UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2016:21

Matrisexponentialfunktionen

Neda Farzaneh

Examensarbete i matematik, 15 hp
Handledare: Martin Herschend
Examinator: Jörgen Östensson
Juni 2016

Department of Mathematics
Uppsala University

Matrisexponentialfunktionen

Neda Farzaneh
Uppsala Universitet

Innehåll

1. Introduktion

Arbetet handlar om matrisexponentialfunktionen. Inom matematiken är matrisexponentialfunktionen en utökning av exponentialfunktionen för komplexa tal till kvadratiska matriser, så att man får en matrisfunktion. Exponentialfunktionen för matriser definieras som

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

Serien konvergerar för alla kvadratiska matriser, X.
I arbetet kommer vi att gå igenom några grundläggande begrepp, några egenskaper hos e^X, beräkning av e^X och matrislogaritmen.
Exponentialfunktionen som till varje komplext tal z tillordnar e^z har många viktiga egenskaper:

$$e^{Xx+y} = e^x e^y$$

$$\frac{\partial}{\partial t} e^{xt} \quad x, y \in \mathbb{C}.$$

I det här arbetet ska vi utvidga denna funktion så även är definierad för kvadratiska matriser X.

Vi går igenom grundläggande begrepp som; vektorrum, norm, normerat rum, Cauchyföljd och kompletta metriskt rum. Därefter bevisar vi några viktiga egenskaper för e^X. Till exempel: rummet av $n \times n$-matriser är komplett, konvergens och absolutkonvergens av e^X, derivatan av e^X och matrislogaritmen. Sedan handlar arbetet om matrislogaritmen. Vilken är en inversfunktion till matrisexponentialen. Det vill säga för varje matris X är e^X inverterbar. Lägg märke till att det bara är nollskilda tal som kan ha logaritm, ty e^z aldrig kan vara noll.

2. Grundläggande begrepp för vektorrum

Grundläggande egenskaper, definitioner och satser gällande gruppteori.

Definition 1 (Produktoperation). **[3]**
En **produktoperation**, även kallad **binär operation**, på en mängd $\mathcal{S}$ är en avbildning $*$ som tilldelar till varje ordnat par av element av $\mathcal{S}$ ett unikt bestämt element av $\mathcal{S}$. Det vill säga

$$* \,:\, \mathcal{S} \times \mathcal{S} \to \mathcal{S}$$

Observera att $\forall a, b \in \mathcal{S} \mid a * b \in \mathcal{S}$, det vill säga denna operation är sluten.

Definition 2 (Grupp). En **grupp** är en mängd $\mathcal{G}$ tillsammans med en binär opertion med följande egenskaper:

(1) Associativ: För alla $a, b, c \in \mathcal{G}$:

$$a * (b * c) = (a * b) * c$$

(2) Existens av identitetselement: Det existerar ett identitetselement, e, sådant att för alla $a \in \mathcal{G}$:

$$a * e = e * a = a$$

(3) Existens av invers: Det existerar för alla $a \in \mathcal{G}$ ett element, kallat för invers, a^{-1}, sådant att

$$a^{-1} * a = a * a^{-1} = e$$

Definition 3 (Abelsk grupp). **[3]**
En **abelsk grupp**, även kallat **kommutativ grupp**, är en grupp som även har den kommutativa egenskapen. Det vill säga:

$$\forall a, b \in \mathcal{G} \mid a * b = b * a$$

För abelska grupper betecknas ofta grupp operationen med +.

Lemma 1 (Identitetselementet är unikt). *Låt $\mathcal{G}$ vara en grupp och låt $e, f \in \mathcal{G}$ vara sådana att för alla $g \in \mathcal{G}$:*

$$e * g = g * e = g$$
$$f * g = g * f = g$$

Isåfall så är $e = f$.

Bevis. Eftersom, e är identitetselement, så har vi att

$$e * f = f$$

men då vi även har definierat f som identitetselement, så har vi att

$$e * f = e$$

Därav så har vi att

$$e = e * f = f$$

□

Lemma 2 (Varje element i en grupp har en och bara en invers). *Låt $\mathcal{G}$ vara en grupp, med e som identitetselement och g, h, k godtyckliga element i $\mathcal{G}$. Antag att*

$$g * h = h * g = e$$
$$g * k = k * g = e$$

Isåfall så är $h = k$.

