Probabilistic upper bounds for the condition number of a matrix

Sparse Days, 6 June 2014

Sarah W. Gaaf
Joint work with Michiel E. Hochstenbach

1 Matrix Condition number

Consider $A\mathbf{x} = \mathbf{b}$, $A \in \mathbb{C}^{n \times n}$, A nonsingular.

Problem:

- A, $\mathbf{b}$ perturbed
- Compute $\mathbf{x} = A^{-1}\mathbf{b}$

The sensitivity of linear system:

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|} \right)$$

How to compute the condition number?

$$\kappa(A) = \|A\| \, \|A^{-1}\| = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)} = \sqrt{\frac{\lambda_{\max}(A^*A)}{\lambda_{\min}(A^*A)}}$$

Singular Value Decomposition

- $\approx 21n^3$ flops

Approximate the condition number

- Approximate $\sigma_{\min}(A)$ and $\sigma_{\max}(A)$

Start with a special case:

- Assume A is Hermitian.
- Eigenvalues of A: $\quad |\lambda_1| \geq \cdots \geq |\lambda_n|$

Condition number of A:

$$\kappa(A) = \|A\| \, \|A^{-1}\| = \frac{|\lambda_1|}{|\lambda_n|}.$$

2 Lanczos Method

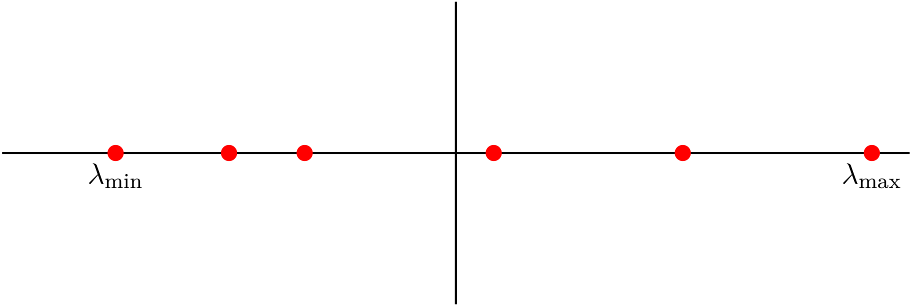
Procedure: $\mathbf{v}_0 \xrightarrow{A} \mathbf{v}_1 \xrightarrow{A} \mathbf{v}_2 \xrightarrow{A} \dots$

- $AV_k = V_{k+1}T_{k+1,k}$
- $\mathbf{v}_k = p_k(A)\mathbf{v}_0$
- $T_k = V_k^* A V_k$ tridiagonal
- Basis for Krylov subspace $\mathcal{K}_k(A, \mathbf{v}_0)$.

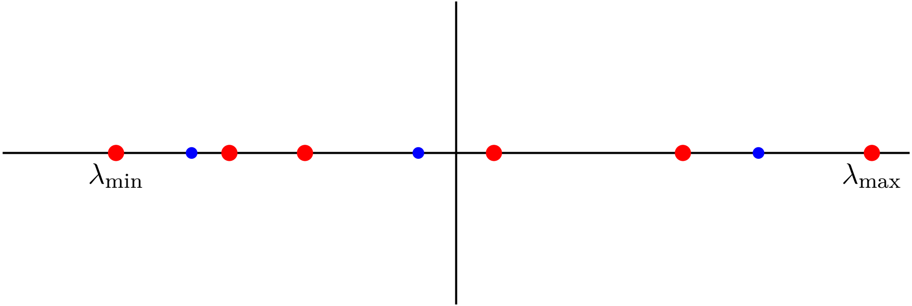
Recall: $\mathcal{K}_k(A, \mathbf{v}_0) = \text{span}\{\mathbf{v}_0, A\mathbf{v}_0, A^2\mathbf{v}_0, \dots, A^{k-1}\mathbf{v}_0\}$.

Procedure: $\mathbf{v}_0 \xrightarrow{A} \mathbf{v}_1 \xrightarrow{A} \mathbf{v}_2 \xrightarrow{A} \dots$

