MANCHESTER 1824

The University of Manchester

The Canonical Polar Decomposition: Theory and Computation

Nick Higham
School of Mathematics
The University of Manchester

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

22nd Biennial Conference on Numerical Analysis
Dundee, June 2007

Moore-Penrose Pseudoinverse

The pseudo-inverse $A^+ \in \mathbb{C}^{n\times m}$ of $A \in \mathbb{C}^{m\times n}$ is the unique matrix $X = A^+$ satisfying the four **Moore–Penrose conditions**

$$\begin{array}{ll} \text{(i)}\;\; AXA = A, & \text{(ii)}\;\; XAX = X, \\ \text{(iii)}\;\; AX = (AX)^*, & \text{(iv)}\;\; XA = (XA)^*. \end{array}$$

When A has rank n, $A^+ = (A^*A)^{-1}A^*$.

Polar Decomposition

Theorem

Let $A \in \mathbb{C}^{m\times n}$ with $m \geq n$. $\exists\, U \in \mathbb{C}^{m\times n}$ with orthonormal columns and a unique Hermitian pos semidef $H \in \mathbb{C}^{n\times n}$ s.t. $A = UH$. $H = (A^*A)^{1/2}$. All U are

$$U = P \begin{bmatrix} I_r & 0 \\ 0 & W \end{bmatrix} Q^*,$$

where $A = P\left[\begin{smallmatrix} \Sigma_r & 0 \\ 0 & 0_{m-r,n-r} \end{smallmatrix}\right] Q^*$ is an SVD, $r = \operatorname{rank}(A)$, and $W \in \mathbb{C}^{(m-r)\times(n-r)}$ is arbitrary with orthonormal columns. If $\operatorname{rank}(A) = n$ then H is pos def and U is unique.

Comments

- 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 $x \in \text{range}(U^*)$; 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*

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

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

Proof

- Existence, but not uniqueness, easily proved using SVD.
- Whole proof can be done using pseudo-inverses.

Proof (flavour of)

$U^* = U^+$

$U = AH^+$ implies, using $HH^+ = H^+H$ and $(H^+)^* = H^+$,

$$UU^*U = AH^+ \cdot H^+A^* \cdot AH^+ = A(H^+)^2H^2H^+ = AH^+ = U.$$

Other Moore–Penrose conditions shown similarly.

Uniqueness

If $A = UH$ and $U^*U = HH^+$ then

$$A^*A = HU^*UH = HHH^+H = H^2,$$

so $H = (A^*A)^{1/2}$ uniquely determined. Then $AH^+ = UHH^+ = UU^*U = UU^+U = U$, gives U uniquely.

Best Approximation

Theorem (Fan & Hoffman, 1955)

Let $A \in \mathbb{C}^{m\times n}$ have polar decomp $A = UH$. Then $\|A - U\| = \min\{\, \|A - Q\| : Q^*Q = I_n \,\}$ for any unitarily invariant norm. The minimizer is unique for the Frobenius norm if A has full rank.

Best Approximation

Theorem (Fan & Hoffman, 1955)

Let $A \in \mathbb{C}^{m \times n}$ have polar decomp $A = UH$. Then $\|A - U\| = \min\{\, \|A - Q\| : Q^*Q = I_n \,\}$ for any unitarily invariant norm. The minimizer is unique for the Frobenius norm if A has full rank.

Theorem (Laskiewicz & Ziętak, 2006)

*Let $A \in \mathbb{C}^{m \times n}$ have **canonical polar decomp** $A = UH$. Then U solves*

$$\min\left\{\, \|A - Q\| : Q \in \mathbb{C}^{m \times n} \text{ is a partial isometry with } \operatorname{range}(Q) = \operatorname{range}(A) \,\right\}$$

for any unitarily invariant norm.

