MANCHESTER 1824

The University of Manchester

Five Theorems in Matrix Analysis, with Applications

Nick Higham
School of Mathematics
The University of Manchester

higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/

Dundee (EMS)—March 17, 2006

Outline

$f(AB)$ and $f(BA)$
- WMFME
- $\Lambda(AB)$ and $\Lambda(BA)$
- $f(\alpha I + AB)$

Symmetrization

Jordan Structure of $f(A)$

Matrix Sign Identities

$f(AB)$ and $f(BA)$

For $A, B \in \mathbb{C}^{n\times n}$, $AB \neq BA$.

- How are AB and BA related?
- How are $f(AB)$ and $f(BA)$ related?
- Same question if $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$.
- Generalize to $f(\alpha I_m + AB)$ and $f(\alpha I_n + BA)$.

Sherman–Morrison–Woodbury Formula

If $U, V \in \mathbb{C}^{n\times p}$ and $I_p + V^*A^{-1}U$ is nonsingular then

$$(A + UV^*)^{-1} = A^{-1} - A^{-1}U(I_p + V^*A^{-1}U)^{-1}V^*A^{-1}.$$

Obtained, using $A + UV^* = A(I + A^{-1}U \cdot V^*)$, from its simpler version

$$(I_m + AB)^{-1} = I - A(I_n + BA)^{-1}B \qquad \begin{cases} A \in \mathbb{C}^{m\times n} \\ B \in \mathbb{C}^{n\times m} \end{cases}$$

World's Most Fundamental Matrix Equation

World's Most Fundamental Matrix Equation

$(I + AB)A = A(I + BA)$, or

$$(AB)A = A(BA).$$

Application of WMFME

$$\begin{aligned} & (AB)A = A(BA) \\ \Rightarrow \quad & (AB)^2A = ABA(BA) = A(BA)^2. \end{aligned}$$

In general, for any poly p,

$$p(AB)A = Ap(BA).$$

- Does the same hold for arbitrary f?

AB and BA

If A, B square and A nonsingular, WMFME implies $AB = A(BA)A^{-1}$, so $\Lambda(AB) = \Lambda(BA)$.

AB and BA

If A, B square and A nonsingular, WMFME implies $AB = A(BA)A^{-1}$, so $\Lambda(AB) = \Lambda(BA)$.

Theorem (Flanders, 1951)

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$.

- *The nonzero eigenvalues of AB have the same Jordan structure as the nonzero eigenvalues of BA.*
- *Any zero eigenvalues appear in Jordan blocks of AB and BA differing in size by at most 1.*

Putnam Problem 1990-A5

If $A, B \in \mathbb{C}^{n \times n}$ **does** $ABAB = 0$ **imply** $BABA = 0$**?**

Putnam Problem 1990-A5

If $A, B \in \mathbb{C}^{n\times n}$ does $ABAB = 0$ imply $BABA = 0$?

Yes for $n \le 2$; no for $n > 2$.

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$(AB)^2 = 0, \qquad (BA)^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Tridiagonal Toeplitz Matrices

$$T_n(c,d,e) = \begin{bmatrix} d & e & & \\ c & d & \ddots & \\ & \ddots & \ddots & e \\ & & c & d \end{bmatrix}.$$

Eigenvalues known explicitly:

$$d + 2(ce)^{1/2}\cos(k\pi/(n+1)), \qquad k = 1\colon n.$$

- What about simple modifications of T_n, e.g. to the (1,1) and (n,n) elements?

Second Difference Matrix

$$T_n = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix},$$

$$\widetilde{T}_n = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}.$$

Second Difference Matrix (cont.)

Define

$$L = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ & & & -1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times n}.$$

Then $T_n = L^T L$, $\widetilde{T}_{n+1} = LL^T$.

So $\Lambda(\widetilde{T}_{n+1}) = \Lambda(T_n) \cup \{0\}$ (Strang, 2005).