Bevis. Vi vet att $g * h = g * k = e$. Utför vi produktoperationen på vänster sida med h så får vi $h * (g * h) = h * (g * k)$. Enligt den associativa egenskapen så är $(h * g) * h = (h * g) * k$, vilket ger oss att $e * h = e * k \quad och \quad h = k$. □

Lemma 3 (Produkten av ett element och dess invers är kommutativ). *Låt $\mathcal{G}$ vara en grupp, med e som identitetselement och g vara ett godtyckligt element i $\mathcal{G}$. Antag att $h \in \mathcal{G}$ uppfyller antingen $h * g = e$ eller $g * h = e$. Isåfall är h den unika inversen av g och både $h * g = e$ och $g * h = e$ är sanna.*

Bevis. För att visa att h är inversen till g så måste vi visa att både $h * g = e$ och $g * h = e$ är sanna. Antag att vi vet att $h * g = e$ är sant, då kan vi visa att detta implicerar $g * h = e$. Eftersom $h * g = e$ så har vi att $g * (h * g) = g * e = g$. Enligt den associativa egenskapen har vi därmed att $(g * h) * g = g$. Enligt definitionen av en grupp så vet vi att gruppen har en invers, och låt oss kalla denna invers för g^{-1}. Vi utför produktoperationen till höger med inversen g^{-1} och vi får

$$\begin{aligned} ((g * h) * g) * g^{-1} &= g * g^{-1} = e \\ (g * h) * \left(g * g^{-1}\right) &= e \\ (g * h) * e &= e \\ g * h &= e \end{aligned}$$

Motsvarande kan vi visa att ifall $g * h = e$ så innebär det att $h * g = e$. □

LEMMA 4 (Egenskaper för inverser). *Låt $\mathcal{G}$ vara en grupp där e är dess identitetselement, där g och h är godtyckliga element av $\mathcal{G}$, isåfall*

$$\left(g^{-1}\right)^{-1} = g \tag{1}$$

$$(g * h)^{-1} = h^{-1} * g^{-1} \tag{2}$$

$$e^{-1} = e \tag{3}$$

BEVIS. 1.

$$\begin{aligned} e &= g^{-1} * g \\ \implies \left(g^{-1}\right)^{-1} * e &= \left(g^{-1}\right)^{-1} * \left(g^{-1} * g\right) \\ \implies \left(g^{-1}\right)^{-1} * e &= \underbrace{\left(\left(g^{-1}\right)^{-1} * g^{-1}\right)}_{=e} * g \\ \implies \left(g^{-1}\right)^{-1} * e &= e * g \\ \implies \left(g^{-1}\right)^{-1} &= g \end{aligned}$$

2.

$$\begin{aligned} &\left(h^{-1} * g^{-1}\right) * (g * h) \\ &= h^{-1} * \left(\left(g^{-1} * g\right) * h\right) \\ &= h^{-1} * (e * h) \\ &h^{-1} * h = e. \end{aligned}$$

Så Lemma 3 ger $h^{-1} * g^{-1} = (g * h)^{-1}$.
3.

$$\begin{aligned} e * e &= e \\ \Rightarrow e &= e^{-1}. \end{aligned}$$

□

DEFINITION 4. (Vektorrum); [**1**, s. 33]
Ett komplext vektorrum är en abelsk grupp $(V, +)$ med en operation $\mathbb{C} \times V \to V, (\lambda, x) \mapsto \lambda x$ där följande påstående gäller;

$$\begin{aligned}
&1.\text{Assosiativitet} : \lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x,\\
&2.\text{Distributivitet} :\\
&(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x,\\
&\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y,\\
&3.\text{Enhetselement} : 1 \cdot x = x.
\end{aligned}$$

DEFINITION 5. (Norm, normerat rum); [**1**, s. 46]
En norm på ett vektorrum V tillordnar varje vektor x ett ickenegativt reellt tal som betecknas $||x||$.
För en norm ska följande uppfyllas;

$$\begin{aligned}
&\cdot \left\| x \right\| = 0 \Leftrightarrow x = 0,\\
&\cdot \left\| ax \right\| = |a| \left\| x \right\|,\\
&\cdot \left\| x + y \right\| \leqslant \left\| x \right\| + \left\| y \right\|.
\end{aligned}$$

EXEMPEL 1. Om $V = \mathbb{C}^n$ så är

$$\|x\| = \sqrt{|x_1^2| + |x_2^2| + |x_3^2| + ... + |x_n^2|}$$

en norm.