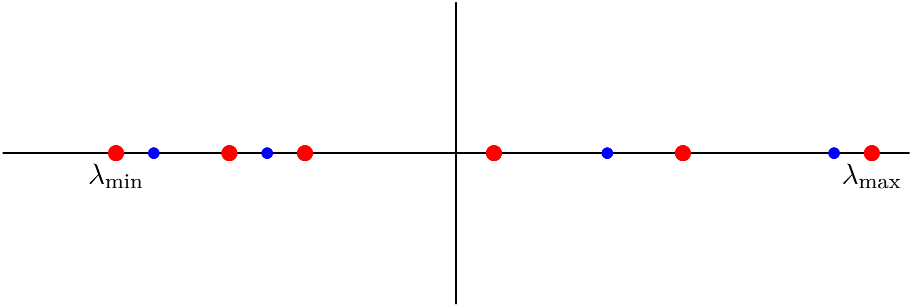
$$A \begin{bmatrix} \mathbf{v}_0 & \mathbf{v}_1 & \dots \end{bmatrix} = \begin{bmatrix} \mathbf{v}_0 & \mathbf{v}_1 & \mathbf{v}_2 & \dots \end{bmatrix} \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \ddots & \\ & & & & \end{bmatrix}$$



● eigenvalues of A



- eigenvalues of A
- eigenvalues of T_k



- eigenvalues of A
- eigenvalues of T_k

van Dorsselaer, Hochstenbach, van der Vorst, 2000:

Probabilistic upper bound for $\lambda_{\min}$ and $\lambda_{\max}$

3 Probabilistic upper bound

Let $\mathbf{v}_0 = \sum_{i=1}^{n} \gamma_i \mathbf{x}_i$,

then

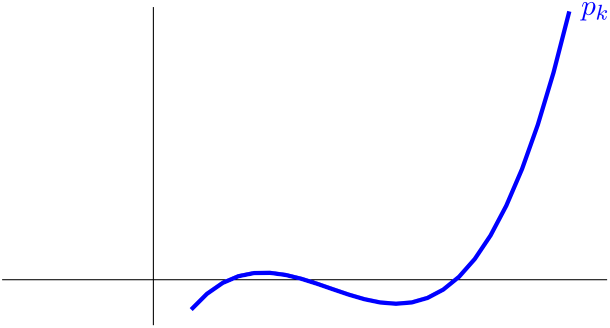
$$1 = \|\mathbf{v}_k\|^2 = \|p_k(A)\mathbf{v}_0\|^2 = \sum_{i=1}^{n} \gamma_i^2 p_k(\lambda_i)^2.$$

Thus

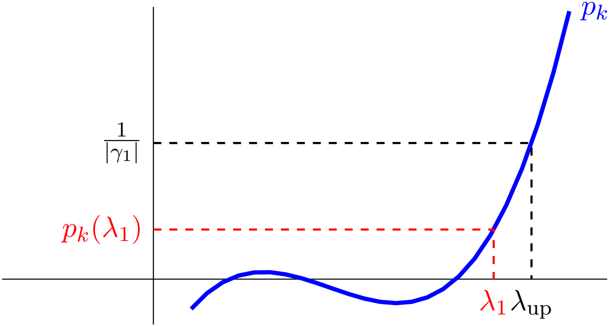
$$1 \geq \gamma_1^2 \, p_k(\lambda_1)^2,$$

and

$$\frac{1}{|\gamma_1|} \geq |p_k(\lambda_1)|.$$

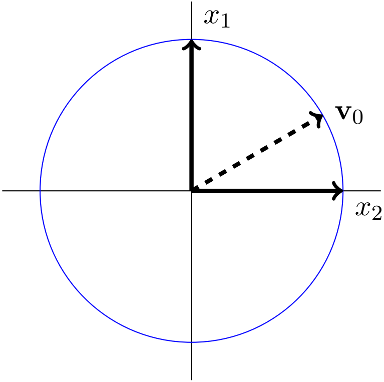


Recall: $\frac{1}{|\gamma_1|} \geq |p_k(\lambda_1)|$

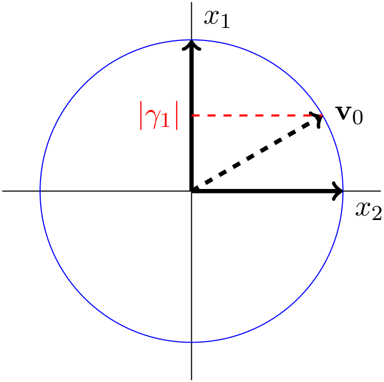


Recall: $\frac{1}{|\gamma_1|} \geq |p_k(\lambda_1)|$

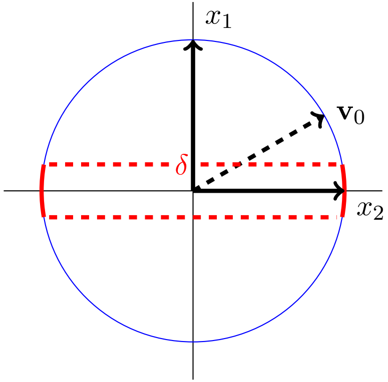
$\mathbb{P}(\frac{1}{\delta} < \frac{1}{|\gamma_1|}) = \mathbb{P}(|\gamma_1| < \delta) = \epsilon$ (ϵ is user-chosen)