Example:

```
n = 6; L = gallery('triw',n,-1,1)';
L = L(:,1:n-1), A = L*L', B = L'*L
```

Definition of Matrix Function

Let A have distinct eigenvalues $\lambda_1, \dots, \lambda_s$, and let n_i be order of the largest Jordan block in which λ_i appears.

Definition (Sylvester, 1883)

$f(A) := r(A)$, where r is the unique Hermite interpolating polynomial of degree less than $\sum_{i=1}^{s} n_i$ that satisfies

$$r^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \qquad j = 0\colon n_i - 1, \quad i = 1\colon s.$$

$f(AB)$ and $f(BA)$

Recall that for any **polynomial** p,

$$Ap(BA) = p(AB)A.$$

$f(AB)$ and $f(BA)$

Recall that for any **polynomial** p,

$$Ap(BA) = p(AB)A.$$

Lemma

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$ and let $f(AB)$ and $f(BA)$ be defined. Then

$$Af(BA) = f(AB)A.$$

$f(AB)$ and $f(BA)$

Recall that for any **polynomial** p,

$$Ap(BA) = p(AB)A.$$

Lemma

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$ and let $f(AB)$ and $f(BA)$ be defined. Then

$$Af(BA) = f(AB)A.$$

Proof. There is a single polynomial p such that $f(AB) = p(AB)$ and $f(BA) = p(BA)$. Hence

$$Af(BA) = Ap(BA) = p(AB)A = f(AB)A.$$

Special Case

Take $f(t) = t^{1/2}$. When AB (and hence also BA) has no eigenvalues on $\mathbb{R}^-$,

$$A(BA)^{1/2} = (AB)^{1/2}A.$$

Special Case

Take $f(t) = t^{1/2}$. When AB (and hence also BA) has no eigenvalues on $\mathbb{R}^-$,

$$A(BA)^{1/2} = (AB)^{1/2}A.$$

Useful, but

$$Af(BA) = f(AB)A$$

cannot be solved for $f(BA)$ in terms of $f(AB)$.

Theorem (Harris 1993; H 2005)

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$, with $m \geq n$, and assume BA is nonsingular. Then

$$f(\alpha I_m + AB) = f(\alpha)I_m + A(BA)^{-1}\big(f(\alpha I_n + BA) - f(\alpha)I_n\big)B.$$

Theorem (Harris 1993; H 2005)

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$, with $m \geq n$, and assume BA is nonsingular. Then

$$f(\alpha I_m + AB) = f(\alpha)I_m + A(BA)^{-1}\big(f(\alpha I_n + BA) - f(\alpha)I_n\big)B.$$

Proof. Define $g(X) = X^{-1}\big(f(\alpha I + X) - f(\alpha I)\big)$.
Then $f(\alpha I + X) = f(\alpha)I + Xg(X)$.
Hence, using the lemma,

$$\begin{aligned} f(\alpha I_m + AB) &= f(\alpha)I_m + ABg(AB) \\ &= f(\alpha)I_m + Ag(BA)B \\ &= f(\alpha)I_m + A(BA)^{-1}\big(f(\alpha I_n + BA) - f(\alpha)I_n\big)B. \end{aligned}$$

Example: Rank 2 Perturbation of I

Consider $f(\alpha I_n + uv^* + xy^*)$, where $u, v, x, y \in \mathbb{C}^n$. Write

$$uv^* + xy^* = [\, u \quad x \,] \begin{bmatrix} v^* \\ y^* \end{bmatrix} \equiv AB.$$

Then

$$C := BA = \begin{bmatrix} v^*u & v^*x \\ y^*u & y^*x \end{bmatrix} \in \mathbb{C}^{2\times 2}.$$

$$f(\alpha I_n + uv^* + xy^*) = f(\alpha) I_n + [\, u \quad x \,]\, C^{-1} \big(f(\alpha I_2 + C) - f(\alpha) I_2 \big) \begin{bmatrix} v^* \\ y^* \end{bmatrix}$$