Om en norm är definierade på ett vektorrum då kallas det för *normerat rum.* [**1**, s. 46]

DEFINITION 6. (metriskt rum)[**1**, s. 1]
Ett *metriskt rum* är en mängd X med en funktion $d : X \times X \to \mathbb{R}$, där följande vilkor gäller;

$$\begin{aligned}
&d(x, y) = d(y, x),\\
&d(x, y) = 0 \Leftrightarrow x = y,\\
&d(x, y) \leqslant d(x, z) + d(z, y).
\end{aligned}$$

Från ovanstående ser vi att *avståndsfunktionen* betecknas som d.

EXEMPEL 2. Om V är ett normerat rum så är

$$d(x, y) = ||x - y||$$

en metrik på V.

DEFINITION 7. (Cauchyföjld) [**1**, s. 51]
Betrakta en följd

$$\{x_i\}_{i=1}^{\infty}$$

i ett metriskt rum (X, d).
Ifall $d(x_n, x_m)$ går mot noll då m, n går mot oändligheten (oberoende av varandra) kallas följden en *Cauchföljd.*
Mer precist gäller:

för varje ϵ finns det ett N att två godtyckliga element med index större än N har ett avstånd som är mindre än ϵ, det vill säga:

$$\forall \epsilon > 0 \quad \exists N_\epsilon \geqslant 1, \forall n, m > N_\epsilon, d(x_n, x_m) < \epsilon.$$

För ett normerat rum $(V, \|\cdot\|)$ gäller;

$$\forall \epsilon > 0 \quad \exists N_\epsilon \geqslant 1, \forall n, m > N_\epsilon, \|x_n - x_m\| < \epsilon.$$

Där $d(x_n, x_m)$ har ersatts med $\|x_n - x_m\|$.

Med *kompletta metriskt rum* menar man ett rum V där alla Cauchyföljder konvergerar.

3. Rummet av matriser

Vi kommer i huvudsak studera det normerade rummet $(V, \|\cdot\|)$ där

$$V = \{alla\ komplexa \quad n \times n - matriser\},$$

och

$$\|A\| = \sqrt{\sum |a_{ij}|^2},$$

där $A \in V$.

SATS 1. *Rummet $(V, \|\cdot\|)$ av komplexa $n \times n$-matriser är komplett. Dessutom gäller*

$$\|A \cdot B\| \leqslant \|A\| \cdot \|B\|.$$

BEVIS. Beviset för denna sats finns i [**2**, s. 28]. □

SATS 2. *För alla $A \in V$ gäller*

$$\|A^k\| \leqslant \|A\|^k.$$

BEVIS. Vi använder induktion för att bevisa denna sats.
Antag att

$$k = 1, \textit{(induktionbas)}.$$

Då har vi att:

$$k = 1, \quad \|A^1\| = \|A\| = \|A\|^1, \quad OK.$$

Antag att

$$\|A^k\| \leqslant \|A\|^k, \textit{(induktionantagandet)}.$$

Om det gäller för k då ska det gälla även för $k+1$:

$$\|A^{k+1}\| = \|A^k \cdot A\| \leqslant \|A^k\| \cdot \|A\| \leqslant \|A\|^k \cdot \|A\|.$$

Alltså

$$\|A^{k+1}\| \leqslant \|A\|^{k+1}, \textit{(induktionsbevis)}.$$

□

Definition 8. (konvergens)[**1**, s. 51]
Låt

$$S = \sum_{k=0}^{\infty} A_k$$

och

$$S_n = \sum_{k=0}^{n} A_k,$$

där A_k är $n \times n$-matriser, eller komplexa tal.
Vi säger att S konvergerar om följden (S_n) konvergerar och skriver

$$S = \lim_{n \to \infty} S_n.$$

Definition 9. (absolutkonvergens)[**2**, s. 32]
Vi säger att

$$S = \sum_{k=0}^{\infty} A_k$$

är absolutkonvergent om

$$\sum_{k=0}^{\infty} \|A_k\|$$

är konvergent.

