Optimal Krylov Space Approximations to Rational Matrix Functions

Anne Greenbaum, joint with: Tyler Chen, Cameron Musco, Christopher Musco, and Natalie Wellen

University of Washington, Department of Applied Mathematics

Mar. 10, 2023

Lanczos and Arnoldi Algorithms

The Lanczos algorithm (for Hermitian matrices A) and the Arnoldi algorithm (for general square matrices A) construct an orthonormal basis for the Krylov space: $\text{span}\{b, Ab, \ldots, A^{k-1}b\}$. Let $Q_k := [q_1, \ldots, q_k]$ denote the n by k matrix whose columns are these basis vectors. Then

$$AQ_k = Q_k H_k + h_{k+1,k} q_{k+1} e_k^T = Q_{k+1} H_{k+1,k},$$

where H_k is a k by k symmetric tridiagonal (for Lanczos) or upper Hessenberg (for Arnoldi) matrix, e_k denotes the kth unit vector, and $H_{k+1,k}$ is the $k+1$ by k matrix obtained by appending the row $[0, \ldots, 0, h_{k+1,k}]$ to H_k.

We wish to use these basis vectors to approximate

$$R(A)b := D(A)^{-1} N(A)b.$$

Minimize the $D(A)^* D(A)$-Norm of the Error

Assume wlog that $\|b\| = 1$, $q_1 = b$. Choose y to minimize

$$\begin{aligned}
\|R(A)b - Q_k y\|^2_{D(A)^* D(A)} &= \langle D(A)^{-1} N(A)b - Q_k y, \\
&\quad D(A)^* D(A)(D(A)^{-1} N(A)b - Q_k y)\rangle \\
&= \langle N(A)b - D(A)Q_k y, N(A)b - D(A)Q_k y\rangle \\
&= \|N(A)Q_k e_1 - D(A)Q_k y\|^2_2.
\end{aligned}$$

Minimize the $D(A)^*D(A)$-Norm of the Error

Assume wlog that $\|b\| = 1$, $q_1 = b$. Choose y to minimize

$$\begin{aligned}
\|R(A)b - Q_k y\|^2_{D(A)^*D(A)} &= \langle D(A)^{-1}N(A)b - Q_k y, \\
&\quad D(A)^*D(A)(D(A)^{-1}N(A)b - Q_k y)\rangle \\
&= \langle N(A)b - D(A)Q_k y, N(A)b - D(A)Q_k y\rangle \\
&= \|N(A)Q_k e_1 - D(A)Q_k y\|_2^2.
\end{aligned}$$

Note that if $D(A) = A$ and $N(A) = I$, this is just GMRES (or MINRES):

$$\min_y \|A^{-1}b - Q_k y\|_{A^*A} = \min_y \|Q_k e_1 - AQ_k y\|_2 = \min_y \|e_1 - H_{k+1,k}y\|.$$

Minimize the $D(A)^*D(A)$-Norm of the Error, Cont.

$$
\begin{aligned}
AQ_k &= Q_{k+1}H_{k+1,k} \\
A^2Q_k &= AQ_{k+1}H_{k+1,k} = Q_{k+2}H_{k+2,k+1}H_{k+1,k} := Q_{k+2}H_{k+2,k} \\
&\vdots \\
A^jQ_k &= Q_{k+j}H_{k+j,k+j-1}H_{k+j-1,k+j-2}\cdots H_{k+1,k} := Q_{k+j}H_{k+j,k}.
\end{aligned}
$$

If $D(z) := \sum_{j=0}^{J} d_j z^j$ and $N(z) := \sum_{\ell=0}^{L} n_\ell z^\ell$, then

$$\|N(A)Q_k e_1 - D(A)Q_k y\|_2 =$$

$$\left\| \sum_{\ell=0}^{L} n_\ell Q_{k+\ell} H_{k+\ell,k} e_1 - \left(\sum_{j=0}^{J} d_j Q_{k+j} H_{k+j,k}\right) y \right\|_2 .$$

