On some known and open matrix nearness problems

Krystyna Ziętak

Institute of Mathematics and Computer Science
Wrocław University of Technology

Householder Symposium XVII

Zeuthen, June 1–6, 2008

Outline

$\mathcal{M}$ closed convex set in $\mathbb{C}^{m\times n}$

$$\min_{X\in\mathcal{M}} ||A-X||, \quad \text{spectral norm}$$

Strict spectral approximation

B is strict spectr. approx. of A if the vector $\sigma(A-B)$ of singular values is minimal with respect to ordinary lexicographic ordering in

$$\{\sigma(A-X) : X \in \mathcal{M}\}.$$

KZ SIMAX 1995,
Householder Symposium 1993 (Lake Arrowhead)

c_p norm

lexicographic ordering

[3, 3, 2, 0] is bigger then [3, 2, 2, 2]

$$\sigma(A) = [\sigma_1, \ldots, \sigma_n], \quad \text{vector of singular values}$$

$$||A||_p = ||\sigma||_p = \left(\sum_j \sigma_j(A)^p \right)^{1/p}, \qquad 1 \leq p \leq \infty$$

$p = \infty$ spectral norm

$p = 1$ trace norm

$p = 2$ Frobenius norm

Another characterization of strict spectral approx.

Theorem, KZ, 1997

$\widehat{X}$ is strict spectr. approx. to A **iff**

$$||A - X||_p > ||A - \widehat{X}||_p, \quad X \neq \widehat{X}, \quad X \in \mathcal{M},$$

for all p sufficiently large

Rogers and Ward 1981
c_p-minimal positive approximant of operator
in finite-dimensional complex Hilbert space

Conjecture

$\mathcal{M}$ linear subspace of matrices

Let

$$\min_{X\in\mathcal{M}} ||A - X||_p = ||A - X_p||_p.$$

Then

$$\lim_{p\to\infty} X_p = X_\infty \quad \text{strict spectral approx.}$$

Canonical trace approximant

c_1 - trace norm , $\mathcal{M}$ convex

Legg, Ward, 1985

$$X_p \to \widehat{X}_1, \quad \text{when} \quad p \to 1$$

where $\widehat{X}_1$ unique canonical trace approximant minimizing

$$\sum_{j=1}^{n} \sigma_j(A - X)\ln(\sigma_j(A - X))$$

over all trace approximants $X \in \mathcal{M}$ of A.

Vector case - strict Chebyshev approximation

Overdetermined real linear system

$$\min_{x\in\mathbb{R}^n} ||Ax-b||_p = ||Ax_p-b||_p, \qquad 1\leq p\leq\infty$$

Rice 1962 - strict Chebyshev solution

Descloux 1963, Pólya algorithm

$$\lim_{p\to\infty} ||Ax_p-b||_p = ||Ax_\infty-b||_\infty$$

Ax_∞ strict Cheb. approx. to b

- some generalizations of Descloux result
  Pólya algorithm on convex sets in $\mathbb{R}^n$
- Egger, Huotari 1989:
  - There exists closed, convex set in $\mathbb{R}^n$ for which best approx. x_p in l_p-norm to fixed $b \in \mathbb{R}^n$ fails to converge as $p \to \infty$.
  - If best approx. x_p converges it need not converge to strict Chebysh. approx.

Approximation by PSD matrices

$$A = B + iC$$

$$B = B^H,\ C = C^H, \quad \text{real and imaginary parts}$$

$$\min_{X \text{ is PSD}} ||A - X||, \qquad \text{spectral norm}$$

$$A = B + iC, \qquad B^H = B, \; C^H = C$$

Halmos approximant (1972)

Let

$$\delta(A) = \inf\{r > 0 : B + (r^2 I - C^2)^{1/2} \text{ and } r^2 I - C^2 \text{ are PSD}\}$$