Example: Rank 2 Perturbation of I

Consider $f(\alpha I_n + uv^* + xy^*)$, where $u, v, x, y \in \mathbb{C}^n$. Write

$$uv^* + xy^* = [\, u \quad x \,] \begin{bmatrix} v^* \\ y^* \end{bmatrix} \equiv AB.$$

Then

$$C := BA = \begin{bmatrix} v^*u & v^*x \\ y^*u & y^*x \end{bmatrix} \in \mathbb{C}^{2\times 2}.$$

$$f(\alpha I_n + uv^* + xy^*) = f(\alpha)I_n + [\, u \quad x \,]\, C^{-1}\big(f(\alpha I_2 + C) - f(\alpha)I_2\big) \begin{bmatrix} v^* \\ y^* \end{bmatrix}$$

For $A \in \mathbb{C}^{2\times 2}$, $f(A) = f(\lambda_1)I + f[\lambda_1, \lambda_2](A - \lambda_1 I).$

Outline

- $f(AB)$ and $f(BA)$
  - WMFME
  - $\Lambda(AB)$ and $\Lambda(BA)$
  - $f(\alpha I + AB)$
- Symmetrization
- Jordan Structure of $f(A)$
- Matrix Sign Identities

Symmetrization

Theorem (Frobenius, 1910)

For any $A \in \mathbb{F}^{n\times n}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) there exist symmetric $S_1, S_2 \in \mathbb{F}^{n\times n}$, either one of which can be taken nonsingular, such that $A = S_1 S_2$.

Symmetrization

Theorem (Frobenius, 1910)

For any $A \in \mathbb{F}^{n\times n}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) there exist symmetric $S_1, S_2 \in \mathbb{F}^{n\times n}$, either one of which can be taken nonsingular, such that $A = S_1 S_2$.

Implication

The generalized eigenproblem $Ax = \lambda Bx$ with symmetric A and B has no special eigenproperties: equivalent to $Cx := B^{-1}Ax = \lambda x$, with C arbitrary.

Proof

Rational canonical form says A is similar to a direct sum of companion matrices: $A = X^{-1}CX$. But $S_1^{-1}C = S_2$:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -\beta_2 \\ 1 & -\beta_2 & -\beta_1 \end{bmatrix} \begin{bmatrix} \beta_2 & \beta_1 & \beta_0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -\beta_2 & 0 \\ 0 & 0 & \beta_0 \end{bmatrix}.$$

Then $A = X^{-1}S_1S_2X = X^{-1}S_1X^{-T} \cdot X^TS_2X \equiv \widetilde{S}_1\widetilde{S}_2$.

Proof

Rational canonical form says A is similar to a direct sum of companion matrices: $A = X^{-1}CX$. But $S_1^{-1}C = S_2$:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -\beta_2 \\ 1 & -\beta_2 & -\beta_1 \end{bmatrix} \begin{bmatrix} \beta_2 & \beta_1 & \beta_0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -\beta_2 & 0 \\ 0 & 0 & \beta_0 \end{bmatrix}.$$

Then $A = X^{-1}S_1S_2X = X^{-1}S_1X^{-T} \cdot X^TS_2X \equiv \widetilde{S}_1\widetilde{S}_2$.

Theorem

For any $A \in \mathbb{F}^{n\times n}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) there exists a nonsingular symmetric S such that $A = S^{-1}A^TS$.

Application to Polynomial Zero-Finding

Lancaster (1961) takes companion linearization $\lambda I - C$ for scalar poly $p(t) = a_k t^k + \cdots + a_1 t + a_0$:

$$C = \begin{bmatrix} -a_{k-1}/a_k & -a_{k-2}/a_k & \dots & -a_0/a_k \\ 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{bmatrix}.$$

We can write $C = S_1^{-1} S_2$ with S_1, S_2 symm. So

- $S_1(\lambda I - C) = \lambda S_1 - S_1 C$ is a **symm. pencil**.
- Ditto $S_1 C^{\ell-1}(\lambda I - C) = \lambda S_1 C^{\ell-1} - S_1 C^{\ell}$ for $\ell \geq 1$.