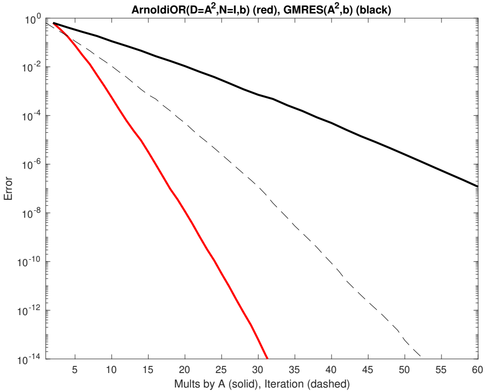
Minimize the $D(A)^*D(A)$-Norm of the Error, Cont.

Equivalently, if $m = \max\{J, L\}$,

$$\min_y \|N(H_{k+m})e_1 - D(H_{k+m})(:, 1:k)y\|_2.$$

Solve this least squares problem using standard QR factorization of the $k+m$ by k coefficient matrix. It is not upper Hessenberg if $\deg(D) > 1$, but $D(H_{k+m})(:, 1:k)$ differs from $D(H_{k+m-1})(:, 1:k-1)$ only in last row and last column. Apply previous Givens rotations to last column and determine new rotations to eliminate entries $k+1, \ldots, k+m$ in column k.

Example: Solve $A^2x = b$



Error Bounds

Let $S := D(A)^*D(A)$. Since $Q_k y = P_{k-1}(A)b$ where P_{k-1} is the $(k-1)^{st}$ degree polynomial that minimizes the S-norm of the error, we can write

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2} \min_{p_{k-1} \in \mathcal{P}_{k-1}} \|R(A) - p_{k-1}(A)\|_2.$$

Error Bounds

Let $S := D(A)^* D(A)$. Since $Q_k y = P_{k-1}(A)b$ where P_{k-1} is the $(k-1)^{st}$ degree polynomial that minimizes the S-norm of the error, we can write

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2} \min_{p_{k-1} \in \mathcal{P}_{k-1}} \|R(A) - p_{k-1}(A)\|_2.$$

If A is diagonalizable with eigendecomposition $A = V \Lambda V^{-1}$, then

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2} \kappa(V) \cdot$$

$$\min_{p_{k-1} \in \mathcal{P}_{k-1}} \max_{\lambda \in \Lambda(A)} |R(\lambda) - p_{k-1}(\lambda)|.$$

Error Bounds

Let $S := D(A)^* D(A)$. Since $Q_k y = P_{k-1}(A)b$ where P_{k-1} is the $(k-1)^{st}$ degree polynomial that minimizes the S-norm of the error, we can write

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2} \min_{p_{k-1} \in \mathcal{P}_{k-1}} \|R(A) - p_{k-1}(A)\|_2.$$

If A is diagonalizable with eigendecomposition $A = V \Lambda V^{-1}$, then

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2} \kappa(V) \cdot$$

$$\min_{p_{k-1} \in \mathcal{P}_{k-1}} \max_{\lambda \in \Lambda(A)} |R(\lambda) - p_{k-1}(\lambda)|.$$

However, as for linear systems, any nonincreasing convergence curve can be obtained with a matrix having any given eigenvalues.

Error Bounds Using the Crouzeix-Palencia Result

If A is not diagonalizable or if $\kappa(V)$ is huge, the following error bound based on [Crouzeix and Palencia, *The Numerical Range is a* $(1+\sqrt{2})$ *Spectral Set*] may prove more useful.

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2}(1+\sqrt{2}) \min_{p_{k-1}\in\mathcal{P}_{k-1}} \max_{z\in W(A)} |R(z)-p_{k-1}(z)|,$$

where $W(A) := \{\langle Aq, q\rangle : \langle q, q\rangle = 1\}$ is the numerical range of A.

Error Bounds Using the Crouzeix-Palencia Result

If A is not diagonalizable or if $\kappa(V)$ is huge, the following error bound based on [Crouzeix and Palencia, *The Numerical Range is a $(1+\sqrt{2})$ Spectral Set*] may prove more useful.