Then

$$P_h(A) = B + (\delta^2 I - C^2)^{1/2}$$

Algorithm - Higham 1988

Approximation by matrices with spectrum in strip

$$\min_{X \in \mathbb{X}(\mathbb{S})} ||A - X||, \qquad \text{spectral norm}$$

$$\mathbb{X}(\mathbb{S}) = \{X \in \mathbb{C}^{n \times n} : \text{spectrum of } X \text{ is in } \mathbb{S}\}$$

$$\mathbb{S} = [0, \infty) \times [0, \infty) = \{x + iy : x \geq 0, \ y \geq 0\}$$

Khalil, Maher, *Numer. Functional Anal. Optim.* 2000

$$\mathbb{S}_a = [0, \infty) \times [0, a], \qquad \text{operators}$$

Khalil, Maher

spectrum of X in $\mathbb{S}_a = [0, \infty) \times [0, a]$

$$\min_X ||A - X||$$

BL, KZ; 2008

$$X = \mathrm{Re}(X) + i\mathrm{Im}(X) \equiv X_1 + iX_2$$

- spectrum of X_1 in $[0, \infty)$
- spectrum of X_2 in $[0, a]$

$$\mathbb{E}_1, \mathbb{E}_2 \quad \text{intervals,} \; [0, \infty) \text{ or } [0, a]$$

$$A = B + iC, \quad B = \mathrm{Re}(A), \quad C = \mathrm{Im}(A)$$

Corrected version of theorem of Khalil, Maher (BL, KZ 2008)

Let

$$\mathbb{K} = \{||A - X|| : X = X_1 + iX_2 \in \mathbb{C}^{n\times n}\}$$

$$X_1 \text{ has spectrum in } \mathbb{E}_1,$$

$$X_2 \text{ has spectrum in } \mathbb{E}_2$$

$$\mathbb{L} = \{r > 0 : B + [r^2 I - (C - \tilde{C})^2]^{1/2} \text{ for some } \tilde{C}\}$$

$$\tilde{C} \quad \text{Hermitian with spectrum in } \mathbb{E}_2.$$

Then

$$\delta(A) = \inf \mathbb{K} = \inf \mathbb{L}.$$

Best approximant

$$\widehat{X} = B + [\delta^2 I - (C - \tilde{C})^2]^{1/2}$$

for some Hermitian $\tilde{C}$ with spectrum in $\mathbb{E}_2$

Conjecture

$\tilde{C}$ is strict spectral approximant of C by Hermitian matrices with spectrum in $\mathbb{E}_2$.

Algorithm

Let $A = B + iC$

$$\mathbb{S} = [0, \infty) \times \mathbb{E}_2$$

- Compute strict spectral approx. $\hat{C}$ to C, spectrum $\hat{C}$ in $\mathbb{E}_2$.
- Compute Halmos approx. $\hat{P}_h$ to $A - i\hat{C}$ by Higham algorithm.
- Compute $\hat{X} = \hat{P}_h + iC$.

Example

Let $A = B + \mathrm{i}C$

$$B = \begin{bmatrix} 3 & -5 & 1 \\ -5 & -3 & 1 \\ 1 & 1 & -1 \end{bmatrix}, \qquad C = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

C_k Hermitian approx. of C with spectrum in $[0, \infty)$:

$$C_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \; C_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; C_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/8 & 1/2 \\ 0 & 1/2 & 5/2 \end{bmatrix}$$

C_1 - strict spectral approx.

C_4 Halmos approximant: $C_4 = \mathrm{diag}(0, 1, 3)$.

Let $X^{(\mathrm{k})} = P_k + \mathrm{i}C_k$

	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$\|\|A - X^{(\mathrm{k})}\|\|_2$	6.2087	6.2140	6.2156	6.2700

Special cases

Let $A = B + iC$

- If B PSD then conjecture true.
- If If B is not PSD and C has spectrum in $\mathbb{E}_2$ then true.

Conjecture

$$\delta(A) = \{r > 0 : B + [r^2 I - (C - \widehat{C})^2]^{1/2} \text{is PSD}\}$$

$\widehat{C}$ strict spectral approx. of C

Numerical experiments

Let $r > 0$ such that

$B + [r^2 I - (C - \tilde{C})^2]^{1/2}$, PSD for some $\tilde{C}$.

Then also

$$B + [r^2 I - (C - \hat{C})^2]^{1/2}$$

is PSD, where $\hat{C}$ is strict Chebysh. approx. to C.

