On $e^A e^B = e^B e^A \Rightarrow AB = BA$

Roger Horn

MIMS Meeting, Manchester, UK, May 16, 2008

Primary Matrix Functions

- $A = SJS^{-1}, J = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_k}(\lambda_k)$

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{bmatrix} \in M_m$$

Primary Matrix Functions

- $A = SJS^{-1}, J = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_k}(\lambda_k)$

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{bmatrix} \in M_m$$

- $f(A) \equiv Sf(J)S^{-1}, f(J) \equiv f\left(J_{n_1}(\lambda_1)\right) \oplus \cdots \oplus f\left(J_{n_k}(\lambda_k)\right)$

Primary Matrix Functions

- $A = SJS^{-1}, J = J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_k}(\lambda_k)$

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{bmatrix} \in M_m$$

- $f(A) \equiv Sf(J)S^{-1}, f(J) \equiv f(J_{n_1}(\lambda_1)) \oplus \cdots \oplus f(J_{n_k}(\lambda_k))$
-

$$f(J_m(\lambda)) \equiv \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!}f^{(2)}(\lambda) & \cdots & \frac{1}{(m-1)!}f^{(m-1)}(\lambda) \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!}f^{(2)}(\lambda) \\ & & & \ddots & f'(\lambda) \\ \mathbf{0} & & & & f(\lambda) \end{bmatrix}$$

Important Facts

- $f(A)$ is well defined (independent of the choice of S)

Important Facts

- $f(A)$ is well defined (independent of the choice of S)
- $JCF(f(J_m(\lambda)) = J_m(f(\lambda))$ if $f'(\lambda) \neq 0$ (blocks do not split)

Important Facts

- $f(A)$ is well defined (independent of the choice of S)
- $JCF(f(J_m(\lambda)) = J_m(f(\lambda))$ if $f'(\lambda) \neq 0$ (blocks do not split)
- $f(A)$ is some polynomial in A. If B commutes with A, then B commutes with $f(A)$.

Existence of a Unique Solution of f(X) = Y (Figure 1)

- $$\begin{aligned} H_1 &: \; f'(z) \neq 0 \text{ on } \mathcal{D}_1 \\ H_2 &: \; f(z_1) = f(z_2) \;\&\; z_1, z_2 \in \mathcal{D}_1 \Leftrightarrow z_1 = z_2 \end{aligned}$$

Existence of a Unique Solution of f(X) = Y (Figure 1)

- $$\begin{aligned} H_1 &: \quad f'(z) \neq 0 \text{ on } \mathcal{D}_1 \\ H_2 &: \quad f(z_1) = f(z_2) \;\&\; z_1, z_2 \in \mathcal{D}_1 \Leftrightarrow z_1 = z_2 \end{aligned}$$

- Suppose $\sigma(A) \subset \mathcal{D}_1$. Then $J(A)$ and $J(f(A))$ have the *same* sets of block sizes (H_1) with respective eigenvalues λ_j and $f(\lambda_j)$, $j = 1, \ldots, k$ (H_2).

Existence of a Unique Solution of f(X) = Y (Figure 1)

- $$\begin{aligned} H_1 &: \quad f'(z) \neq 0 \text{ on } \mathcal{D}_1 \\ H_2 &: \quad f(z_1) = f(z_2) \;\&\; z_1, z_2 \in \mathcal{D}_1 \Leftrightarrow z_1 = z_2 \end{aligned}$$

- Suppose $\sigma(A) \subset \mathcal{D}_1$. Then $J(A)$ and $J(f(A))$ have the *same* sets of block sizes (H_1) with respective eigenvalues λ_j and $f(\lambda_j)$, $j = 1, \ldots, k$ (H_2).
- $g(f(A)) = A$ is a primary matrix function, so A is a polynomial in $f(A)$.

Existence of a Unique Solution of f(X) = Y (Figure 1)

- $$\begin{aligned} H_1 &: f'(z) \neq 0 \text{ on } \mathcal{D}_1 \\ H_2 &: f(z_1) = f(z_2) \ \& \ z_1, z_2 \in \mathcal{D}_1 \Leftrightarrow z_1 = z_2 \end{aligned}$$
- Suppose $\sigma(A) \subset \mathcal{D}_1$. Then $J(A)$ and $J(f(A))$ have the *same* sets of block sizes (H_1) with respective eigenvalues λ_j and $f(\lambda_j)$, $j = 1, \ldots, k$ (H_2).
- $g(f(A)) = A$ is a primary matrix function, so A is a polynomial in $f(A)$.
- Suppose $\sigma(B) \subset \mathcal{D}_1$ and $f(B) = f(A)$. Then block sizes and eigenvalues of B are same as those of A, so $JCF(B) = JCF(A)$ and hence B is similar to A: $B = SAS^{-1}$

