Probabilistic upper bounds for the condition number of a matrix

Student Krylov Day 2015

Delft, 02 February 2015

Sarah W. Gaaf
Joint work with Michiel E. Hochstenbach

1 Matrix Condition number

Consider $A\mathbf{x} = \mathbf{b}$, $A \in \mathbb{R}^{n \times n}$, 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?

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

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 symmetric.
- 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 Bidiagonalization

For the general case:

- A nonsymmetric, $\qquad \sigma_i(A)^2 = \lambda_i(A^TA)$

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

Lanczos Bidiagonalization, procedure:

$$\mathbf{v}_0 \xrightarrow{A} \mathbf{u}_0 \xrightarrow{A^T} \mathbf{v}_1 \xrightarrow{A} \mathbf{u}_1 \xrightarrow{A^T} \ldots$$

Lanczos Bidiagonalization

- $$\begin{aligned} AV_k &= U_k B_k \\ A^T U_k &= V_k B_k^T + \beta \mathbf{v}_{k+1} \mathbf{e}_k^T \end{aligned}$$
- $\mathbf{v}_k = p_k(A^T A)\mathbf{v}_0$
- $B_k = U_k^T A V_k$ bidiagonal
- Basis for Krylov subspace $\mathcal{K}_k(A^T A, \mathbf{v}_0)$.

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

Extended **Lanczos Bidiagonalization, procedure:**

$$\mathbf{v}_0 \xrightarrow{A} \mathbf{u}_0 \xrightarrow{A^T} \mathbf{v}_1 \xrightarrow{A^{-T}} \mathbf{u}_1 \xrightarrow{A^{-1}} \dots$$

$$\begin{aligned} A^TAV &= VH^TH \\ (A^TA)^{-1}V &= VKK^T \\ AA^TU &= UHH^T \\ (AA^T)^{-1}U &= UK^TK \end{aligned}$$

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

- H and K are tridiagonal, $H \cdot K = I$
- p_k and q_k are Laurent polynomials
- $\mathcal{K}_{m,m}(A^TA, \mathbf{v}_0) = \text{span}\{\dots, (A^TA)^{-1}\mathbf{v}_0, \mathbf{v}_0, A^TA\mathbf{v}_0, \dots\}$.

Extended **Lanczos Bidiagonalization**

- Largest singular value θ_1 of H approximates $\sigma_1(A)$
- Smallest singular value θ_k of H approximates $\sigma_n(A)$
- Lower bound for condition number:

$$\frac{\theta_1}{\theta_k} \leq \kappa(A)$$

3 Probabilistic upper bound

Let $\mathbf{v}_0 = \sum_{i=1}^{n} \gamma_i \mathbf{y}_i, \qquad (\mathbf{y}_i \text{ right singular vectors of } A)$

then

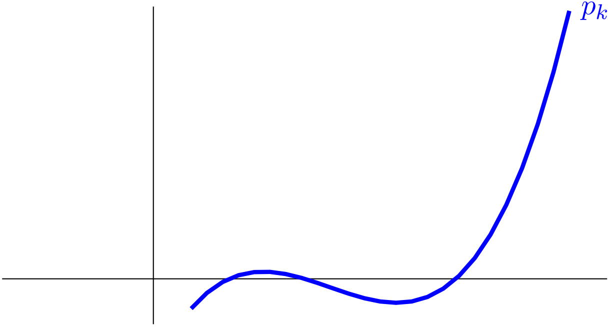
$$1 = \|\mathbf{v}_k\|^2 = \|p_k(A^T A)\mathbf{v}_0\|^2 = \sum_{i=1}^{n} \gamma_i^2 p_k(\sigma_i^2)^2.$$

Thus

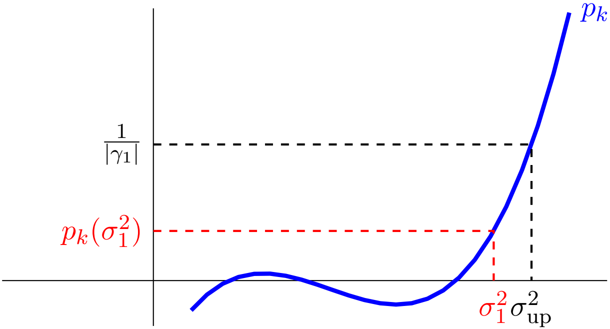
$$1 \geq \gamma_1^2\, p_k(\sigma_1^2)^2,$$

and

$$\frac{1}{|\gamma_1|} \geq |p_k(\sigma_1^2)|.$$



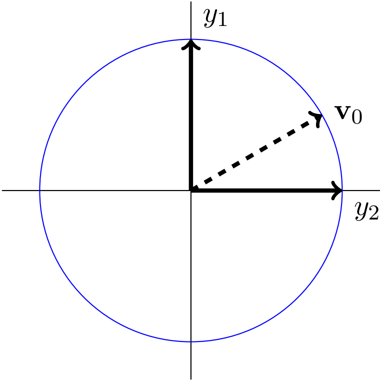
Recall: $\frac{1}{|\gamma_1|} \geq |p_k(\sigma_1^2)|$