Part II - partial isometry approximants

Polar decomposition

$$A = UH, \qquad A \in \mathbb{C}^{m \times n}, \quad m \geq n$$

U orthonormal columns
H - Hermitian positive definite

Approximation by unitary matrices

$$||A - U|| = \min_{Z-unitary} ||A - Z||$$

Fan, Hoffman 1955

$||\cdot|| - unitarily\ invariant$

Canonical polar decomposition

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

U - partial isometry (subunitary matrix)
H - Hermitian positive semidefinite

Partial isometry

$$||Ux||_2 = ||x||_2, \qquad x \in \mathrm{range}(U^H)$$

Partial isometry

Equivalent conditions

- $UU^HU = U$
- $U^H = U^\dagger$ Moore-Penrose inverse
- UU^H is an orthogonal projector
- singular values of U are 0 or 1

Ben-Israel, Greville, *Generalized Inverses*

Approximation by partial isometries

$$A = P\Sigma Q^H \in \mathbb{C}^{m\times n}$$

Theorem (B.L;K.Z., 2006)

- *for all partial isometries E of rank $r = \mathrm{rank}(A)$ we have* $||A - \hat{E}|| \leq ||A - E||$, *where* $\hat{E} = P \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q^H$

- *for all partial isometries E we have* $||A - \hat{X}|| \leq ||A - E|| \leq ||A + \tilde{E}||$, *where*

$$\hat{X} = P \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix} Q^H, \quad \tilde{E} = P \begin{bmatrix} I_n \\ 0 \end{bmatrix} Q^H.$$

q number of $\sigma_j(A) \geq \frac{1}{2}$

Algorithms

Algorithm I:

$\hat{X}$ is computed directly from the SVD of A

Algorithm II:

$\hat{X}$ is the limit of the sequence $X_k, X_0 = A$, generated by Gander's method with $f = 19/13$

Algorithm III:

Stage 1: computing polar decomposition $A = EH$
Stage 2: computing unitary polar factor E_C of $C = 2H - I$
Stage 3: computing $\hat{X} = \frac{1}{2}E(E_C + I_n)$

In algorithm III we apply Higham's method for computing polar factors

Algorithms

Algorithm I:

$\hat{X}$ is computed directly from the SVD of A

Algorithm II:

$\hat{X}$ is the limit of the sequence $X_k, X_0 = A$, generated by Gander's method with $f = 19/13$

Algorithm III:

Stage 1: computing polar decomposition $A = EH$
Stage 2: computing unitary polar factor E_C of $C = 2H - I$
Stage 3: computing $\hat{X} = \frac{1}{2}E(E_C + I_n)$

In algorithm III we apply Higham's method for computing polar factors

Algorithms

Algorithm I:

$\hat{X}$ is computed directly from the SVD of A

Algorithm II:

$\hat{X}$ is the limit of the sequence $X_k, X_0 = A$, generated by Gander's method with $f = 19/13$

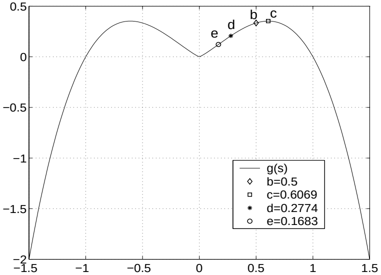
Algorithm III:

Stage 1: computing polar decomposition $A = EH$
Stage 2: computing unitary polar factor E_C of $C = 2H - I$
Stage 3: computing $\hat{X} = \frac{1}{2}E(E_C + I_n)$