Existence of a Unique Solution of f(X) = Y (Figure 1)

- $$\begin{aligned} H_1 &: \quad f'(z) \neq 0 \text{ on } \mathcal{D}_1 \\ H_2 &: \quad f(z_1) = f(z_2) \;\&\; z_1, z_2 \in \mathcal{D}_1 \Leftrightarrow z_1 = z_2 \end{aligned}$$
- Suppose $\sigma(A) \subset \mathcal{D}_1$. Then $J(A)$ and $J(f(A))$ have the *same* sets of block sizes (H_1) with respective eigenvalues λ_j and $f(\lambda_j)$, $j = 1, \ldots, k$ (H_2).
- $g(f(A)) = A$ is a primary matrix function, so A is a polynomial in $f(A)$.
- Suppose $\sigma(B) \subset \mathcal{D}_1$ and $f(B) = f(A)$. Then block sizes and eigenvalues of B are same as those of A, so $JCF(B) = JCF(A)$ and hence B is similar to A: $B = SAS^{-1}$
- $f(A) = f(B) = f(SAS^{-1}) = Sf(A)S^{-1}$, so $Sf(A) = f(A)S$, and hence $SA = AS$ since A is a polynomial in $f(A)$.

Existence of a Unique Solution of f(X) = Y (Figure 1)

- $$\begin{aligned} H_1 &: \quad f'(z) \neq 0 \text{ on } \mathcal{D}_1 \\ H_2 &: \quad f(z_1) = f(z_2) \;\&\; z_1, z_2 \in \mathcal{D}_1 \Leftrightarrow z_1 = z_2 \end{aligned}$$
- Suppose $\sigma(A) \subset \mathcal{D}_1$. Then $J(A)$ and $J(f(A))$ have the *same* sets of block sizes (H_1) with respective eigenvalues λ_j and $f(\lambda_j)$, $j = 1, \ldots, k$ (H_2).
- $g(f(A)) = A$ is a primary matrix function, so A is a polynomial in $f(A)$.
- Suppose $\sigma(B) \subset \mathcal{D}_1$ and $f(B) = f(A)$. Then block sizes and eigenvalues of B are same as those of A, so $JCF(B) = JCF(A)$ and hence B is similar to A: $B = SAS^{-1}$
- $f(A) = f(B) = f(SAS^{-1}) = Sf(A)S^{-1}$, so $Sf(A) = f(A)S$, and hence $SA = AS$ since A is a polynomial in $f(A)$.
- Finally, $B = SAS^{-1} = ASS^{-1} = A$

Application to f(z) = exp(z) (Figure 2)

- Suppose $\sigma(A), \sigma(B) \subset \mathcal{D}_1$. Then A is a polynomial in e^A and B is a polynomial in e^B, so e^A commutes with e^B (if and) only if A commutes with B.

Application to f(z) = exp(z) (Figure 2)

- Suppose $\sigma(A), \sigma(B) \subset \mathcal{D}_1$. Then A is a polynomial in e^A and B is a polynomial in e^B, so e^A commutes with e^B (if and) only if A commutes with B.
- The implication is entirely conceptual; no need for power series, limits, or other analytic arguments.

Application to f(z) = exp(z) (Figure 2)

- Suppose $\sigma(A), \sigma(B) \subset \mathcal{D}_1$. Then A is a polynomial in e^A and B is a polynomial in e^B, so e^A commutes with e^B (if and) only if A commutes with B.
- The implication is entirely conceptual; no need for power series, limits, or other analytic arguments.
- But this is only a local result about the entire analytic function e^z and its associated primary matrix function e^A, defined for all $A \in M_n$.

Application to f(z) = exp(z) (Figure 2)

- Suppose $\sigma(A), \sigma(B) \subset \mathcal{D}_1$. Then A is a polynomial in e^A and B is a polynomial in e^B, so e^A commutes with e^B (if and) only if A commutes with B.
- The implication is entirely conceptual; no need for power series, limits, or other analytic arguments.
- But this is only a local result about the entire analytic function e^z and its associated primary matrix function e^A, defined for all $A \in M_n$.
- Next: a global result

Kronecker Sums and Commutativity

- $$\operatorname{vec}(YXZ) = \left(Z^T \otimes Y\right) \operatorname{vec} X \qquad \text{(VTP)}$$

