Exponential of a matrix

Arundhati Roy

August 16, 2020

Outline

- Convergence of the series for arbitrary nxn real matrix

  A Cauchy sequence in a complete metric space is convergent.

Outline

- Convergence of the series for arbitrary nxn real matrix

  A Cauchy sequence in a complete metric space is convergent.

- Two properties of the matrix exponential in $M_n(\mathbb{R})$

A norm on the space $M_n(\mathbb{R})$

We associate a real number

$$||A|| := \sup_{v \in S^{n-1}} ||Av|| = \sup_{v \in \mathbb{R}^n - \mathbf{0}} ||Av||/||v||$$

to every $A \in M_n(\mathbb{R})$.

We shall show that the function thus defined from $M_n(\mathbb{R})$ to $\mathbb{R}$ is a *norm* on the space $M_n(\mathbb{R})$, i.e. it has the following three properties.

- $||A|| \geq 0$ with equality iff $A = \mathbf{0}_n$

  Proof: $||Av|| \geq 0 \ \forall v \in \mathbb{R}^n$, so $||A|| \geq 0$

  $A = O_n \Rightarrow Av = O_{\mathbb{R}^n}$ or $||Av|| = 0 \ \forall v \in \mathbb{R}^n \Rightarrow ||A|| = 0$

  And,

  $||A|| = 0 \Rightarrow ||Av||/||v|| = 0$ or $||Av|| = 0 \Rightarrow Av = 0 \Rightarrow A = O_n$

  $\therefore ||A|| = 0 \Leftrightarrow A = O_n$

A norm on the space $M_n(\mathbb{R})$(Contd.)

- $||cA|| = |c|.||A||$

  Proof: $(cA)v = c(Av)\ \forall v \in \mathbb{R}^n,\ c \in \mathbb{R}$, so

  $$||(cA)v|| = ||c(Av)|| = |c|.||Av||\ \forall v \in \mathbb{R}^n$$

  $$||Av|| \leq ||A||\ \forall v \in S^{n-1}\text{ , so}$$

  $$|c|.||Av|| \leq |c|.||A|| \text{ or } ||(cA)v|| \leq |c|.||A||\ \forall v \in S^{n-1}$$

  $$\Rightarrow \sup_{v \in S^{n-1}} ||(cA)v|| \leq |c|.||A|| \text{ or } ||cA|| \leq |c|.||A||$$

  Also, $||cA|| \geq ||(cA)v||$ or $||cA|| \geq |c|.||Av||\ \forall v \in S^{n-1}$

  $$\sup_{v \in S^{n-1}} |c|.||Av|| = |c| \sup_{v \in S^{n-1}} ||Av|| = |c|.||A||$$

  $$\Rightarrow ||cA|| \geq |c|.||A||$$

  $\therefore ||cA|| = |c|.||A||$

A norm on the space $M_n(\mathbb{R})$(Contd.)

- $||A+B|| \leq ||A|| + ||B||$

  Proof: $\forall v \in S^{n-1}$, $||(A+B)v|| = ||Av + Bv|| \leq ||Av|| + ||Bv||$

  By definition, $||A|| \geq ||Av||$ and $||B|| \geq ||Bv||$ $\forall v \in S^{n-1}$

  $\Rightarrow ||Av|| + ||Bv|| \leq ||A|| + ||B||$ $\forall v \in S^{n-1}$

  $\Rightarrow ||(A+B)v|| \leq ||A|| + ||B||$ $\forall v \in S^{n-1}$

$$\Rightarrow \sup_{v \in S^{n-1}} ||(A+B)v|| \leq ||A|| + ||B||$$

$\therefore ||A+B|| \leq ||A|| + ||B||$

A norm on the space $M_n(\mathbb{R})$(Contd.)