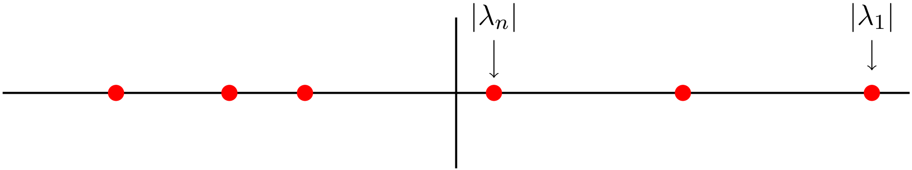


$\mathbb{P}(\frac{1}{\delta} < \frac{1}{|\gamma_1|}) = \mathbb{P}(|\gamma_1| < \delta) = \epsilon$ (ϵ is user-chosen)

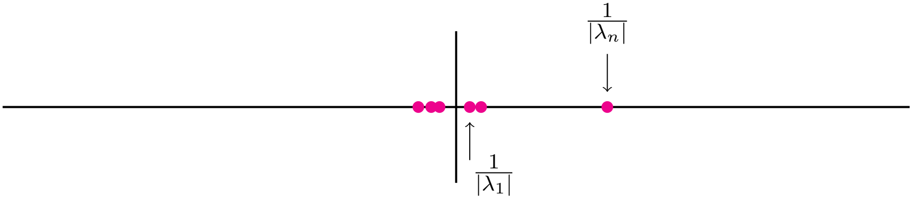


$\mathbb{P}(\frac{1}{\delta} < \frac{1}{|\gamma_1|}) = \mathbb{P}(|\gamma_1| < \delta) = \epsilon$ (ϵ is user-chosen)





● eigenvalues of A



● eigenvalues of A^{-1}

4 Extended Lanczos Method

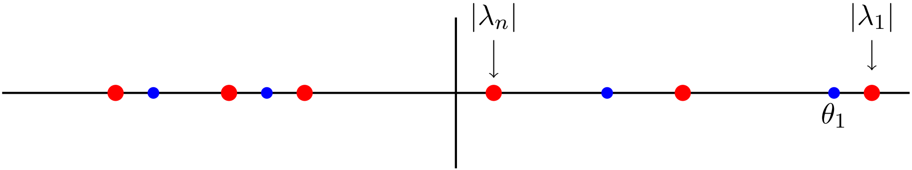
Procedure: $\mathbf{v}_0 \xrightarrow{A} \mathbf{v}_1 \xrightarrow{A^{-1}} \mathbf{v}_2 \xrightarrow{A} \mathbf{v}_3 \xrightarrow{A^{-1}} \mathbf{v}_4 \xrightarrow{A} \ldots$

- $AV_{k-1} = V_k H_{k,k-1}$ and $A^{-1}V_k = V_{k+1}G_{k+1,k}$
- $\mathbf{v}_k = p_k(A)\mathbf{v}_0$ (p_k Laurent)
- H_{k-1}, G_k pentadiagonal
- Basis for ***extended*** Krylov subspace $\mathcal{K}_{m,m}(A, \mathbf{v}_0)$.

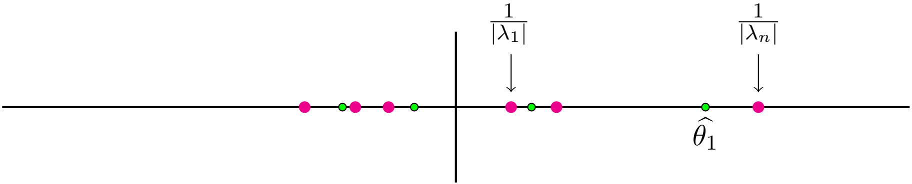
Note: $\mathcal{K}_{m,m}(A, \mathbf{v}_0) = \text{span}\{A^{-m+1}\mathbf{v}_0, \ldots, A^{-1}\mathbf{v}_0, \mathbf{v}_0, A\mathbf{v}_0, \ldots, A^{m-1}\mathbf{v}_0\}$.

