Ralph, the Matrix Sign Function, and the Polar Decomposition

Nick Higham
School of Mathematics
The University of Manchester

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

Memorial Workshop for Ralph Byers
TU Berlin, May 31-June 1, 2008

Matrix Sign Function

Let $A \in \mathbb{C}^{n\times n}$ have no pure imaginary eigenvalues. Given Jordan 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,

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

Alternative definition:

$$\operatorname{sign}(A) = A(A^2)^{-1/2}.$$

Introduced by Roberts (1971).

Polar Decomposition

Autonne (1902): for $A \in \mathbb{C}^{m\times n}$, $m \geq n$, $A = UH$, where $U \in \mathbb{C}^{m\times n}$ has orthonormal columns and $H \in \mathbb{C}^{n\times n}$ is Hermitian positive semidefinite.

- $H = (A^*A)^{1/2}$ is unique.
- U is unique iff A has full rank and then

$$U = A(A^*A)^{-1/2}.$$

Newton's Method in 1984

For matrix sign (RB):

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

For unitary polar factor of nonsingular $A \in \mathbb{C}^{n\times n}$ (NJH):

$$X_{k+1} = \tfrac{1}{2}(X_k + X_k^{-*}), \qquad X_0 = A.$$

Ralph's (1984) observation:

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

First written down in H (1994).

Ralph's 1987 Paper (1)

Ralph Byers. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.*, 85:267–279, 1987.

Hermitian stabilizing solution of

$$XFX - A^*X - XA - G = 0, \qquad F = F^*,\ G = G^*$$

obtained from sign of

$$W = \begin{bmatrix} A^* & G \\ F & -A \end{bmatrix}.$$

Ralph's 1987 Paper (2)

- Determinantal scaling:

$$X_{k+1} = \frac{1}{2}(\mu_k X_k + \mu_k^{-1} X_k^{-1}), \qquad \mu_k = |\det(X_k)|^{-1/n}.$$

  Minimizes $d(\mu X)$, where $d(X) = \sum_{i=1}^{n} (\log |\lambda_i(X)|)^2$.
- JW is Hermitian for $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ so can use LINPACK Hermitian inversion code.
- Iterative refinement of the ARE.
- Implementation details.

Ralph's 1987 Paper (3)

- "The first twelve iterations of (1) differ insignificantly from $W^{(k+1)} := W^{(k)}/2$. At $O(n^3)$ floating-point operations per iteration, this is an expensive way to divide by 2."
- "Trivial cases are handled trivially by (2)."
- "The matrix sign function can be regarded simply as a way to obtain a good initial guess for Newton's method."

What do you get if you cross a matrix with a patrix?

Sign Function Developments

- Numerical stability of sign for invariant subspaces: Byers (1986), Byers, He & Mehrmann (1997), Bai & Demmel (1998) .
- Generalization.
- Parallelization.
- Analysis of scalings (Kenney & Laub).
- Padé family of iterations (Kenney & Laub).
- Conditioning of sign.
- Lattice quantum chromodynamics application.

Sign Function Stability and Termination

H (2008):

- (Roughly): all superlinearly convergent sign iterations $X_{k+1} = g(X_k)$ are stable and have idempotent Fréchet derivative at $S = \mathrm{sign}(X_0)$ given by $L_g(S, E) = \frac{1}{2}(E - SES) = L(S, E)$ (Fréchet derivative of sign function at S).
- Terminate Newton iteration when

$$\|X_{k+1} - X_k\| \le \left(2\eta \frac{\|X_{k+1}\|}{\|X_k^{-1}\|}\right)^{1/2}.$$

Based on quadratic convergence.

Polar Decomposition

$A \in \mathbb{C}^{m\times n}$, $A = UH$.

- Defined only for $m \geq n$.
- Unique U only if A has full rank.

Partial Isometry

$U \in \mathbb{C}^{m\times n}$ is a **partial isometry** if $\|Ux\|_2 = \|x\|_2$ for all $x \in \mathrm{range}(U^*)$; i.e., U norm-preserving on orthogonal complement of its null space.

Lemma

For $U \in \mathbb{C}^{m\times n}$ each of the following conditions is equivalent to U being a partial isometry:

(a) $U^+ = U^*$,

(b) $UU^*U = U$,

(c) *the singular values of U are all* 0 *or* 1.