Lancaster takes

$$S_1 = \begin{bmatrix} & & & a_k \\ & & \iddots & a_{k-1} \\ & \iddots & \iddots & \vdots \\ a_k & a_{k-1} & \dots & a_1 \end{bmatrix}.$$

Matrix Polynomial Case

This construction generalizes immediately to **matrix polynomials** and provides **block symmetric** pencils $\lambda X + Y$ $[X_{ij} = X_{ji}, i \neq j]$.

- What space of pencils is generated?
- What happens as ℓ increases?
- Is there anything special about this particular S_1?
- How are ei'vecs of the pencils related to those of P?

Matrix Polynomial Case

This construction generalizes immediately to **matrix polynomials** and provides **block symmetric** pencils $\lambda X + Y$ $[X_{ij} = X_{ji},\ i \neq j]$.

- What space of pencils is generated?
- What happens as ℓ increases?
- Is there anything special about this particular S_1?
- How are ei'vecs of the pencils related to those of P?

Answered via a new theory of vector spaces of linearizations:
H, D. S. Mackey, N. Mackey, Mehl, Mehrmann, Tisseur (2005)

Outline

$f(AB)$ and $f(BA)$

- WMFME
- $\Lambda(AB)$ and $\Lambda(BA)$
- $f(\alpha I + AB)$

Symmetrization

Jordan Structure of $f(A)$

Matrix Sign Identities

Function of Jordan block

$$A = Z\mathrm{diag}(J_1, \dots, J_p)Z^{-1} \Rightarrow f(A) = Z\mathrm{diag}(f(J_1), \dots, f(J_p))Z^{-1}.$$

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k \times m_k},$$

$$f(J_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \dots & \dfrac{f^{(m_k-1)})(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}.$$

Theorem

Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues λ_k.

1. *If $f'(\lambda_k) \neq 0$ then for every $J(\lambda_k)$ in A there is a Jordan block of the **same** size in $f(A)$ for $f(\lambda_k)$.*

Theorem

Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues λ_k.

1. *If $f'(\lambda_k) \neq 0$ then for every $J(\lambda_k)$ in A there is a Jordan block of the **same** size in $f(A)$ for $f(\lambda_k)$.*
2. *Let $f'(\lambda_k) = f''(\lambda_k) = \cdots = f^{(\ell-1)}(\lambda_k) = 0$ but $f^{(\ell)}(\lambda_k) \neq 0$, where $\ell \geq 2$, and consider $J(\lambda_k)$ of size r in A.*

   *(i) If $\ell \geq r$, $J(\lambda_k)$ splits into r 1×1 **Jordan blocks** for $f(\lambda_k)$ in $f(A)$.*

   (ii) If $\ell \leq r-1$, $J(\lambda_k)$ splits into Jordan blocks for $f(\lambda_k)$ in $f(A)$ as follows:

   - $\ell - q$ ***Jordan blocks of size*** p,
   - q ***Jordan blocks of size*** $p+1$,

   where $r = \ell p + q$ with $0 \leq q \leq \ell - 1$, $p > 0$.

Theorem

Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues λ_k.