Procedure: $\mathbf{v}_0 \xrightarrow{A} \mathbf{v}_1 \xrightarrow{A^{-1}} \mathbf{v}_2 \xrightarrow{A} \mathbf{v}_3 \xrightarrow{A^{-1}} \mathbf{v}_4 \xrightarrow{A} \dots$

$$A \begin{bmatrix} \mathbf{v}_0 & \mathbf{v}_1 & \dots \end{bmatrix} = \begin{bmatrix} \mathbf{v}_0 & \mathbf{v}_1 & \mathbf{v}_2 & \dots \end{bmatrix} \begin{bmatrix} \times & \times & & & & \\ \times & \times & \times & \times & & \\ & \times & \times & \times & & \\ & \times & \times & \times & \times & \times \\ & & & \times & \times & \times \\ & & & \times & \times & \ddots \end{bmatrix}$$



- eigenvalues of A
- approximations to eigenvalues of A



- eigenvalues of A^{-1}
- approximations to eigenvalues of A^{-1}

Bounds:

$$\theta_1 \cdot \widehat{\theta_1} \leq \kappa(A) \qquad \text{(guaranteed lower bound)}$$

$$\kappa(A) \leq |\lambda_1|^{\mathrm{up}} \cdot \left(\frac{1}{|\lambda_n|}\right)^{\mathrm{up}} \qquad \text{(probabilistic upper bound)}$$

5 Bidiagonalization

For the general case:

- A non-Hermitian, $\sigma_i(A)^2 = \lambda_i(A^*A)$

Thus $\kappa(A) = \|A\| \, \|A^{-1}\| = \sqrt{\dfrac{|\lambda_1(A^*A)|}{|\lambda_n(A^*A)|}}$

Lanczos Bidiagonalization, procedure:

$$\mathbf{v}_0 \xrightarrow{A} \mathbf{u}_0 \xrightarrow{A^*} \mathbf{v}_1 \xrightarrow{A} \mathbf{u}_1 \xrightarrow{A^*} \dots$$

Hochstenbach, 2013: Probabilistic upper bound for $\|A\|$

***Extended* Lanczos Bidiagonalization, procedure:**

$$\mathbf{v}_0 \xrightarrow{A} \mathbf{u}_0 \xrightarrow{A^*} \mathbf{v}_1 \xrightarrow{(A^*)^{-1}} \mathbf{u}_1 \xrightarrow{A^{-1}} \dots$$

$$\begin{aligned} A^*AV &= VH^*H \\ (A^*A)^{-1}V &= VKK^* \\ AA^*U &= UHH^* \\ (AA^*)^{-1}U &= UK^*K \end{aligned}$$

$$\mathbf{v}_k = p_k(A^*A)\mathbf{v}_0$$
$$\mathbf{u}_k = q_k(AA^*)\mathbf{u}_0$$

Extended Krylov subspace:

- $\mathcal{K}_{m,m}(A^*A, \mathbf{v}_0) = \text{span}\{\dots, (A^*A)^{-1}\mathbf{v}_0, \mathbf{v}_0, A^*A\mathbf{v}_0, \dots\}$.

6 Conclusions and results

Method		Bounds for
Lanczos	$\rightarrow$	$\lvert\lambda_1\rvert$
Extended Lanczos	$\rightarrow$	$\lvert\lambda_1\rvert$ and $\lvert\lambda_n\rvert$
Lanczos Bidiagonalization	$\rightarrow$	σ_1
Extended Lanczos Bidiagonalization	$\rightarrow$	σ_1 and σ_n

Matrix A	κ	**Ext. Lan. BD** κ_{low}	κ_{up}	**Ratio** κ_{up}/κ_{low}	**CPU** (sec's)
tols2000	$5.99 \cdot 10^6$	$5.97 \cdot 10^6$	$6.22 \cdot 10^6$	1.04	0.15
tols4000	$2.36 \cdot 10^7$	$2.34 \cdot 10^7$	$2.44 \cdot 10^7$	1.04	0.26
grcar10000	$3.63 \cdot 10^0$	$3.61 \cdot 10^0$	$3.87 \cdot 10^0$	1.07	0.37
memplus	$3.81 \cdot 10^6$	$1.29 \cdot 10^5$	$1.32 \cdot 10^5$	1.02	0.65
af23560	$1.99 \cdot 10^4$	$1.99 \cdot 10^4$	$2.00 \cdot 10^4$	1.00	4.58

- For $k = 10$ (i.e. Krylov subspace $\mathcal{K}_{10,10}$)
- We choose $\epsilon = 0.01$
- Matrix af23560, condest(A) time: 2.65 seconds

7 References

Hochstenbach, 2013

Halko, Martinsson, and Tropp, 2011

Jagels and Reichel, 2011

Knizhnerman and Simoncini, 2010

Jagels and Reichel, 2009

Simoncini, 2007

van Dorsselaer, Hochstenbach, and van der Vorst, 2000

Druskin and Knizhnerman, 1998

Kuczyński and Woźniakowski, 1992