Canonical Polar Decomposition

Theorem

*$A \in \mathbb{C}^{m\times n}$ has a **unique** decomp $A = UH$ with $U \in \mathbb{C}^{m\times n}$ a partial isometry, $H \in \mathbb{C}^{n\times n}$ Hermitian pos semidef, and*

$$\textbf{equivalent conditions} \begin{cases} \mathrm{range}(U^*) = \mathrm{range}(H) \\ U^*U = HH^+ \\ \mathrm{null}(U) = \mathrm{null}(H) \\ \mathrm{range}(U) = \mathrm{range}(A) \\ UU^+ = AA^+ \end{cases}$$

$H = (A^*A)^{1/2}$ and $U = AH^+$. Moreover, if $A = P\left[\begin{smallmatrix}\Sigma_r & 0 \\ 0 & 0_{m-r,n-r}\end{smallmatrix}\right]Q^*$ is an SVD then

$$U = P\begin{bmatrix} I_r & 0 \\ 0 & 0_{m-r,n-r}\end{bmatrix}Q^*, \qquad H = Q\begin{bmatrix}\Sigma_r & 0 \\ 0 & 0_{n-r}\end{bmatrix}Q^*.$$

Discussion

- **Polar decomp**:
  - $m \geq n$.
  - U has orthonormal cols.
  - U is nearest matrix with orthonormal columns.
- **Canonical polar decomp**:
  - Any m and n.
  - U and H both rank deficient if A is.
  - U is nearest partial isometry satisfying range condition.
- Some authors refer to canonical polar decomp as "*generalized polar decomp*".

Newton Iteration

Nonsingular $A \in \mathbb{C}^{n\times n}$:

$$X_{k+1} = \tfrac{1}{2}(X_k + X_k^{-*}), \qquad X_0 = A.$$

Scaled iteration (H, 1986):

$$X_{k+1} = \tfrac{1}{2}\left(\mu_k X_k + \mu_k^{-1} X_k^{-*}\right), \quad \mu_k = \left(\frac{\|X_k^{-1}\|_1 \|X_k^{-1}\|_\infty}{\|X_k\|_1 \|X_k\|_\infty}\right)^{1/4}.$$

$\mu_k \approx \mu_k^{\mathrm{opt}} = \big(\sigma_1(X_k)\sigma_n(X_k)\big)^{-1/2}$, where latter gives convergence in d iterations, where A has d distinct singular values.

Stability of Scaled Newton

- Observed to be excellent, starting with H (1986).
- Computed factors proved backward stable, assuming matrix inverses are mixed backward–forward stable, by Byers & Xu (2008), following earlier analysis of Kielbasiński & Ziętak (2003) .

  See following talk for more. . .

Sign Connections

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

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

Very useful for deriving new polar & square root iterations/results from existing sign iterations: H (1997), H, Mackey, Mackey & Tisseur (2004, 2005).

E.g.,

$$\mathrm{sign}(A) = \frac{2}{\pi} A \int_0^\infty (t^2 I + A^2)^{-1}\, dt$$

implies

$$U = \frac{2}{\pi} A \int_0^\infty (t^2 I + A^* A)^{-1}\, dt.$$

Our “Gatlinburg” Papers

- Ralph Byers. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.*, 85: 267–279, 1987.
- Nicholas J. Higham. Computing the polar decomposition—with applications. *SIAM J. Sci. Statist. Comput.*, 7(4):1160–1174, 1986.

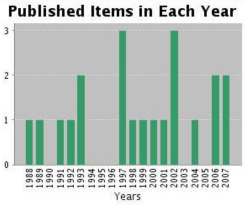
Web of Science®

<< Back to previous results list

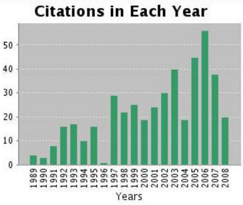
Citation Report Author=(byers r)

Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI.
Refined by: Subject Areas=(MATHEMATICS, APPLIED)

This report reflects citations to source items indexed within Web of Science. Perform a Cited Reference Search to include citations to items not indexed within Web of Science.



The latest 20 years are displayed.
View a graph with all years.



The latest 20 years are displayed.
View a graph with all years.