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \le \kappa(S)^{1/2}(1+\sqrt{2}) \min_{p_{k-1}\in\mathcal{P}_{k-1}} \max_{z\in W(A)} |R(z)-p_{k-1}(z)|,$$

where $W(A) := \{\langle Aq, q\rangle : \langle q, q\rangle = 1\}$ is the numerical range of A.

If one prefers to work directly with the S-norm, then

$$\frac{\|R(A)b - Q_k y\|_S}{\|b\|_S} \le (1+\sqrt{2}) \min_{p_{k-1}\in\mathcal{P}_{k-1}} \max_{z\in W(S^{1/2}AS^{-1/2})} |R(z)-p_{k-1}(z)|.$$

Error Bounds Using Results from Crouzeix and G.

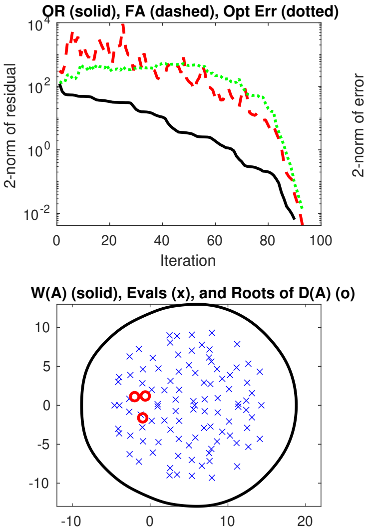
The previous bounds are not useful if R has a pole in $W(A)$ (or in $W(S^{1/2}AS^{-1/2})$). Crouzeix and G. [*Spectral Sets: Numerical Range and Beyond*] showed that one could remove a disk about such a pole $\xi \in W(A)$ of radius $1/w((\xi I - A)^{-1})$, where w is the numerical radius, and still have a $(3+2\sqrt{3})$-spectral set. Thus

$$\frac{\|R(A)b - Q_k y\|_2}{\|b\|_2} \leq \kappa(S)^{1/2}(3+2\sqrt{3}) \cdot$$

$$\min_{p_{k-1} \in \mathcal{P}_{k-1}} \max_{z \in W(A) \setminus \mathcal{D}(\xi, 1/w((\xi I - A)^{-1}))} |R(z) - p_{k-1}(z)|.$$

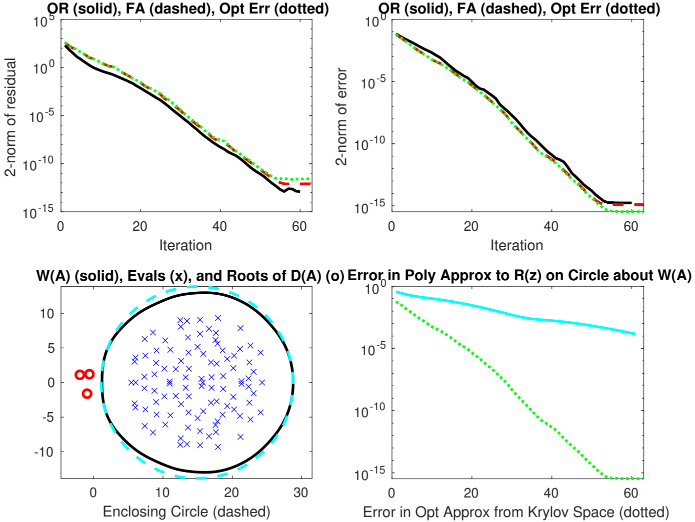
Example: Random A, Random Quadratic N, Random Cubic D

$A = \text{randn}(100) + 5\text{I}$, $\kappa(V) = 98$, $\kappa(D(A)) = 228$.