1. *If $f'(\lambda_k) \neq 0$ then for every $J(\lambda_k)$ in A there is a Jordan block of the* ***same*** *size in $f(A)$ for $f(\lambda_k)$.*
2. *Let $f'(\lambda_k) = f''(\lambda_k) = \cdots = f^{(\ell-1)}(\lambda_k) = 0$ but $f^{(\ell)}(\lambda_k) \neq 0$, where $\ell \geq 2$, and consider $J(\lambda_k)$ of size r in A.*

   (i) If $\ell \geq r$, $J(\lambda_k)$ splits into r 1×1 ***Jordan blocks*** *for $f(\lambda_k)$ in $f(A)$.*

   (ii) If $\ell \leq r-1$, $J(\lambda_k)$ splits into Jordan blocks for $f(\lambda_k)$ in $f(A)$ as follows:

   - *$\ell - q$* ***Jordan blocks of size*** *p,*
   - *q* ***Jordan blocks of size*** *$p+1$,*

   where $r = \ell p + q$ with $0 \leq q \leq \ell - 1$, $p > 0$.

Application: Matrix Logarithm

Find all solutions to $e^X = A$.

- Let A have JCF $A = Z\mathrm{diag}(J_k(\lambda_k))Z^{-1} = ZJZ^{-1}$.

Application: Matrix Logarithm

Find all solutions to $e^X = A$.

- Let A have JCF $A = Z\mathrm{diag}(J_k(\lambda_k))Z^{-1} = ZJZ^{-1}$.
- Since $\frac{d}{dx}e^x \neq 0$, X has Jordan form $J_X = \mathrm{diag}(J_k(\mu_k))$, where $\exp(\mu_k) = \lambda_k$ and hence $\mu_k = \log \lambda_k + 2j_k\pi i$.

Application: Matrix Logarithm

Find all solutions to $e^X = A$.

- Let A have JCF $A = Z\mathrm{diag}(J_k(\lambda_k))Z^{-1} = ZJZ^{-1}$.
- Since $\frac{d}{dx}e^x \neq 0$, X has Jordan form $J_X = \mathrm{diag}(J_k(\mu_k))$, where $\exp(\mu_k) = \lambda_k$ and hence $\mu_k = \log \lambda_k + 2j_k\pi i$.
- Now consider $L = \mathrm{diag}(L_k)$, where $L_k = \log(J_k(\lambda_k)) + 2j_k\pi iI$. Then $e^L = J$, so by same argument as above, L has Jordan form J_X, i.e., $X = WLW^{-1}$, some W.

Application: Matrix Logarithm

Find all solutions to $e^X = A$.

- Let A have JCF $A = Z\mathrm{diag}(J_k(\lambda_k))Z^{-1} = ZJZ^{-1}$.
- Since $\frac{d}{dx}e^x \neq 0$, X has Jordan form $J_X = \mathrm{diag}(J_k(\mu_k))$, where $\exp(\mu_k) = \lambda_k$ and hence $\mu_k = \log \lambda_k + 2j_k\pi i$.
- Now consider $L = \mathrm{diag}(L_k)$, where $L_k = \log(J_k(\lambda_k)) + 2j_k\pi iI$. Then $e^L = J$, so by same argument as above, L has Jordan form J_X, i.e., $X = WLW^{-1}$, some W.
- But $e^X = A$ implies $WJW^{-1} = We^LW^{-1} = ZJZ^{-1}$, or $(Z^{-1}W)J = J(Z^{-1}W)$.

Application: Matrix Logarithm

Theorem (Gantmacher, 1959)

Let $A \in \mathbb{C}^{n\times n}$ be nonsing. with JCF $A = Z\mathrm{diag}(J_k(\lambda_k))Z^{-1}$. All solutions to $e^X = A$ are given by

$$X = ZU\mathrm{diag}(L_1^{(j_1)}, L_2^{(j_2)}, \dots, L_p^{(j_p)})U^{-1}Z^{-1},$$

where

$$L_k^{(j_k)} = \log(J_k(\lambda_k)) + 2j_k\pi i I,$$

$\log(J_k(\lambda_k))$ *is the principal logarithm, j_k is an arbitrary integer, and U is an arbitrary nonsingular matrix commuting with J.*

Outline

$f(AB)$ and $f(BA)$
- WMFME
- $\Lambda(AB)$ and $\Lambda(BA)$
- $f(\alpha I + AB)$

Symmetrization

Jordan Structure of $f(A)$

Matrix Sign Identities

Matrix Sign Function

For $A \in \mathbb{C}^{n\times n}$ with Jordan canonical form

$$A = Z \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} Z^{-1},$$

where $\lambda(J_1) \in$ open LHP, $\lambda(J_2) \in$ open RHP,

$$\mathrm{sign}(A) = Z \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} Z^{-1}.$$

Introduced by Roberts (1971), who proposed Newton iter.