Sats 3. *(absolutkonvergens av e^X)* [**2**, s. 32]
Betrakta serien

$$S = \sum_{k=0}^{\infty} A_k$$

och summan

$$S_n = \sum_{k=0}^{n} A_k,$$

där A_k är $n \times n$-matriser.
Om S är absolutkonvergent så är S konvergent.

Bevis. Om S absolutkonvergerar så är S_n en Cauchyföljd som konvergerar eftersom $(V, \|\cdot\|)$ är *komplett.*
Att S_n är en Cauchyföljd kan ses på följande vis: Antag att $m \geq n$. Så är

$$\begin{aligned} ||S_m - S_n|| = \left\| \sum_{k=n+1}^{m} A_k \right\| &\leq \sum_{k=n+1}^{m} ||A_k|| \\ &\leq \sum_{k=n+1}^{\infty} ||A_k|| \end{aligned}$$

som går mot noll om $n \to \infty$. □

4. Matrisexponentialfunktionen

Vi ska nu visa att

$$S = e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k$$

absolutkonvergerar.

Exempel 3. Vi uppskattar serien

$$\sum_{k=0}^{\infty} \left\| \frac{1}{k!} X^k \right\|.$$

Eftersom $\frac{1}{k!}$ bara är en konstant då kan vi ta ut det från absolut-tecknet:

$$\begin{aligned} \sum_{k=0}^{\infty} \left\| \frac{1}{k!} X^k \right\| &= \sum_{k=0}^{\infty} \frac{1}{k!} \left\| X^k \right\| \\ &\leqslant \sum_{k=0}^{\infty} \frac{1}{k!} \left\| X \right\|^k = e^{\|X\|} < \infty. \end{aligned}$$

Vi ser att serien är begränsad av konvergerar mot något tal nämligen $e^{\|X\|}$ vilket är mindre än oändligheten så serien absolutkonvergerar.

Sats 4. *(konvergens av e^X)* [**2**, s. 31]
Serien

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k$$

konvergerar.

Bevis. Om en serie är absolutkonvergent är den även konvergent, enligt Sats 3. Så från exemplet ovan ser vi att S är absolutkonvergens och alltså konvergerar. □

Enligt Sats 4 har vi alltså en funktion som till varje $n \times n$-matris X tillordnar matrisen e^X. Vi kallar denna funktion för *matrisexponentialfunktionen.*
Vi går nu igenom några viktiga egenskaper hos e^X.

Sats 5. [**2**, s. 29]
Låt X och Y vara två godtyckliga $n \times n -$ matriser.
Då har vi följande:

1.$e^0 = I.$

2.$(e^X)^* = e^{X^*}.$

3.e^X *är inverterbar och* $(e^X)^{-1} = e^{-X}.$

4.$e^{(\alpha+\beta)X} = e^{\alpha X} e^{\beta X}$ *för alla* α *och* β *i* $\mathbb{C}.$

5.*Om* C *är inverterbar, då* $e^{X+Y} = e^X e^Y = e^Y e^X.$

6.*Om* C *är inverterbar, då* $e^{CXC^{-1}} = Ce^X C^{-1}.$

7.$\left\| e^X \right\| \leq e^{\|X\|}.$

Bevis. 1. Antag $X = 0$.
Då är

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k = \underbrace{\frac{1}{0!} X^0}_{=I} + \underbrace{\sum_{k=1}^{\infty} \frac{1}{k!} X^k}_{=0} = I.$$

2.

$$e^{X^*} = \sum_{k=0}^{\infty} \frac{1}{k!} (X^*)^k = \frac{1}{0!} (X^*)^0 + \frac{1}{1!} (X^*)^1 + \cdots = \frac{1}{0!} (X^0)^* + \frac{1}{1!} (X^1)^* + \cdots$$
$$= \sum_{k=0}^{\infty} (X^k)^* = (e^X)^*.$$