We shall see that this metric has another additional property, i.e.

$$||AB|| \leq ||A||.||B||$$

Proof: $v \longmapsto Bv$ is a linear map on $\mathbb{R}^n$, so $Bv = O_n \ \forall v \in \mathbb{R}^n - \mathbf{0}$

$$\Leftrightarrow B = O_n \Leftrightarrow ||B|| = 0 \text{ and } ||AB|| = 0 \Rightarrow ||AB|| = ||A||.||B|| = 0$$

Otherwise, $\exists u \in \mathbb{R}^n - \mathbf{0}$ such that $Bv \in \mathbb{R}^n - \mathbf{0}$, so that $||Bv|| \neq 0$

For any such u, $\frac{||ABu||}{||Bu||} \leq \sup_{v \in S^{n-1}} \frac{||Av||}{||v||}$, or $\frac{||ABu||}{||Bu||} \leq ||A||$

$$\Rightarrow ||ABu|| \leq ||A||.||Bu|| \text{ or } \frac{||ABu||}{||u||} \leq ||A||.\frac{||Bu||}{||u||}$$

$$\Rightarrow \sup_{v \in \mathbb{R}^n - \mathbf{0}} \frac{||ABv||}{||v||} \leq \sup_{v \in \mathbb{R}^n - \mathbf{0}} ||A||.\frac{||Bv||}{||v||}$$

$$\Rightarrow ||AB|| \leq ||A||. \sup_{v \in \mathbb{R}^n - \mathbf{0}} \frac{||Bv||}{||v||}$$

$$\therefore ||AB|| \leq ||A||.||B||$$

A suitable metric on $M_n(\mathbb{R})$

Next, we claim that $d(A, B) = ||A - B||$ is a metric on $M_n(\mathbb{R})$ such that $(M_n(\mathbb{R}), d)$ is a complete metric space.

Proof: We take an arbitrary Cauchy sequence $\{A_n\}$ in $M_n(\mathbb{R})$,

so that for $\epsilon > 0$, $\exists N_0 \in \mathbb{N}$ such that $d(A_k, A_m) < \epsilon \ \forall m, k \geq N_0$

i.e. $||A_k - A_m|| < \epsilon \ \forall m, k \geq N_0$

$$||A_k - A_m|| = \sup_{v \in S^{n-1}} ||(A_k - A_m)v|| = \sup_{v \in S^{n-1}} ||A_k v - A_m v||$$

$$\geq ||A_k v - A_m v|| \ \forall v \in S^{n-1} \text{ for any m,k} \in \mathbb{N}$$

If $B = \{e_1, e_2, ..., e_n\}$ is the standard ordered basis of the vector space $\mathbb{R}^n$, $e_i \in S^{n-1}$ $\forall i \in \{1, 2, 3, ..., n\}$ and $(A_k e_i)_j = (A_k)_{ji}$ $\forall j \in \{1, 2, 3, ..., n\}$

$||A_k e_i|| = (\sum_{j=1}^{n} (A_k)_{ji}^2)^{\frac{1}{2}} \geq |(A_k)_{ji}|$ for any $j \in \{1, 2, ..., n\}$

$\Rightarrow ||(A_k - A_m)e_i|| \geq |(A_k - A_m)_{ji}|$

i.e. $||(A_k - A_m)e_i|| \geq |(A_k)_{ji} - (A_m)_{ji}|$ for any i,j $\in \{1, 2, ..., n\}$, m,k $\in \mathbb{N}$

$\Rightarrow ||A_k - A_m|| \geq |(A_k)_{ji} - (A_m)_{ji}|$ $\forall i, j \in \{1, 2, ..., n\}$, m,k $\in N$

$\Rightarrow |(A_k)_{ji} - (A_m)_{ji}| < \epsilon$ $\forall m, k \geq N_0$ for any i,j $\in \{1, 2, ..., n\}$

So, $\{(A_k)_{ji}\}_{k \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$.
$\Rightarrow$ Since $\mathbb{R}$ is a complete metric space, $\exists a \in \mathbb{R}$ such that $(A_k)_{ji} \to a$ as $k \to \infty$ for each i,j $\in \{1, 2, ..., n\}$

We define

$$(A)_{ji} := \lim_{k \to \infty} (A_k)_{ji}$$

A suitable metric on $M_n(\mathbb{R})$(Contd.)

