\right) \cdot \left(\sum_{k=0}^{\infty} \frac{(sA)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{((t+s)A)^k}{k!}$

(iii) $\sum_{k=0}^{\infty} \frac{A^k}{k!} = \hat{A}$ on kääntyvä ja $(\hat{A})^{-1} = \sum_{k=0}^{\infty} \frac{(-A)^k}{k!}$

Tod

(i) Sarja $\sum_{k=0}^{\infty} \frac{A^k}{k!}$ suppenee jos se suppenee itseisesti, eli $\sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\|$ suppenee.

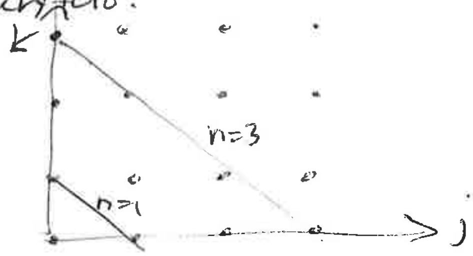
Matriisinormin ominaisuuksien nojalla

$$\sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\| \leq \sum_{k=0}^{\infty} \left| \frac{1}{k!} \right| \|A^k\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k = e^{\|A\|}$$

joten sarjan suppeneminen seuraa reaalien eksponenttifunktion sarjakehitelmän suppenemisesta.

(ii) Itseisesti suppenevien sarjojen tulon antaa Cauchy-tulo:

$$\left(\sum_{k=0}^{\infty} B_k\right)\left(\sum_{j=0}^{\infty} C_j\right) = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} B_i C_{n-i}\right)$$



Näin ollen

$$\left(\sum_{k=0}^{\infty} \frac{(tA)^k}{k!}\right)\left(\sum_{j=0}^{\infty} \frac{(sA)^j}{j!}\right) = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} \frac{t^i s^{n-i}}{i!(n-i)!} A^i A^{n-i}\right)$$

$$= \sum_{n=0}^{\infty} \left(\frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} t^i s^{n-i}\right) A^n = \sum_{n=0}^{\infty} \frac{(t+s)^n}{n!} A^n$$

(iii) Soveltamalla (ii)-kohtaa kertoimilla $t=1$ ja $s=-1$ nähdään, että

$$\left(\sum_{k=0}^{\infty} \frac{A^k}{k!}\right)\left(\sum_{j=0}^{\infty} \frac{(-A)^j}{j!}\right) = \left(\sum_{n=1}^{\infty} \frac{(A+(-A))^n}{n!}\right) + I = I \qquad \square$$

Lemman 4.3 nojalla voidaan määritellä

Määr 4.4 Matriisien eksponenttifunktio on

$$\exp: M_n(\mathbb{K}) \to GL(n,\mathbb{K}), \quad \exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

Lemma 4.3 sanoo tällöin että $\forall t,s \in \mathbb{K}$ ja $A \in M_n(\mathbb{K})$

$$\exp(tA)\exp(sA) = \exp((t+s)A) \quad \text{ja}$$
$$\exp(A)^{-1} = \exp(-A)$$

Matriisikertolaskun epäkommutatiivisuuden takia sen sijaan $\exp(A)\exp(B)$ ei välttämättä ole sama kuin $\exp(A+B)$.

Lemma 4.5
Jos $AB=BA$, niin $\exp(A+B) = \exp(A)\exp(B) = \exp(B)\exp(A)$.

Tod
Jos $AB=BA$, niin
$$(A+B)^n = \sum_{i=0}^{n} \binom{n}{i} A^i B^{n-i}$$
ja Lemman 4.3 (ii) todistus voidaan toistaa. $\square$

Lemma 4.6

Jos $\|A-I\|<1$, niin $\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(A-I)^k$ suppenee itseisesti.

Tod

$$\sum_{k=1}^{\infty} \left\|\frac{(-1)^{k-1}}{k}(A-I)^k\right\| \le \sum_{k=1}^{\infty} \frac{1}{k}\|A-I\|^k \le \sum_{k=1}^{\infty} \|A-I\|^k$$

ja tämä potenssisarja suppenee jos $\|A-I\|<1$. □

Määr 4.7 Matriisilogaritmi on

$$\log : \underbrace{B_{\|\cdot\|}(I,1)}_{=\{A\in M_n(\mathbb{K}) :\ \|A-I\|<1\}} \to M_n(\mathbb{K}), \quad \log(A) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(A-I)^k$$

Lemma 4.8

$\exp\circ\log = id$ ja $\log\circ\exp = id$ kun ne ovat määriteltyjä. Toisin sanoen,

(i) Jos $\|A-I\|<1$, niin $\exp\circ\log(A) = A$

(ii) Jos $\|\exp(A)-I\|<1$, niin $\log\circ\exp(A) = A$

Tod

Sivuutetaan. Heuristisesti: reaaliset eksponentiaali ja logaritmi ovat toistensa käänteisfunktiot joten myös sarjojen suppenemissäteiden sisällä $\forall x\in\mathbb{R}$, $|x-1|<1$

$$e^{\log x} = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(x-1)^k\right)^n = x$$

Matriisien A ja $-I$ kommutoinnin nojalla vastaavan osoittaminen matriiseille onnistuu.