3.
Eftersom

$$X \cdot (-X) = (-X) \cdot X = -X^2 \quad \textit{så} \quad e^X e^{-X} = e^{X+(-X)}, \quad \text{enligt 5.}$$

Men

$$e^{X+(-X)} = e^{X-X} = e^0 = I.$$

Alltså

$$(e^X)^{-1} = e^{-X}.$$

4.
Eftersom

$$(\alpha X)(\beta X) = \alpha\beta X^2 = (\beta X)(\alpha X), \quad \text{så} \quad e^{\alpha X + \beta X} = e^{\alpha X} e^{\beta X}, \quad \textit{enligt 5.}$$

5.

$$e^X e^Y = \left(\sum_{k=0}^{\infty} \frac{1}{k!} X^k \right) \left(\sum_{k=0}^{\infty} \frac{1}{k!} Y^k \right)$$
$$= \left(\frac{1}{0!} X^0 + \frac{1}{1!} X^1 + \frac{1}{2!} X^2 + \frac{1}{3!} X^3 + \dots \right) \left(\frac{1}{0!} Y^0 + \frac{1}{1!} Y^1 + \frac{1}{2!} Y^2 + \frac{1}{3!} Y^3 + \dots \right)$$

Om vi väljer en term från första parantesen: $\frac{1}{m!} X^m$,
och en term från andra parantesen: $\frac{1}{n!} Y^n$, så ger det
upphov till $\frac{1}{m!} \frac{1}{n!} X^m Y^n$. Alltså är

$$e^X e^Y = \sum_{m,n=0}^{\infty} \frac{1}{m!\, n!} X^m\, Y^n.$$

Eftersom produkten är kommutativ, dvs $XY = YX$, så är

$$e^{X+Y} = \sum_{k=0}^{\infty} \frac{1}{k!} (X+Y)^k = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} X^j Y^{k-j}$$

enligt binomialstasen.
Dessutom

$$\frac{1}{k!}\binom{k}{j} = \frac{1}{k!}\frac{k!}{j!(k-j)!} = \frac{1}{j!(k-j)!}.$$

Alltså är

$$e^{X+Y} = \sum_{k=0}^{\infty}\sum_{j=0}^{k}\frac{1}{j!(k-j)!}X^jY^{k-j},$$

Om vi nu ersätter $m = j$ och $n = k - j$.
Så är

$$\frac{1}{j!(k-j)!}X^jY^{k-j} = \frac{1}{m!n!}X^mY^n.$$

Alltså är

$$e^Ye^X = e^{X+Y}.$$

6.

$$e^{CXC^{-1}} = \sum_{k=0}^{\infty}\frac{1}{k!}(CXC^{-1})^k.$$

Eftersom

$$\begin{aligned}(CXC^{-1})^k &= CX\underbrace{C^{-1}C}_{=I}X\underbrace{C^{-1}C}_{=I}XC^{-1} = \\ &= CXIXI\cdots IXC^{-1} = CX^kC^{-1},\end{aligned}$$

så är

$$e^{CX^kC^{-1}} = \sum\frac{1}{k!}CXC^{-1}.$$

Eftersom C är konstant vi tar ut den från summan,

$$= C(\sum_{k=0}^{\infty}\frac{1}{k!}X^k)C^{-1} = Ce^XC^{-1}.$$

7.

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

□

Sats 6. *(Derivatan av e^X)* [**2**, s. 30]
Låt X vara en $n \times n$ komplex matris.
Då är $f(t) = e^{tX}$ en deriverbar funktion, och

$$\frac{\partial}{\partial t}e^{tX} = Xe^{tX} = e^{tX}X.$$

Speciellt,

$$\frac{\partial}{\partial t}e^{tX}\Big|_{t=0} = X.$$

Bevis.