$\therefore \exists N_1 \in \mathbb{N}$ such that $\forall k \geq N_1$, $|(A_k)_{ij} - (A)_{ij}| < \frac{\epsilon}{2n}$

$\forall v \in S^{n-1}$, $||(A_k - A)v|| = [\sum_{i=1}^{n}(\sum_{j=1}^{n}((A_k)_{ij} - (A)_{ij})v_j)^2]^{\frac{1}{2}}$,

where $(v_1, v_2, ..., v_n)$ is the coordinate vector of v with respect to the basis B, $v_j \in \mathbb{R}\ \forall j \in \{1, 2, ..., n\}$ and $||v|| = (\sum_{j=1}^{n} v_j^2)^{\frac{1}{2}} = 1$

Now, $\forall k \geq N_1$, $(A_k)_{ij} - (A)_{ij} < \frac{\epsilon}{2n}$,

so $\sum_{j=1}^{n}((A_k)_{ij} - (A)_{ij})v_j < \sum_{j=1}^{n} \frac{\epsilon}{2n} v_j$

$\Rightarrow (\sum_{j=1}^{n}((A_k)_{ij} - (A)_{ij})v_j)^2 < \frac{\epsilon^2}{(2n)^2}(\sum_{j=1}^{n} v_j)^2$

$\Rightarrow ||(A_k - A)v|| < (\sum_{i=1}^{n} \frac{\epsilon^2}{(2n)^2}(\sum_{j=1}^{n} v_j)^2)^{\frac{1}{2}}$

$\Rightarrow ||(A_k - A)v|| < \frac{\epsilon}{2n}.\sqrt{n}.|\sum_{j=1}^{n} v_j|$

A suitable metric on $M_n(\mathbb{R})$(Contd.)

$|\frac{1}{n}\sum_{j=1}^{n} v_j| \leq \frac{1}{n}|\sum_{j=1}^{n} v_j| \leq (\frac{1}{n}\sum_{j=1}^{n} v_j^2)^{\frac{1}{2}}$

$\Rightarrow \frac{1}{n}|\sum_{j=1}^{n} v_j| \leq \frac{1}{\sqrt{n}}(\sum_{j=1}^{n} v_j^2)^{\frac{1}{2}}$

$\Rightarrow |\sum_{j=1}^{n} v_j| \leq \sqrt{n}$

$\Rightarrow \frac{\epsilon}{2n}.\sqrt{n}.|\sum_{j=1}^{n} v_j| \leq \frac{\epsilon}{2}$

$\Rightarrow ||(A_k - A)v|| < \frac{\epsilon}{2}$

$\Rightarrow \sup_{v\in S^{n-1}} ||(A_k - A)v|| \leq \frac{\epsilon}{2} < \epsilon$

$\therefore ||(A_k - A)|| < \epsilon$ or $d(A_k, A) < \epsilon \ \forall k \geq N_1$

i.e. $\{A_k\}$ converges to A $\in M_n(\mathbb{R})$.

Hence, $(M_n(\mathbb{R}), d)$ is a complete metric space.

Convergence of the exponential series on $M_n(\mathbb{R})$

We define

$$S_n := \sum_{i=0}^{n-1} \frac{A^i}{i!}, n \in \mathbb{N}$$

Without loss of generality, let $m < n, m, n \in \mathbb{N}$
Let $\epsilon > 0$. $d(S_n, S_m) = ||S_n - S_m|| = ||\sum_{i=m}^{n-1} \frac{A^i}{i!}||$

$$\leq \sum_{i=m}^{n-1} ||\frac{A^i}{i!}||$$

$$\leq \sum_{i=m}^{n-1} \frac{||A||^i}{i!}$$

$\{\frac{||A||^n}{n!}\}$ is a sequence in $\mathbb{R}$.