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}), \quad X_0 = A.$$

X_k converges quadratically to $\mathrm{sign}(A)$.

Matrix Sign Relations

For nonsingular $A \in \mathbb{C}^{n \times n}$ (Byers, 1984):

$$\operatorname{sign}\left(\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & U \\ U^* & 0 \end{bmatrix},$$

where $A = UH$ is the polar decomposition.

Matrix Sign Relations

For nonsingular $A \in \mathbb{C}^{n\times n}$ (Byers, 1984):

$$\mathrm{sign}\left(\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & U \\ U^* & 0 \end{bmatrix},$$

where $A = UH$ is the polar decomposition.

For $A \in \mathbb{C}^{n\times n}$ with no eigenvalues on $\mathbb{R}^-$ (H, 1997):

$$\mathrm{sign}\left(\begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}.$$

More General Matrix Sign Relation

Theorem (H, Mackey, Mackey, Tisseur, 2005)

Let $A, B \in \mathbb{C}^{n\times n}$ and suppose AB has no eigenvalues on $\mathbb{R}^-$. Then

$$\operatorname{sign}\left(\begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & C \\ C^{-1} & 0 \end{bmatrix},$$

where $C = A(BA)^{-1/2}$.

Proof. $P = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$ has no pure imaginary ei'vals. Hence

$$\begin{aligned}
\mathrm{sign}(P) = P(P^2)^{-1/2} &= \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \begin{bmatrix} AB & 0 \\ 0 & BA \end{bmatrix}^{-1/2} \\
&= \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \begin{bmatrix} (AB)^{-1/2} & 0 \\ 0 & (BA)^{-1/2} \end{bmatrix} \\
&= \begin{bmatrix} 0 & A(BA)^{-1/2} \\ B(AB)^{-1/2} & 0 \end{bmatrix} =: \begin{bmatrix} 0 & C \\ D & 0 \end{bmatrix}.
\end{aligned}$$

Now

$$I = (\mathrm{sign}(P))^2 = \begin{bmatrix} 0 & C \\ D & 0 \end{bmatrix}^2 = \begin{bmatrix} CD & 0 \\ 0 & DC \end{bmatrix},$$

so $D = C^{-1}$. $\square$

Puzzle

Proof of previous theorem shows that

$$A(BA)^{-1/2} = \left[B(AB)^{-1/2}\right]^{-1} = (AB)^{1/2}B^{-1}.$$

Why do we have equality?

Puzzle

Proof of previous theorem shows that

$$A(BA)^{-1/2} = \left[B(AB)^{-1/2}\right]^{-1} = (AB)^{1/2}B^{-1}.$$

Why do we have equality?

Recall

$$Af(BA) = f(AB)A.$$

Now

$$\begin{aligned} A(BA)^{-1/2} \cdot B(AB)^{-1/2} &= (AB)^{-1/2}A \cdot B(AB)^{-1/2} \\ &= (AB)^{-1/2}AB(AB)^{-1/2} \\ &= I. \end{aligned}$$

Application: Matrix Iterations

Apply any iteration for the matrix sign function to

$$\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}$$

and read off from the (1,2) block an iteration for polar factor U or $A^{1/2}$.

Applying the lemma to

$$\begin{bmatrix} 0 & A \\ A^\star & 0 \end{bmatrix}$$

can derive new iterations for the *generalized* polar decomposition this way (HMMT, 2005).