$$\frac{\partial}{\partial t}(e^{tX}) = \frac{\partial}{\partial t}\sum_{k=0}^{\infty}\frac{1}{k!}(tX)^k = \frac{\partial}{\partial t}\sum_{k=0}^{\infty}\frac{t^k}{k!}X^k$$
$$= \sum_{k=0}^{\infty}\frac{\partial}{\partial t}\left(\frac{t^k}{k!}X^k\right),$$

vi vet att för $k > 0$ är

$$\frac{\partial}{\partial t}(t^k\frac{1}{k!}X^k) = kt^{k-1}, \frac{1}{k!}X^k = t^{k-1}\frac{k}{k!}X^k = t^{k-1}\frac{1}{(k-1)!}X^k.$$

Dessutom är $\frac{\partial}{\partial t}(t^0\frac{1}{0!}X^0) = 0$.
Alltså har vi att

$$\frac{\partial}{\partial t}(e^{tX}) = \sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}X^k = \sum_{l=0}^{\infty}\frac{t^l}{l!}X^{l+1},$$

där $k = l+1$ och $l = k-1$.
Dessutom är

$$\sum_{l=0}^{\infty}\frac{t^l}{l!}X^{l+1} = \sum_{l=0}^{\infty}\frac{t^l}{l!}X^l \cdot X^1$$
$$= X\left(\sum_{l=0}^{\infty}\frac{t^l}{l!}X^l\right) = \left(\sum_{l=0}^{\infty}\frac{t^l}{l!}X^l\right)X$$
$$= Xe^{tX} = e^{tX}X.$$

Därav har vi bevisat att

$$\frac{\partial}{\partial t}e^{tX} = Xe^{tX} = e^{tX}X.$$

□

En *tillämpning* av derivatan av matrisexponentialfunktioner är att lösa ekvationen $Y' = XY$.

Sats 7. *Ekvationen $Y' = XY$, där $Y = Y(t)$, har lösningar $Y = e^{tX}C$ där C är en konstant $n \times n$ matris.*

Bevis.

$$Y' = \frac{\partial}{\partial t}Y(t) = \frac{\partial}{\partial t}(e^{tX}C)$$
$$= \left(\frac{\partial}{\partial t}e^{tX}\right)C = X\underbrace{e^{tX}C}_{Y} = XY, \quad \text{enligt sats 6},$$

så

$$Y' = XY.$$

□

5. Beräkning av e^X

Här anger vi metoder för beräkning av e^X. Vi går igenom tre fall.
Fall 1:
X är *diagonaliserbar.*
Anta att X är en komplex $n \times n$ reell eller matris och att X är diagonaliserbar. Då existerar det en inverterbar komplex matris C så att $X = CDC^{-1}$. Där

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

e^D är diagonalmatrisen med egenvärden $e^{\lambda_1}, \cdots, e^{\lambda_n}$, och vi har

$$e^X = C \begin{bmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda n} \end{bmatrix} C^{-1},$$

eftersom

$$C^{-1} e^X C = e^{C^{-1} X C} = e^D.$$

Genom att diagonalisera X kan vi tydligt lösa ut e^X.

Exempel 4. Matrisen

$$X = \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}$$

är given.
Egenvektorer av X är $\begin{bmatrix} 1 \\ i \end{bmatrix}$ och $\begin{bmatrix} i \\ 1 \end{bmatrix}$ med egenvärden $-ia$ och ia.
Alltså den inverterbara matrisen

$$C = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$

och

$$D = \begin{bmatrix} -ia & 0 \\ 0 & ia \end{bmatrix}.$$

Vi beräknar $C^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{bmatrix}$, och

$$e^X = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{bmatrix} =$$

$$= \frac{1}{2} \begin{bmatrix} e^{-ia} & ie^{ia} \\ ie^{-ia} & e^{ia} \end{bmatrix} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix} =$$

$$= \frac{1}{2} \begin{bmatrix} e^{-ia} + e^{ia} & -ie^{-ia} + ie^{ia} \\ ie^{-ia} - ie^{ia} & e^{-ia} + e^{ia} = \end{bmatrix}.$$