In algorithm III we apply Higham's method for computing polar factors

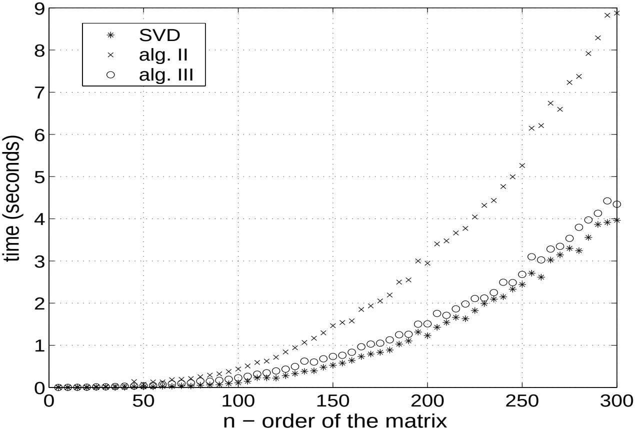
Test matrices

$$\text{Gander's method} \quad f = 19/13$$



A random with singular values:
$(0, b)$ 30 per cent; $(c, 1)$ 40 per cent; $(1, 3/2)$ rest

computing best partial isometry: average time



Part III - minimal rank approximants

Minimal rank approximation $A \in \mathbb{C}^{m \times n}$

$$\min_{B \text{ minimal rank}} ||A - B||_2 < \delta,$$

spectral norm, δ given, Golub 1968

Algorithm IV

- computing Hermitian polar factor H of A
- computing unitary polar factor E_D of $D = H - \delta I$
- computing $\hat{B} = \frac{1}{2}A(E_D + I)$

Minimal rank approximation $A \in \mathbb{C}^{m \times n}$

Algorithm IV-bis

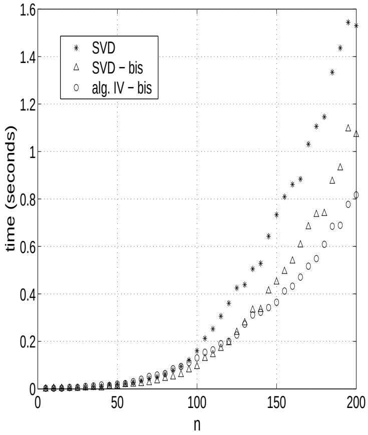
- computing unitary polar factor E of $A^H A - \delta^2 I$
- computing $\hat{B} = \frac{1}{2} A(E + I)$

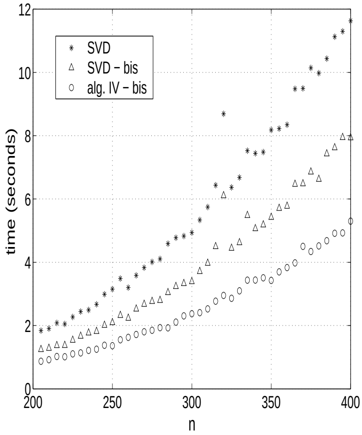
- **SVD**: computing $\hat{B}$ by means SVD applied to A
- **SVD-bis**: computing $\hat{B}$ by means SVD applied to $A^H A$

$$A \quad 2n \times n$$

Numerical tests for rectangular A, $2n \times n$

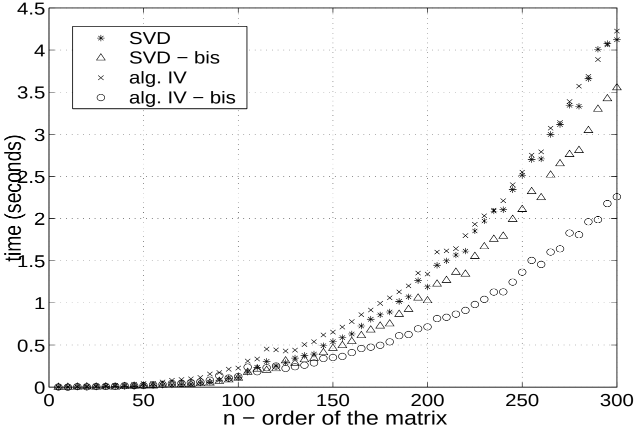
minimal rank approximant: average time





Numerical tests for square A

average time of computing minimal rank approximant



References

- KZ, Strict approximation of a matrix,
  SIMAX 16 (1995)
- BL, KZ, Approximation of matrices and family of Gander methods for polar decomposition,
  BIT Numer. Math. 46 (2006)
- BL, KZ, Approximation by matrices with restricted spectra,
  LAA 428 (2008)

Thank you for your atention!!!