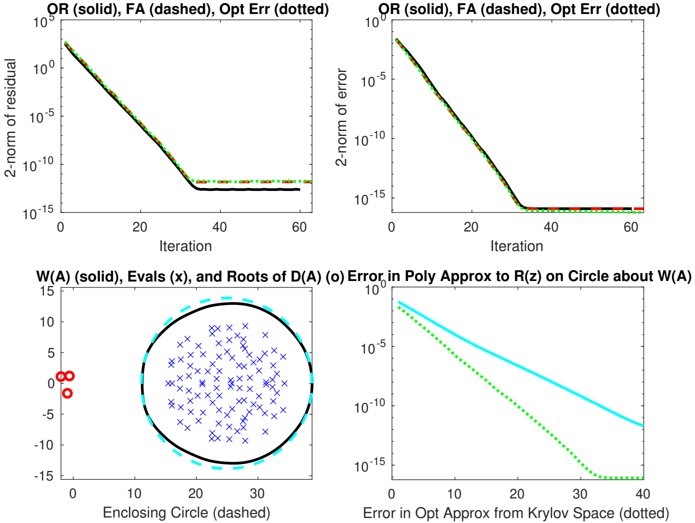
Example: Random A, Random Quadratic N, Random Cubic D

$A = \text{randn}(100) + 15\text{I}$



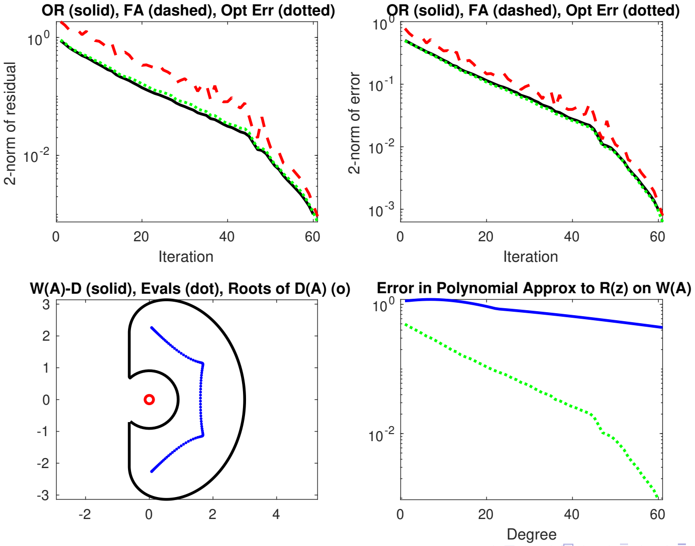
Example: Random A, Random Quadratic N, Random Cubic D

$A = \text{randn}(100) + 25\text{I}$



Example: Grcar Matrix, $N = I$, $D = A$

$A =$ gallery('grcar',100), $\kappa(V) =$ huge, $\kappa(D(A)) = 3.6$.



Summary

- Using $\max\{\deg(D(A)), \deg(N(A))\} - 1$ extra steps, one can find the optimal (in $D(A)^*D(A)$-norm) approx. to $D(A)^{-1}N(A)b$ from the Krylov space.

Summary

- Using $\max\{\deg(D(A)), \deg(N(A))\} - 1$ extra steps, one can find the optimal (in $D(A)^*D(A)$-norm) approx. to $D(A)^{-1}N(A)b$ from the Krylov space.
- Solve a $k + \deg(D(A))$ by k least squares problem to determine this optimal approx. Solve via QR factorization with Givens rotations; apply previous rotations to last column and choose new rotations to eliminate entries below diagonal in column k. The residual norm in the least squares problem is $\|R(A)b - Q_k y\|_{D(A)^*D(A)}$, so need not form $Q_k y$ until tolerance is met.

Summary

- Using $\max\{\deg(D(A)), \deg(N(A))\} - 1$ extra steps, one can find the optimal (in $D(A)^*D(A)$-norm) approx. to $D(A)^{-1}N(A)b$ from the Krylov space.
- Solve a $k + \deg(D(A))$ by k least squares problem to determine this optimal approx. Solve via QR factorization with Givens rotations; apply previous rotations to last column and choose new rotations to eliminate entries below diagonal in column k. The residual norm in the least squares problem is $\|R(A)b - Q_k y\|_{D(A)^*D(A)}$, so need not form $Q_k y$ until tolerance is met.
- For highly nonnormal problems, *a priori* error bounds can be based on how well one can approximate $R(z)$ using a polynomial of degree at most $k-1$ on $W(A)$ or other K-spectral sets.