Vi vet att

$$e^{-ia} + e^{ia} = \cos(-a) + i\sin(-a) + \cos(a) + i\sin(a)$$
$$= 2\cos a + 0 = 2\cos a,$$

$$i(e^{-ia} - e^{ia}) = i(\cos(-a) + i\sin(-a) - \cos(a) - i\sin(a))$$
$$= i(-2i\sin a) = 2\sin a.$$

Så

$$e^X = \frac{1}{2}\begin{bmatrix} e^{-ia} + e^{ia} & -ie^{-ia} + ie^{ia} \\ ie^{-ia} - ie^{ia} & e^{-ia} + e^{ia} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2\cos a & -2\sin a \\ 2\sin a & 2\cos a \end{bmatrix} = \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix}.$$

Notera att om X, och därav a, är reell då e^X är reell.

Fall 2:
X är nilpotent.
En $n \times n$ matris X är nilpotent om $X^m = 0$ för någon positivt tal m.
Vidare om $X^m = 0$, då $X^l = 0$ för alla $l > m$.
I detta fall efter de första termerna serien

$$\sum_{k=0}^{\infty} \frac{1}{k!} X^k$$

avslutas och kan beräknas exakt.

Exempel 5. Vi beräknar e^X där

$$X = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}.$$

Notera att

$$X^2 = \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

och att $X^3 = 0$. Alltså

$$e^X = \begin{bmatrix} 1 & a & b + \frac{1}{2}ac \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}.$$

Fall 3:
X är godtycklig.
En general matris X kan vara varken nilpotent eller diagonaliserbar.
Men vi har följande sats:

Sats 8. [**2**, s. 295]
Låt A vara en komplex $n \times n$ matris.
Då existerar det ett unikt par (S, N) av matriser med föjande egenskaper:
1. $A = S + N$,
2. $SN = NS$,
3. S är diagonaliserbar och
4. N är nilpotent.

Så varje matris X kan skrivas i formen $X = S + N$, enligt Sats 8, med S diagonaliserbar och N nilpotent och $SN = NS$.
Eftersom N och S kommuterar, har vi

$$e^X = e^{S+N} = e^S e^N.$$

Vi kan beräkna e^S och e^N som i fall 1 och fall 2.

Exempel 6. Matrisen

$$X = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$$

är given. Då är

$$X = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix},$$

där

$$S = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \quad \text{och} \quad N = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$$

uppfyller vilkoren i Sats 8.
Vi har alltså

$$e^X = e^S e^N = \begin{bmatrix} e^a & 0 \\ 0 & e^a \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e^a & e^a b \\ 0 & e^a \end{bmatrix}.$$

Exempel 7. Låt $a = 2$ och $b = 3$.
det vill säga vi har matrisen

$$X = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}.$$

Lägg märke till att

$$X = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}}_{=S} + \underbrace{\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}}_{=N},$$

där (S, N) är som i sats 8.

$$e^X = e^S e^N = \begin{bmatrix} e^2 & 0 \\ 0 & e^2 \end{bmatrix} e^N : \quad N = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} \quad N^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

$$e^N = \frac{1}{0!} I + \frac{1}{1!} N + \underbrace{\frac{1}{2!} N^2 + \cdots}_{=0}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

$$e^X = e^S e^N = \begin{bmatrix} e^2 & 0 \\ 0 & e^2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e^2 & 3e^2 \\ 0 & e^2 \end{bmatrix}.$$

6. Matrislogaritm

Matrislogaritm är en invers funktion till matrisexponentialen.
Vi tittar på logaritmen av komplexa talen för att se vad är det som är rimligt att förvänta sig i matrisfallet.
Eftersom e^z aldrig kan vara noll, kan bara nollskilda tal ha en logaritm.
Varje nollskild komplex tal kan skrivas som e^z för någon z, men z är inte unik. Det finns inte någon kontinuerligt väg till att definiera logaritmen i mängden av alla nollskilda komplexa tal.
Situationen är liknande för matriser.
För varje matris X är e^X inverterbar. Så det är bara inverterbara matriser som kan ha logaritm.