Summary

- $\lambda(AB)$ vs. $\lambda(BA)$: Flanders (1951).
- $f(\alpha I_m + AB)$: Harris (1993), H (2005).
- $A = S_1 S_2$: Frobenius (1910).
- Jordan structure of $f(J)$.
- $\mathrm{sign}(\left[\begin{smallmatrix} 0 & A \\ B & 0 \end{smallmatrix}\right])$: H, Mackey, Mackey, Tisseur (2005).

Bibliography I

A. J. Bosch.
Note on the factorization of a square matrix into two Hermitian or symmetric matrices.
SIAM Rev., 29(3):463–468, 1987.

Harley Flanders.
Elementary divisors of AB and BA.
Proc. Amer. Math. Soc., 2(6):871–874, 1951.

F. R. Gantmacher.
The Theory of Matrices, volume two.
Chelsea, New York, 1959.

Bibliography II

- P. R. Halmos.
  Bad products of good matrices.
  Linear and Multlinear Algebra, 29:1–20, 1991.

- Lawrence A. Harris.
  Computation of functions of certain operator matrices.
  Linear Algebra Appl., 194:31–34, 1993.

- Nicholas J. Higham.
  Functions of a Matrix: Theory and Computation.
  Book in preparation.

Bibliography III

- Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur.
  Functions preserving matrix groups and iterations for the matrix square root.
  SIAM J. Matrix Anal. Appl., 26(3):849–877, 2005.

- Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur.
  Symmetric linearizations for matrix polynomials.
  MIMS EPrint 2005.25, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, November 2005.
  Submitted to SIAM J. Matrix Anal. Appl.

Bibliography IV

Roger A. Horn and Dennis I. Merino.
Contragredient equivalence: A canonical form and some applications.
Linear Algebra Appl., 214:43–92, 1995.

Charles R. Johnson and Eric Schreiner.
The relationship between *AB* and *BA*.
Amer. Math. Monthly, 103(7):578–582, 1996.

Leonard F. Klosinski, Gerald L. Alexanderson, and Loren C. Larson.
The fifty-first William Lowell Putnam mathematical competition.
Amer. Math. Monthly, 98(8):719–727, 1991.

Bibliography V

- Peter Lancaster.
  Symmetric transformations of the companion matrix.
  NABLA: Bulletin of the Malayan Math. Soc., 8:146–148, 1961.
- Heydar Radjavi.
  Products of Hermitian matrices and symmetries.
  Proc. Amer. Math. Soc., 21(2):369–372, 1969.
- O. Taussky.
  The role of symmetric matrices in the study of general matrices.
  Linear Algebra Appl., 5:147–154, 1972.

Bibliography VI

- Olga Taussky and Hans Zassenhaus.
  On the similarity transformation between a matrix and its transpose.
  Pacific J. Math., 9:893–896, 1959.

- R. C. Thompson.
  On the matrices AB and BA.
  Linear Algebra Appl., 1:43–58, 1968.

World's Most Fundamental Matrix Equation

$$(I + AB)^{-1} = I - A(I + BA)^{-1}B.$$

$$\begin{aligned} I &= I + AB - (I + AB)A(I + BA)^{-1}B \\ &= I + AB - A(I + BA)(I + BA)^{-1}B \\ &= I + AB - AB \\ &= I \quad \surd \end{aligned}$$

World's Most Fundamental Matrix Equation

$$(I + AB)^{-1} = I - A(I + BA)^{-1}B.$$

$$\begin{aligned} I &= I + AB - (I + AB)A(I + BA)^{-1}B \\ &= I + AB - A(I + BA)(I + BA)^{-1}B \\ &= I + AB - AB \\ &= I \quad \surd \end{aligned}$$

Key equation: $(I + AB)A = A(I + BA)$, or

$$(AB)A = A(BA).$$