Ehto $\|\exp(A)-I\|<1$ toteutuu kun $\|A\|<\log 2$:

$$\|\exp(A)-I\| = \left\|\sum_{k=1}^{\infty} \frac{1}{k!}A^k\right\| \le \sum_{k=1}^{\infty} \frac{\|A\|^k}{k!} = e^{\|A\|}-1 < e^{\log 2}-1 = 1.$$

Eksponentiaalien määrittäminen

Lemma 4.9

$\exp(BAB^{-1}) = B\exp(A)B^{-1}$ $\forall A\in M_n(\mathbb{K})$, $B\in GL(n,\mathbb{K})$

Tod

HT

(Blokki)diagonaalimatriisille $A = \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_n \end{bmatrix}$

$$A^k = \begin{bmatrix} B_1^k & & \\ & \ddots & \\ & & B_n^k \end{bmatrix} \Rightarrow \exp(A) = \begin{bmatrix} \exp(B_1) & & \\ & \ddots & \\ & & \exp(B_n) \end{bmatrix}$$

Näin ollen jos $A\in M_n(\mathbb{K})$ on diagonalisoituva, eli $A = UDU^{-1}$ joillekin $U\in GL(n,\mathbb{K})$ ja $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$,

$$\exp(A) = \exp(UDU^{-1}) = U\exp D\cdot U^{-1} = U\cdot\operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})\, U^{-1}$$

Yleisesti matriisit eivät ole diagonalisoituvia, missä tapauksessa puhutaan yleisemmästä Jordan-muodon käsitteestä.
Kuitenkin monet tärkeät erityistapaukset (esim. symmetriset ja hermiittiset matriisit) ovat diagonalisoituvia.

Lause 4.10

Matriisi $A\in M_n(\mathbb{K})$ on diagonalisoituva $\Leftrightarrow$ on olemassa kanta $x_1, \dots, x_n \in \mathbb{K}^n$ ominaisvektoreita $Ax_k = \lambda x_k$.

Tod

"$\Rightarrow$" Olkoon $A = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}U^{-1}$, $U\in GL(n,\mathbb{K})$.

Olkoot $x_1, \dots, x_n$ matriisin U sarakkeet. Tällöin $\{x_1, \dots, x_n\}$ on kanta ja

$$Ax_j = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}U^{-1}x_j = U\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}e_j = U\lambda_j e_j = \lambda_j x_j,$$

joten vektorit $x_1, \dots, x_n$ ovat ominaisvektoreita.

"$\Leftarrow$" Asetetaan $U = [x_1 \dots x_n]$. Tällöin

$$U^{-1}AUe_j = U^{-1}Ax_j = U^{-1}\lambda_j x_j = \lambda_j e_j,$$

joten $U^{-1}AU$ on diagonaalimatriisi. $\square$

To 25.1

G

Esim 4.11

Määritetään $\exp(A)$ matriisille $A = \begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & 0 \\ 2 & -2 & -1 \end{bmatrix}$

Ominaisarvot $\det(A-\lambda I) \overset{\text{lasku}}{=} -\lambda(\lambda-1)^2 = 0 \Leftrightarrow \lambda=0$ tai $\lambda=1$.

Ominaisvektorit:

$\lambda=0$: $\quad Ax=0 \Leftrightarrow \begin{cases} x_2=0 \\ 2x_1-x_3=0 \end{cases} \overset{\text{valitaan kantavek.}}{\Longrightarrow} u_1=(1,0,2)$

$\lambda=1$: $\quad (A-I)x=0 \Leftrightarrow x_1-x_2-x_3=0 \overset{\text{valitaan kanta}}{\Longrightarrow} u_2=(1,1,0)$, $u_3=(1,0,1)$

Muodostetaan ominaisvektoreista kannanvaihtomatriisi

$$U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \overset{\text{lasku}}{\Longrightarrow} U^{-1} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & -2 & -1 \end{bmatrix}$$

Näille $U^{-1}AU = \operatorname{diag}(0,1,1)$, joten

$$\exp(A) = U \underbrace{\operatorname{diag}(1,e,e)}_{=\exp(\operatorname{diag}(0,1,1))} U^{-1} = \begin{bmatrix} 2e-1 & -e+1 & -e+1 \\ 0 & e & 0 \\ 2e-2 & -2e+2 & -e+2 \end{bmatrix}$$

On kaksi syytä miksi ominaisvektorit eivät aina muodosta kantaa

(1) Algebrallinen epätäydellisyys (tapaus $\mathbb{K}=\mathbb{R}$) polynomilla $\det(A-\lambda I)$ on liian vähän juuria

(2) Matriisilla on yleistettyjä ominaisvektoreita, eli vektoreita $u \in \mathbb{K}^n$ joille $(A-\lambda I)u \neq 0$, mutta $(A-\lambda I)^k u = 0$ jollekin $k>1$.