SATS 9. [**2**, s. 32]
Funktionen

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}$$

är definierad och analytik i en cirkel med radien 1 *och centrum* $z = 1$.
För alla z *med* $|z-1| < 1$, *gäller*

$$e^{\log z} = z.$$

BEVIS. Denna sats kan bevisas med analytisk utvidgning.
se [**2**, s. 32], alltså Lemma 2.5. □

DEFINITION 10. [**2**, s. 33]
För en $n \times n$ matris A definerar vi $\log A$ med

$$\log A = \sum_{m=1}^{\infty} (-1)^{(m+1)} \frac{(A-I)^m}{m}$$

närhelst serien konvergerar.

Eftersom serien

$$\sum_{m=1}^{\infty} (-1)^{(m+1)} \frac{(z-1)^m}{m}$$

har konvergensradie 1 och eftersom

$$||(A-I)^m|| \leq ||A-I||^m,$$

kommer serien

$$\sum_{m=1}^{\infty} (-1)^{(m+1)} \frac{(A-I)^m}{m}$$

att konvergera om $||A-I|| < 1$.

SATS 10. *Låt* A *vara en godtycklig komplex* $n \times n$ *matris. Då finns diagonaliserbara* $n \times n$ *matriser* A_m $m \geq 1$ *så att*

$$\lim_{m \to \infty} A_m = A.$$

BEVIS. Se [**2**, s. 59] avsnitt, alltså exercise 5. □

Sats 11. [**2**, s. 34]
Funktionen

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m}$$

är definierad och kontinuerlig på uppsättningen av alla komplexa $n \times n$ matriser A med $||A - I|| < 1$.
För alla A med $||A - I|| < 1$, gäller

$$e^{\log A} = A.$$

Bevis. Eftersom

$$||(A-I)^m|| \leq ||(A-I)||^m$$

och eftersom serien

$$\sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}$$

har konvergensradie 1, kommer serien

$$\sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m}$$

att absolutkonvergera för alla A med $||A - I|| < 1$.
Vi visar $\exp(\log A) = A$ för alla A med $||A - I|| < 1$ genom två fall.
Fall 1
A är diagonaliserbar.
Anta att

$$A = CDC^{-1},$$

med D diagonal.
Då

$$A - I = CDC^{-1} - I = C(D-I)C^{-1}.$$

Det följer att $(A - I)^m$ är på formen

$$(A-I)^m = C \begin{bmatrix} (z_1 - 1)^m & & 0 \\ & \ddots & \\ 0 & & (z_n - 1)^m \end{bmatrix} C^{-1},$$

där $z_1, \cdots, z_n$ är egenvärden av A.

Om $||A - I|| < 1$, då måste varje egenvärde z_k av A uppfylla $|z_k - 1| < 1$.
Så

$$\sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} = C \begin{bmatrix} \log z_1 & & 0 \\ & \ddots & \\ 0 & & \log z_n \end{bmatrix} C^{-1},$$

och

$$e^{\log A} = C \begin{bmatrix} e^{\log z_1} & & 0 \\ & \ddots & \\ 0 & & e^{\log z_n} \end{bmatrix} C^{-1} = A.$$

Fall 2
A är inte diagonaliserbar.
Välj A_m $m \in \mathbb{N}$ enligt Sats 10 så att

$$\lim_{m\to\infty} A_m = A$$

och A_m diagonaliserbar.

$$\lim_{m\to\infty} ||A_m - I|| = ||A - I|| < 1.$$

Alltså finns det $N > 0$ så att för $m \geq N$ gäller $||A_m - I|| < 1$,
så för $m \geq N$ gäller

$$\begin{aligned} e^{\log(A_m)} &= A_m \\ \text{och} \quad \lim_{m\to\infty} e^{\log(A_m)} &= \lim A_m = A \\ \text{men} \quad e^{\log(\lim_{m\to\infty} A_m)} &= e^{\log A}. \quad \text{Så} \\ e^{\log A} &= A. \end{aligned}$$

□

Litteraturförteckning

[1] *B. Choudhary, Sudarsan Nanda, Functional Analysis with Applications, Wiley Easten Ltd, New Dehli, 1989*

[2] *Brian C. Hall, Lie Groups, Lie algebras and Representations, Graduate Texts in Mathematics, Springer, 2003*

[3] *Jones, H.F Groups, representations and physics, Institute of Physics, 1998*