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

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

Theorem

The Newton iterates X_k converge quadratically to the unitary polar factor U of A, with

$$\|X_{k+1} - U\| \le \frac{1}{2}\|X_k^{-1}\|\,\|X_k - U\|^2.$$

Stopping Test

$$\delta_{k+1} := \frac{\|X_{k+1} - X_k\|}{\|X_{k+1}\|} \le \eta.$$

Let $X_* = \lim_{k\to\infty} X_k$. We have

$$X_{k+1} - X_* = (X_{k+1} - X_k) + (X_k - X_*),$$

and for suff. fast convergence $\|X_{k+1} - X_*\| \ll \|X_k - X_*\|$.
Hence δ_{k+1} actually approximates $\|X_k - X_*\|/\|X_*\|$.

- Test is approximating error in X_k.
- Yet X_{k+1} is our approximation.
- Tends to stop one iteration too late.

New Stopping Test

Suppose, for large k,

$$\|X_{k+1} - X_*\| \le c\|X_k - X_*\|^2.$$

As argued on prev slide, $\|X_{k+1} - X_k\| \approx \|X_k - X_*\|$. Hence

$$\|X_{k+1} - X_*\| \lesssim c\|X_{k+1} - X_k\|^2.$$

New Stopping Test

Suppose, for large k,

$$\|X_{k+1} - X_*\| \le c\|X_k - X_*\|^2.$$

As argued on prev slide, $\|X_{k+1} - X_k\| \approx \|X_k - X_*\|$. Hence

$$\|X_{k+1} - X_*\| \lesssim c\|X_{k+1} - X_k\|^2 \le \eta\|X_{k+1}\|.$$

New Stopping Test

Suppose, for large k,

$$\|X_{k+1} - X_*\| \le c\|X_k - X_*\|^2.$$

As argued on prev slide, $\|X_{k+1} - X_k\| \approx \|X_k - X_*\|$. Hence

$$\|X_{k+1} - X_*\| \lesssim c\|X_{k+1} - X_k\|^2 \le \eta\|X_{k+1}\|.$$

To get relative error $\le \eta$, stop when

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

For polar, $c = \frac{1}{2}\|X_k^{-1}\|$ so

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

Numerical Example

Binomial matrix, $n = 16$. $\kappa_U = 1.7 \times 10^3$, $\kappa_2(A) = 4.7 \times 10^3$.

k	$\frac{\|X_k - U\|_F}{\|U\|_F}$	$\|X_k^* X_k - I\|_F$	δ_{k+1}	$\|X_k\|_F$
1	1.7e+1	2.4e+3	2.6e+2	7.0e+1
2	1.4e+0	2.2e+1	7.0e+0	9.2e+0
3	1.3e-1	1.1e+0	1.2e+0	4.5e+0
4	2.6e-3	2.1e-2	1.3e-1	4.0e+0
5	1.4e-6	1.1e-5	2.6e-3	4.0e+0
6	1.3e-12	1.0e-11	1.4e-6	4.0e+0
7	3.3e-14	1.2e-15	1.3e-12	4.0e+0
8	3.3e-14	1.4e-15	2.3e-16	4.0e+0

My convergence test:

$$\text{if } \|X_{k+1} - X_k\|_F \leq (\text{tol\_cgce})^{1/2} \text{ or } \delta_{k+1} > \delta_k/2.$$

Summary

- There are two polar decomps:
  - (usual) one defined for $m \geq n$ only,
  - the **canonical polar decomp**, defined and unique for all m, n.
- New convergence test for Newton usually saves an iteration; applicable to any matrix Newton iteration.
- **Open problem**: develop ways to efficiently update the polar decomp after a small-normed or low rank perturbation.

References I

- N. J. Higham.
  Functions of Matrices: Theory and Computation.
  2008.
  Book to appear.

- B. Laskiewicz and K. Ziętak.
  Approximation of matrices and a family of Gander methods for polar decomposition.
  BIT, 46(2):345–366, 2006.

- R. Penrose.
  A generalized inverse for matrices.
  Proc. Cambridge Philos. Soc., 51(3):406–413, 1955.