Kronecker Sums and Commutativity

- $$\operatorname{vec}(YXZ) = \left(Z^T \otimes Y\right) \operatorname{vec} X \qquad \text{(VTP)}$$

- $$\begin{aligned} \operatorname{vec}(XY - YX) &= \operatorname{vec}(IXY - YXI) \\ &= \underbrace{(Y^T \otimes I - I \otimes Y)}_{K_Y} \operatorname{vec} X \qquad \text{(K)} \end{aligned}$$

Kronecker Sums and Commutativity

- $$\operatorname{vec}(YXZ) = \left(Z^T \otimes Y\right) \operatorname{vec} X \qquad \text{(VTP)}$$

- $$\begin{aligned} \operatorname{vec}(XY - YX) &= \operatorname{vec}(IXY - YXI) \\ &= \underbrace{(Y^T \otimes I - I \otimes Y)}_{K_Y} \operatorname{vec} X \end{aligned} \qquad \text{(K)}$$

- $$XY - YX = 0 \Leftrightarrow K_Y \operatorname{vec} X = 0$$

An Entire Analytic Function

- $f(z) = \frac{e^z - 1}{z}$ if $z \neq 0, f(0) = 1$

An Entire Analytic Function

- $f(z) = \frac{e^z - 1}{z}$ if $z \neq 0$, $f(0) = 1$
- $f(z)z = e^z - 1$ for all $z \in \mathbb{C}$

An Entire Analytic Function

- $f(z) = \frac{e^z - 1}{z}$ if $z \neq 0, f(0) = 1$
- $f(z)z = e^z - 1$ for all $z \in \mathbb{C}$
- For any $Y \in M_n$ we have

$$\begin{aligned} f(K_Y)K_Y &= e^{K_Y} - I = e^{Y^T \otimes I - I \otimes Y} - I \\ &= e^{Y^T \otimes I} e^{-I \otimes Y} - I \\ &= e^{Y^T} \otimes e^{-Y} - I \end{aligned}$$

An Entire Analytic Function

- $f(z) = \frac{e^z - 1}{z}$ if $z \neq 0$, $f(0) = 1$
- $f(z)z = e^z - 1$ for all $z \in \mathbb{C}$
- For any $Y \in M_n$ we have

$$\begin{aligned} f(K_Y)K_Y &= e^{K_Y} - I = e^{Y^T \otimes I - I \otimes Y} - I \\ &= e^{Y^T \otimes I} e^{-I \otimes Y} - I \\ &= e^{Y^T} \otimes e^{-Y} - I \end{aligned}$$

- $f(K_Y)$ is singular if and only if $2m\pi i \in \sigma(K_Y)$ for some $\pm$integer $m \neq 0$

An Entire Analytic Function

- $f(z) = \frac{e^z - 1}{z}$ if $z \neq 0$, $f(0) = 1$
- $f(z)z = e^z - 1$ for all $z \in \mathbb{C}$
- For any $Y \in M_n$ we have

$$\begin{aligned} f(K_Y)K_Y &= e^{K_Y} - I = e^{Y^T \otimes I - I \otimes Y} - I \\ &= e^{Y^T \otimes I} e^{-I \otimes Y} - I \\ &= e^{Y^T} \otimes e^{-Y} - I \end{aligned}$$

- $f(K_Y)$ is singular if and only if $2m\pi i \in \sigma(K_Y)$ for some $\pm$integer $m \neq 0$
- If $f(K_Y)$ is nonsingular, then

$$K_Y = f(K_Y)^{-1} \left(e^{Y^T} \otimes e^{-Y} - I \right) \qquad (\bigstar)$$

First Step

- **Claim:** If $e^A e^B = e^B e^A$ (H) then $Ae^B = e^B A$ *provided that* $f(K_A)$ is nonsingular

First Step

- **Claim:** If $e^A e^B = e^B e^A$ (H) then $Ae^B = e^B A$ *provided that* $f(K_A)$ is nonsingular
- $$\begin{aligned}
\operatorname{vec}(e^B A - Ae^B) &\overset{K}{=} K_A \operatorname{vec} e^B \overset{\bigstar}{=} f(K_A)^{-1}\left(e^{A^T} \otimes e^{-A} - I\right) \operatorname{vec} e^B \\
&= f(K_A)^{-1}\left((e^{A^T} \otimes e^{-A}) \operatorname{vec} e^B - \operatorname{vec} e^B\right) \\
&\overset{VTP}{=} f(K_A)^{-1} \operatorname{vec}\left(e^{-A} e^B e^A - e^B\right) \\
&\overset{H}{=} f(K_A)^{-1} \operatorname{vec}\left(e^{-A} e^A e^B - e^B\right) \\
&= f(K_A)^{-1} \operatorname{vec}\left(e^B - e^B\right) = 0
\end{aligned}$$