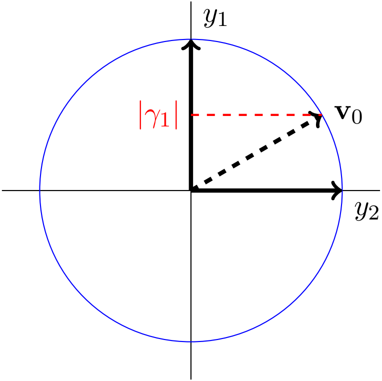
Recall: $\frac{1}{|\gamma_1|} \geq |p_k(\sigma_1^2)|$

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

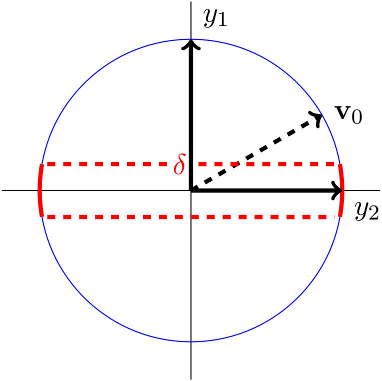
$\mathbb{P}(\frac{1}{\delta} < \frac{1}{|\gamma_1|}) = \mathbb{P}(|\gamma_1| < \delta) = \epsilon$ (ϵ is user-chosen)



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



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



4 Conclusions and results

Bounds for the condition number:

- lower bound: $\kappa_{\text{low}} = \frac{\theta_1}{\theta_k} \leq \kappa(A)$
- probabilistic upper bound: $\kappa(A) \leq \frac{\sigma_1^{\text{prob}}}{\sigma_n^{\text{prob}}} = \kappa_{\text{up}}$

User chosen values:

- ε, probabilistic bound holds with probability at least $1 - 2\varepsilon$
- ζ, method adaptively performs k steps such that

$$\frac{\kappa_{\text{up}}}{\kappa_{\text{low}}} \leq \zeta$$

Matrix A	Dim.	κ	κ_{low}	κ_{up}	k	CPU	LU	CPU^1
`utm5940`	5940	$4.35 \cdot 10^8$	$3.98 \cdot 10^8$	$7.21 \cdot 10^8$	4	0.13	61	0.12
`grcar10000`	10000	$3.63 \cdot 10^0$	$3.59 \cdot 10^0$	$5.80 \cdot 10^0$	6	0.07	31	0.05
`af23560`	23560	$1.99 \cdot 10^4$	$1.93 \cdot 10^4$	$2.82 \cdot 10^4$	6	0.98	74	0.88
`rajat16`	96294	$\star$	$5.63 \cdot 10^{12}$	$5.69 \cdot 10^{12}$	5	9.34	97	9.19
`torso1`	116158	$\star$	$1.41 \cdot 10^{10}$	$1.42 \cdot 10^{10}$	3	26.8	93	28.5
`dc1`	116835	$\star$	$2.39 \cdot 10^8$	$4.59 \cdot 10^8$	5	6.05	93	5.57
`xenon2`	157464	$\star$	$4.29 \cdot 10^4$	$8.14 \cdot 10^4$	7	20.1	82	19.6
`scircuit`	170998	$\star$	$2.40 \cdot 10^9$	$4.69 \cdot 10^9$	7	2.05	54	1.39
`transient`	178866	$\star$	$1.02 \cdot 10^{11}$	$2.00 \cdot 10^{11}$	8	7.70	86	7.12
`stomach`	213360	$\star$	$4.62 \cdot 10^1$	$9.02 \cdot 10^1$	6	13.8	80	13.7

- For $\zeta = 2$ (i.e. $\kappa_{\text{up}}/\kappa_{\text{low}} \leq 2$)
- $\varepsilon = 0.01$ (i.e. upper bound holds with probability at least 98%)
- CPU^1 indicates time of `condest`

Matrix A	Dim.	κ	κ_{low}	κ_{up}	k	CPU	LU
utm5940	5940	$4.35 \cdot 10^8$	$4.35 \cdot 10^8$	$4.71 \cdot 10^8$	10	0.19	42
grcar10000	10000	$3.63 \cdot 10^0$	$3.62 \cdot 10^0$	$3.97 \cdot 10^0$	13	0.13	21
af23560	23560	$1.99 \cdot 10^4$	$1.99 \cdot 10^4$	$2.12 \cdot 10^4$	9	1.14	66
rajat16	96294	$\star$	$5.63 \cdot 10^{12}$	$5.69 \cdot 10^{12}$	5	9.34	97
torso1	116158	$\star$	$1.41 \cdot 10^{10}$	$1.42 \cdot 10^{10}$	3	26.8	93
dc1	116835	$\star$	$2.39 \cdot 10^8$	$2.45 \cdot 10^8$	8	6.52	91
xenon2	157464	$\star$	$4.32 \cdot 10^4$	$4.67 \cdot 10^4$	14	23.6	70
scircuit	170998	$\star$	$2.45 \cdot 10^9$	$2.67 \cdot 10^9$	16	3.28	33
transient	178866	$\star$	$1.03 \cdot 10^{11}$	$1.11 \cdot 10^{11}$	21	9.47	70
stomach	213360	$\star$	$4.82 \cdot 10^1$	$5.24 \cdot 10^1$	14	17.5	63

- For $\zeta = 1.1$ (i.e. $\kappa_{\text{up}}/\kappa_{\text{low}} \leq 1.1$)
- $\varepsilon = 0.01$ (i.e. upper bound holds with probability at least 98%)

5 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