Results found: 24

Sum of the Times Cited [?] : 445

View Citing Articles
View without self-citations

Average Citations per Item [?] : 18.54

h-index [?] : 14

Results: **24** Page 1 of 3 Go Sort by: Times Cited

	Use the checkboxes to remove individual items from this Citation Report or restrict to items processed between 1945-1954 and 2008 Go	2004	2005	2006	2007	2008	Total	Average Citations per Year
		19	45	56	38	20	445	19.35
1.	Title: A BISECTION METHOD FOR MEASURING THE DISTANCE OF A STABLE MATRIX TO THE UNSTABLE MATRICES Author(s): BYERS R Source: **SIAM JOURNAL ON SCIENTIFIC AND STATISTICAL COMPUTING** Volume: **9** Issue: **5** Pages: **875-881** Published: **SEP 1988**	0	4	4	4	4	60	2.86
2.	Title: SOLVING THE ALGEBRAIC RICCATI EQUATION WITH THE MATRIX SIGN FUNCTION Author(s): BYERS R Source: **LINEAR ALGEBRA AND ITS APPLICATIONS** Volume: **85** Pages: **267-279** Published: **JAN 1987**	1	3	6	1	2	56	2.55
3.	Title: A HAMILTONIAN QR ALGORITHM Author(s): BYERS R Source: **SIAM JOURNAL ON SCIENTIFIC AND STATISTICAL COMPUTING** Volume: **7** Issue: **1** Pages: **212-229** Published: **JAN 1986**	0	4	2	2	1	52	2.26
4.	Title: A CHART OF NUMERICAL-METHODS FOR STRUCTURED EIGENVALUE PROBLEMS Author(s): BUNSEGERSTNER A, BYERS R, MEHRMANN V Source: **SIAM JOURNAL ON MATRIX ANALYSIS AND APPLICATIONS** Volume: **13** Issue: **2** Pages: **419-453** Published:	0	3	5	5	1	34	2.00

Google Scholar BETA

matrix sign function | Search

Advanced Scholar Search
Scholar Preferences
Scholar Help

Scholar **All articles** - Recent articles — Results **1** - **10**

[CITATION] Solving the algebraic Riccati equation with the **matrix sign function**
R Byers - Linear Algebra Appl, 1987
Cited by 123 - Related Articles - Web Search

A generalization of the **matrix-sign-function** solution for algebraic Riccati equations - all 6 versions »
JD Gardiner, AJ Laub - International Journal of Control, 1986 - informaworld.com
... The **matrix sign function** is applied to the solution of generalizedalgebraic Riccati equations without the need to reduce them to standard form. **...**
Cited by 77 - Related Articles - Web Search - Find It via JRUL

The **matrix sign function** - Find It via JRUL - all 5 versions »
CS Kenney, AJ Laub - Automatic Control, IEEE Transactions on, 1995 - ieeexplore.ieee.org
The **matrix sign function**. Kenney, CS Laub, AJ Dept. of Electr. & Comput. Eng., California Univ., Santa Barbara, CA; This paper appears **...**
Cited by 82 - Related Articles - Web Search - BL Direct

[CITATION] The **matrix sign function** and computations in systems - Find It via JRUL
ED Denman, AN Beavers - Appl. Math. Comput, 1976
Cited by 76 - Related Articles - Web Search

Solving stable generalized Lyapunov equations with the **matrix sign function** - Find It via JRUL - all 5 versions »
P Benner, ES Quintana-Ortí - Numerical Algorithms, 1999 - Springer
... with the **matrix sign function** ∗ Peter Benner a and Enrique S. Quintana-Ortı b **...** 2. The **matrix sign function** and Lyapunov equations **...**
Cited by 77 - Related Articles - Web Search - BL Direct

Linear model reduction and solution of the algebraic Riccati equation by use of the **sign function**† - all 2 versions »
JD Roberts - International Journal of Control, 1980 - informaworld.com
This definition also leads immediately to a **matrix sign function** of any **matrix** Z

References I

Z. Bai and J. W. Demmel.
Using the matrix sign function to compute invariant subspaces.
SIAM J. Matrix Anal. Appl., 19(1):205–225, 1998.

R. Byers.
Numerical stability and instability in matrix sign function based algorithms.
In C. I. Byrnes and A. Lindquist, editors, *Computational and Combinatorial Methods in Systems Theory*, pages 185–200. Elsevier Science Publishers B.V. (North-Holland), Amsterdam, The Netherlands, 1986.

References II

- R. Byers, C. He, and V. Mehrmann.
  The matrix sign function method and the computation of invariant subspaces.
  SIAM J. Matrix Anal. Appl., 18(3):615–632, 1997.

- R. Byers and H. Xu.
  A new scaling for Newton's iteration for the polar decomposition and its backward stability.
  SIAM J. Matrix Anal. Appl., 2008.
  To appear.

- N. J. Higham.
  The matrix sign decomposition and its relation to the polar decomposition.
  Linear Algebra Appl., 212/213:3–20, 1994.

References III

N. J. Higham.
Stable iterations for the matrix square root.
Numer. Algorithms, 15(2):227–242, 1997.

N. J. Higham.
Functions of Matrices: Theory and Computation.
Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur.
Computing the polar decomposition and the matrix sign decomposition in matrix groups.
SIAM J. Matrix Anal. Appl., 25(4):1178–1192, 2004.

References IV

N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur.
Functions preserving matrix groups and iterations for the matrix square root.
SIAM J. Matrix Anal. Appl., 26(3):849–877, 2005.

A. Kie␣lbasiński and K. Ziętak.
Numerical behaviour of Higham's scaled method for polar decomposition.
Numer. Algorithms, 32:105–140, 2003.