We define

$$D_n := \sum_{i=0}^{n-1} \frac{||A||^i}{i!}, \forall n \in \mathbb{N}$$

We can show that the series $\sum_{i=0}^{\infty} \frac{||A||^i}{i!}$ converges in $\mathbb{R}$ by ratio test, as

Convergence of the exponential series on $M_n(\mathbb{R})$ (Contd.)

$\limsup_{n\to\infty} \left|\frac{\frac{||A||^{n+1}}{(n+1)!}}{\frac{||A||^n}{n!}}\right| = \lim_{n\to\infty} \frac{||A||}{n+1} = 0$

Hence the sequence of partial sums, $\{D_n\}$ is a Cauchy sequence in $\mathbb{R}$.

$\Rightarrow \exists N_2 \in \mathbb{N}$ such that if $N_2 \leq m < n$, $|D_n - D_m| < \epsilon$

or $\sum_{i=m}^{n-1} \frac{||A||^i}{i!} < \epsilon$

$\Rightarrow \forall m, n \in \mathbb{N}$ such that $n > m \geq N_2$, $d(S_n - S_m) < \epsilon$

$\Rightarrow \{S_n\}$ is a Cauchy sequence in $(M_n(\mathbb{R}), d)$.

$\Rightarrow \{S_n\}$ converges in $(M_n(\mathbb{R}), d)$.

Hence, the series $e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!}$ is convergent in $(M_n(\mathbb{R}), d)$

Property I

- $e^{PAP^{-1}} = Pe^A P^{-1}$

We define

$$S_n(B) = \sum_{i=0}^{n-1} \frac{B^i}{i!} \forall n \in \mathbb{N}, \forall B \in M_n(\mathbb{R})$$

$\Rightarrow S_n(A) = \sum_{i=0}^{n-1} \frac{A^i}{i!}$, and $S_n(PAP^{-1}) = \sum_{i=0}^{n-1} \frac{(PAP^{-1})^i}{i!}$

$PS_n(A)P^{-1} = P(\sum_{i=0}^{n-1} \frac{A^i}{i!})P^{-1} = \sum_{i=0}^{n-1} \frac{PA^iP^{-1}}{i!}$

$\Rightarrow PS_n(A)P^{-1} - S_n(PAP^{-1}) = \sum_{i=0}^{n-1} \frac{(PA^iP^{-1})-(PAP^{-1})^i}{i!}$

Now, $(PA^jP^{-1})(PAP^{-1}) = (PA^j)(P^{-1}P)(AP^{-1})$
$= (PA^j)(AP^{-1})$
$= PA^{j+1}P^{-1} \ \forall j \in \mathbb{N}$

$\Rightarrow (PA^jP^{-1}) = (PAP^{-1})^j \ \forall j \in \mathbb{N}$

Property I (Contd.)

$\Rightarrow PS_n(A)P^{-1} - S_n(PAP^{-1}) = 0$ or $PS_n(A)P^{-1} = S_n(PAP^{-1})\ \forall n \in \mathbb{N}$

$\lim_{n\to\infty} S_n(A) = e^A$, $\lim_{n\to\infty} S_n(PAP^{-1}) = e^{PAP^{-1}}$

Let $\epsilon > 0$. $\exists K \in \mathbb{N}$ such that $d(S_n(A), e^A) < \frac{\epsilon}{||P||.||P^{-1}||}$,
i.e. $||S_n(A) - e^A|| < \frac{\epsilon}{||P||.||P^{-1}||}\ \forall n \geq K$

Now,
$$\begin{aligned} ||PS_n(A)P^{-1} - Pe^AP^{-1}|| &= ||P(S_n(A) - e^A)P^{-1}|| \\ &\leq ||P||.||S_n(A) - e^A||.||P^{-1}|| \end{aligned}$$

$\Rightarrow d(PS_n(A)P^{-1}, Pe^AP^{-1}) = ||PS_n(A)P^{-1} - Pe^AP^{-1}|| < \epsilon\ \forall n \geq K$