Second Step

- **Claim:** If $Ae^B = e^B A$ (H) then $AB = BA$ *provided that* $f(K_B)$ is nonsingular

- **Claim:** If $Ae^B = e^B A$ (H) then $AB = BA$ *provided that* $f(K_B)$ is nonsingular
- $$\begin{aligned}
\operatorname{vec}(AB - BA) &\overset{K}{=} K_B \operatorname{vec} A \overset{\bigstar}{=} f(K_B)^{-1}\left(e^{B^T} \otimes e^{-B} - I\right)\operatorname{vec} A \\
&= f(K_B)^{-1}\left(\left(e^{B^T} \otimes e^{-B}\right)\operatorname{vec} A - \operatorname{vec} A\right) \\
&\overset{VTP}{=} f(K_B)^{-1}\operatorname{vec}\left(e^{-B}Ae^{B} - A\right) \\
&\overset{H}{=} f(K_B)^{-1}\operatorname{vec}\left(e^{-B}e^{B}A - A\right) \\
&= f(K_B)^{-1}\operatorname{vec}(A - A) = 0
\end{aligned}$$

The Nonsingularity Condition

- $f(K_Y)$ is nonsingular if and only if $2m\pi i \notin \sigma(K_Y)$ for every $\pm$integer $m \neq 0$

The Nonsingularity Condition

- $f(K_Y)$ is nonsingular if and only if $2m\pi i \notin \sigma(K_Y)$ for every ±integer $m \neq 0$
- $\sigma(K_Y)$ consists entirely of differences of eigenvalues of Y (not all differences need occur)

The Nonsingularity Condition

- $f(K_Y)$ is nonsingular if and only if $2m\pi i \notin \sigma(K_Y)$ for every $\pm$integer $m \neq 0$
- $\sigma(K_Y)$ consists entirely of differences of eigenvalues of Y (not all differences need occur)
- If we insist that no difference of eigenvalues of A, and no difference of eigenvalues of B, is a nonzero $\pm$ integer multiple of $2\pi i$, then $f(K_Y)$ is nonsingular and $e^A e^B = e^B e^A \Rightarrow AB = BA$.

The Nonsingularity Condition

- $f(K_Y)$ is nonsingular if and only if $2m\pi i \notin \sigma(K_Y)$ for every $\pm$integer $m \neq 0$
- $\sigma(K_Y)$ consists entirely of differences of eigenvalues of Y (not all differences need occur)
- If we insist that no difference of eigenvalues of A, and no difference of eigenvalues of B, is a nonzero $\pm$integer multiple of $2\pi i$, then $f(K_Y)$ is nonsingular and $e^A e^B = e^B e^A \Rightarrow AB = BA$.
- Another approach: Make an assumption on the entries of A and B that makes it impossible for any $\pm$integer multiple of $2\pi i$ to be in the field generated by the zeroes of their characteristic polynomials.

The Nonsingularity Condition

- $f(K_Y)$ is nonsingular if and only if $2m\pi i \notin \sigma(K_Y)$ for every $\pm$integer $m \neq 0$
- $\sigma(K_Y)$ consists entirely of differences of eigenvalues of Y (not all differences need occur)
- If we insist that no difference of eigenvalues of A, and no difference of eigenvalues of B, is a nonzero $\pm\text{integer}$ multiple of $2\pi i$, then $f(K_Y)$ is nonsingular and $e^A e^B = e^B e^A \Rightarrow AB = BA$.
- Another approach: Make an assumption on the entries of A and B that makes it impossible for any $\pm\text{integer}$ multiple of $2\pi i$ to be in the field generated by the zeroes of their characteristic polynomials.
- For example, if all the entries of A and B are algebraic numbers, then the zeroes of their characteristic polynomials are all algebraic numbers, so $\pm\text{integer}$ multiples of $2\pi i$ are excluded.

References

R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

R. Horn and G. Piepmeyer, Two applications of the theory of primary matrix functions, *Linear Algebra Appl.* 361 (2003) 99-106.

C. Schmoeger, Remarks on Commuting Exponentials in Banach Algebras, *Proc. Amer. Math. Soc.* 127 (1999) 1337-1338.

E. Wermuth, A Remark on Commuting Operator Exponentials, *Proc. Amer. Math. Soc.* 125 (1997) 1685-1688.