Tapaus (1) tapahtuu olennaisesti ainoastaan reaalisen matriisin

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

kanssa. Tälle $\det(A-\lambda I) = +\lambda^2+1 > 0 \quad \forall \lambda \in \mathbb{R}$, ja

$$\exp(A) = \begin{bmatrix} \cos 1 & -\sin 1 \\ \sin 1 & \cos 1 \end{bmatrix}$$

Syy tähän kummallisen oloiseen eksponenttiaaliin on se, että $A = \rho(i)$ kompleksiupotukselle $\rho : \mathbb{C} \to M_2(\mathbb{R})$.

Koska ρ on jatkuva rengashomomorfismi ja eksponentiaali määritellään sarjana,

$$\exp \circ \rho = \rho \circ \exp$$

$$\begin{tikzcd} \mathbb{C} \arrow[r, "\rho"] \arrow[d, "\exp"'] & M_2(\mathbb{R}) \arrow[d, "\exp"] \\ \mathbb{C}^* \arrow[r, "\rho"'] & GL(2,\mathbb{R}) \end{tikzcd}$$

$$\Rightarrow \exp(A) = \exp(\rho(i)) = \rho(\exp(i))$$
$$= \rho(e^i) = \rho(\cos 1 + i \sin 1)$$

Tämän takia myös reaalisessa tapauksessa pitää ottaa kompleksiset ominaisarvot ja -vektorit huomioon.

Yleistetyn ominaisvektorin tapauksessa matriisi ei ole diagonalisoituva vaan matriisin ns. Jordan muoto sisältää epätriviaaleja Jordan blokkeja

$$J(\lambda,r) = \begin{bmatrix} \lambda & 1 & 0 & & \\ 0 & \lambda & 1 & \ddots & \\ & 0 & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix} \in M_r(\mathbb{C})$$

Tällaiselle matriisille λ on ainoa ominaisarvo:

$$p(x) = \det\left(J(\lambda,r) - xI\right) = (\lambda - x)^r$$

mutta ainoastaan $e_1 \in \mathbb{C}^r$ on ominaisvektori.

Sen sijaan muille kantavektoreille

$$J(\lambda,r)e_k = e_{k-1} + \lambda e_k \quad \Rightarrow \quad \left(J(\lambda,r) - \lambda I\right)^k e_k = 0$$

$$\left(J(\lambda,r) - \lambda I\right)^{k-1} e_k \neq 0$$

eli e_k on k:n kertaluvun ominaisvektori.

$$e_k \xrightarrow{J-\lambda I} e_{k-1} \xrightarrow{J-\lambda I} e_{k-2} \to \dots \to e_1 \xrightarrow{J-\lambda I} 0$$

Jordan muodon konstruktiossa yleistetyistä ominaisvektoreista etsitään kanta jolla on ylläolevan kaltainen hierarkia.

Tätä konstruktiota voi pitää ortonormalisaation kaltaisena prosessina.

Esim 4.12

Olkoon $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. $\det(A - \lambda I) = (2-\lambda)^3$, joten ainoa ominaisarvo on $\lambda = 2$.

Standardikanta e_1, e_2, e_3 ei hajoa ketjuihin halutulla tavalla

$$e_2 \xrightarrow{A-2I} e_1 \xrightarrow{A-2I} 0$$

$$e_3 \xrightarrow{A-2I} e_1$$

Sen sijaan korvaamalla e_3 vektorilla $e_3 - e_2$ saadaan

$$e_2 \xrightarrow{A-2I} e_1 \xrightarrow{A-2I} 0$$

$$e_3 - e_2 \xrightarrow{A-2I} 0$$

ja $\{e_1, e_2, e_3 - e_2\}$ on edelleen $\mathbb{C}^3$:in kanta.

Asettamalla $U = [e_1 \quad e_2 \quad e_3 - e_2] \in GL(3,\mathbb{C})$, saadaan

$$A = U \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right] U^{-1} = U \begin{bmatrix} J(2,2) & 0 \\ 0 & J(2,1) \end{bmatrix} U^{-1}.$$