$\Rightarrow \lim_{n\to\infty} PS_n(A)P^{-1} = Pe^AP^{-1}$ or $\lim_{n\to\infty} S_n(PAP^{-1}) = Pe^AP^{-1}$

$\therefore e^{PAP^{-1}} = Pe^AP^{-1}$

Property II

- $e^A e^{-A} = I$

If two matrices A and B in $M_n(\mathbb{R})$ commute,

$$(A+B)^k = \sum_{a+b=k, a,b\geq 0} \binom{k}{a} A^a B^b \ [\because AB = BA]$$

$$\Rightarrow S_n(A+B) = \sum_{k=0}^{n-1} \frac{(A+B)^k}{k!} = \sum_{k=0}^{n-1} \frac{1}{k!} \sum_{a+b=k, a,b\geq 0} \binom{k}{a} A^a B^b$$

$$= \sum_{k=0}^{n-1} \sum_{a+b=k, a,b\geq 0} \frac{A^a}{a!} \cdot \frac{B^b}{b!}$$

$$\Rightarrow e^{A+B} = \sum_{k=0}^{\infty} \sum_{a+b=k, a,b\geq 0} \frac{A^a}{a!} \cdot \frac{B^b}{b!} = \sum_{a,b\geq 0} \frac{A^a}{a!} \cdot \frac{B^b}{b!}$$

Also, $S_n(A)S_n(B) = \sum_{0\leq a,b\leq n-1} \frac{A^a}{a!} \cdot \frac{B^b}{b!}$,

so $\lim_{n\to\infty} S_n(A)S_n(B) = \sum_{a,b\geq 0} \frac{A^a}{a!} \cdot \frac{B^b}{b!} = e^{A+B}$

Property II(Contd.)

$$\begin{aligned} d(S_n(A)S_n(B), e^Ae^B) &= ||S_n(A)S_n(B) - e^Ae^B|| \\ &\le ||S_n(A)S_n(B) - e^AS_n(B)|| + ||e^AS_n(B) - e^Ae^B|| \\ &\le ||S_n(A) - e^A||.||S_n(B)|| + ||e^A||.||S_n(B) - e^B|| \end{aligned}$$

Let $\varepsilon > 0$. $\lim_{n\to\infty} S_n(B) = e^B$, $\lim_{n\to\infty} S_n(A) = e^A$.

$\Rightarrow \exists N_3 \in \mathbb{N}$ such that $||S_n(B) - e^B|| < \frac{\varepsilon}{2||e^A||}$, $\forall n \ge N_3$

$\{S_n(B)\}$ converges in $M_n(\mathbb{R})$, so it is bounded, i.e. $\exists \mathbb{K} \in \mathbb{R}$ such that

$d(S_n(B), S_1(B)) = ||S_n(B) - I|| < \mathbb{K}\ \forall n \in \mathbb{N}$

$\Rightarrow ||S_n(B)|| < \mathbb{K} + ||I||$ or $||S_n(B)|| < \mathbb{K} + 1\ \forall n \in \mathbb{N}$
$[||I|| = \sup_{v\in S^{n-1}} ||v|| = 1]$

$\lim_{n\to\infty} S_n(A) = e^A$, so $\exists N_4 \in \mathbb{N}$ such that $\forall n \ge N_4$,

$||S_n(A) - e^A|| < \frac{\varepsilon}{2(\mathbb{K}+1)}$. So $||S_n(A)S_n(B) - e^Ae^B|| < \varepsilon\ \forall n \ge \max\{N_3, N_4\}$

Property II(Contd.)

$\Rightarrow \lim_{n\to\infty} S_n(A)S_n(B) = e^A e^B$

$\Rightarrow e^{A+B} = e^A e^B$

A and (-A) commute $\forall A \in M_n(\mathbb{R})$.

So, $e^A e^{-A} = e^{A+(-A)} = e^0 = I$

$\therefore e^A e^{-A} = I$