Bounds for the Matrix Condition Number

CIME-EMS School, Exploiting Hidden Structure in Matrix Computations. Algorithms and Applications

Cetraro 22-26 June 2015

Sarah W. Gaaf
Joint work with Michiel Hochstenbach (TU/e)

Uses of $\kappa(A)$

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

Problem:

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

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?

Matrix 2-norm:

$$\|A\|_2 \;=\; \max_{x\neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \;=\; \sigma_1(A) \;=\; \sqrt{\lambda_1(A^TA)} \;=\; \lambda_1\left(\begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix}\right)$$

Matrix condition number:

$$\kappa(A) \;=\; \|A\|_2\,\|A^{-1}\|_2 \;=\; \frac{\sigma_1(A)}{\sigma_n(A)} \;=\; \sqrt{\frac{\lambda_1(A^TA)}{\lambda_n(A^TA)}}$$

Bounds for largest singular value

Lower bound θ:

- Krylov subspace methods (Lanczos bidiagonalization)

Upper bounds:

- Bauer-Fike theorem (1960): *There is a singular value in* $[\theta, \theta + C]$.
  - not guaranteed σ_1
- Probabilistic bound [Hochstenbach 2013]

Problem:

- Find upper and lower bound for smallest singular value

Extended Lanczos Bidiagonalization

Procedure:

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

Equations:

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

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

Characteristics:

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

Proposition: The matrix H is tridiagonal and of the form

$$\begin{bmatrix} \alpha_0 & \beta_0 & & & & & \\ & \alpha_1 & & & & & \\ & \beta_{-1} & \alpha_{-1} & \beta_1 & & & \\ & & & \alpha_2 & & & \\ & & & \beta_{-2} & \alpha_{-2} & \beta_2 & \\ & & & & & \alpha_3 & \\ & & & & & & \ddots \end{bmatrix}, \qquad (1)$$

where its entries satisfy

$$\begin{aligned} h_{2j,2j} &= \alpha_j &= \|A^{-T}\mathbf{v}_j\|^{-1} = \|A^T\mathbf{u}_{-j}\|, \\ h_{2j+1,2j} &= \beta_{-j} &= \mathbf{u}_j^T A\mathbf{v}_j, \\ h_{2j+1,2j+1} &= \alpha_{-j} &= \mathbf{u}_j^T A\mathbf{v}_{-j}, \\ h_{2j+1,2j+2} &= \beta_j &= \|A^T\mathbf{u}_j - (\mathbf{u}_j^T A\mathbf{v}_j)\mathbf{v}_j - (\mathbf{u}_j^T A\mathbf{v}_{-j})\mathbf{v}_{-j}\|. \end{aligned}$$

Lower bound for $\kappa(A)$

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:

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

Probabilistic upper bound

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

then

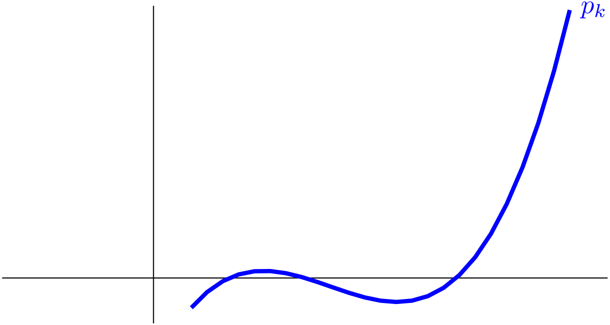
$$1 = \|\mathbf{v}_k\|^2 = \|p_k(A^T A)\mathbf{v}_0\|^2 = \sum_{i=1}^{n} \gamma_i^2 p_k(\sigma_i^2)^2.$$

Thus

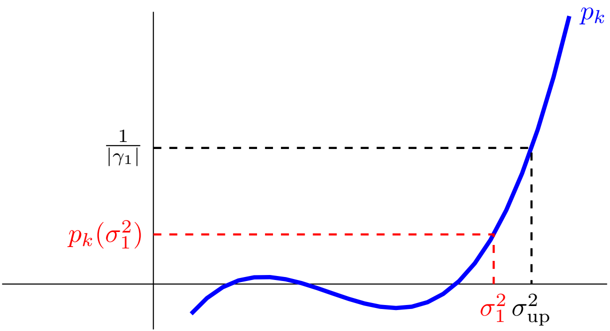
$$1 \geq \gamma_1^2 \, p_k(\sigma_1^2)^2,$$

and

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



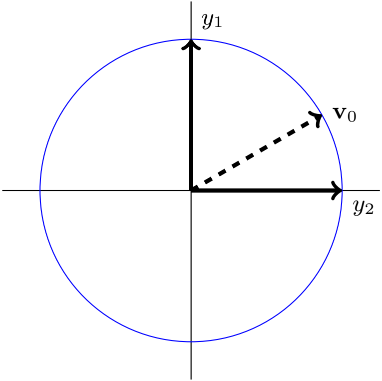
Recall: $\frac{1}{|\gamma_1|} \geq |p_k(\sigma_1^2)|$



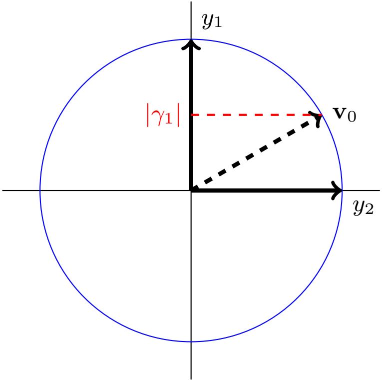
Recall: $\frac{1}{|\gamma_1|} \geq |p_k(\sigma_1^2)|$

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

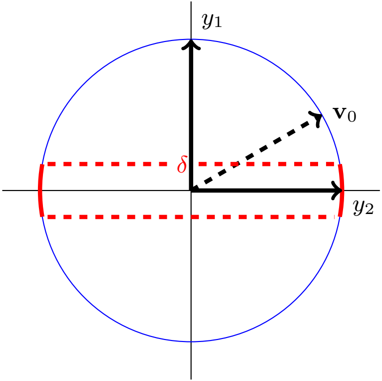
$\mathbb{P}(\frac{1}{\delta} < \frac{1}{|\gamma_1|}) = \mathbb{P}(|\gamma_1| < \delta) = \epsilon$ (ϵ is user-chosen)



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



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



Theorem:

- starting vector $\mathbf{v}_0$ chosen randomly (uniform distribution over S^{n-1})
- $\varepsilon \in (0,1)$ user chosen
- δ be given by $\varepsilon = \mathcal{P}(|\gamma_1| \leq \delta) = \dfrac{B_{\text{inc}}(\frac{n-1}{2}, \frac{1}{2}, \delta^2)}{B_{\text{inc}}(\frac{n-1}{2}, \frac{1}{2}, 1)}$,

where $B_{\text{inc}}(x,y,z) = \int_0^z t^{x-1}(1-t)^{y-1}dt$ (incomplete Beta function)

Then $\sigma_{\text{up}}^{\text{prob}}$, the square root of the largest zero of the polynomial

$$f_1^{\delta}(t) = |p_k(t)| - \tfrac{1}{\delta},$$

is upper bound for σ_1 with probability at least $1-\varepsilon$.

Conclusions

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

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

[SG, M. E. Hochstenbach, *Probabilistic bounds for the matrix condition number with extended Lanczos bidiagonalization* (Submitted)]

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%)

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