ERROR AND SENSITIVTY ANALYSIS FOR SYSTEMS OF LINEAR EQUATIONS

- **Conditioning of linear systems.**
- **Estimating errors for solutions of linear systems**
- **(Normwise) Backward error analysis**
- **Estimating condition numbers ..**

Perturbation analysis for linear systems ($Ax = b$)

Question addressed by perturbation analysis: determine the variation of the solution x when the data, namely A and b, undergoes small variations. Problem is Ill-conditioned if small variations in data cause very large variation in the solution.

Rigorous norm-based error bounds

➤ We perturb $\boldsymbol{A}$ into $\boldsymbol{A}+\boldsymbol{E}$ and $\boldsymbol{b}$ into $\boldsymbol{b}+\boldsymbol{e_b}$. Can we bound the perturbation to the solution?

Preparation: We begin with a lemma for a simple case:

LEMMA: If $\|\boldsymbol{E}\| < 1$ then $\boldsymbol{I}-\boldsymbol{E}$ is nonsingular and

$$\|(\boldsymbol{I}-\boldsymbol{E})^{-1}\| \le \frac{1}{1-\|\boldsymbol{E}\|}$$

Proof is based on following 5 steps

a) Show: If $\|\boldsymbol{E}\| < 1$ then $\boldsymbol{I}-\boldsymbol{E}$ is nonsingular

b) Show: $(\boldsymbol{I}-\boldsymbol{E})(\boldsymbol{I}+\boldsymbol{E}+\boldsymbol{E}^2+\cdots+\boldsymbol{E}^k) = \boldsymbol{I}-\boldsymbol{E}^{k+1}$.

c) From which we get:

$$(\boldsymbol{I} - \boldsymbol{E})^{-1} = \sum_{i=0}^{k} \boldsymbol{E}^{i} + (\boldsymbol{I} - \boldsymbol{E})^{-1} \boldsymbol{E}^{k+1} \rightarrow$$

d) $(\boldsymbol{I} - \boldsymbol{E})^{-1} = \lim_{k \to \infty} \sum_{i=0}^{k} \boldsymbol{E}^{i}$. We write this as

$$(\boldsymbol{I} - \boldsymbol{E})^{-1} = \sum_{i=0}^{\infty} \boldsymbol{E}^{i}$$

e) Finally:

$$\|(I-E)^{-1}\| = \left\|\lim_{k\to\infty}\sum_{i=0}^{k} E^i\right\| = \lim_{k\to\infty}\left\|\sum_{i=0}^{k} E^i\right\|$$

$$\leq \lim_{k\to\infty}\sum_{i=0}^{k}\|E^i\| \leq \lim_{k\to\infty}\sum_{i=0}^{k}\|E\|^i$$

$$\leq \frac{1}{1-\|E\|}$$

➤ Can generalize result:

LEMMA: If A is nonsingular and $\|A^{-1}\| \, \|E\| < 1$ then $A+E$ is non-singular and

$$\|(A+E)^{-1}\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\| \ \|E\|}$$

➤ Proof is based on relation $A + E = A(I + A^{-1}E)$ and use of previous lemma.

➤ Now we can prove the main theorem:

THEOREM 1: Assume that $(A + E)y = b + e_b$ and $Ax = b$ and that $\|A^{-1}\|\|E\| < 1$. Then $A + E$ is nonsingular and

$$\frac{\|x-y\|}{\|x\|} \le \frac{\|A^{-1}\| \ \|A\|}{1-\|A^{-1}\| \ \|E\|}\left(\frac{\|E\|}{\|A\|}+\frac{\|e_b\|}{\|b\|}\right)$$

Proof: From $(A + E)y = b + e_b$ and $Ax = b$ we get $(A + E)(y - x) = e_b - Ex$. Hence:

$$y - x = (A + E)^{-1}(e_b - Ex)$$

Taking norms $\rightarrow$ $\|y - x\| \le \|(A + E)^{-1}\| \left[\|e_b\| + \|E\|\|x\|\right]$
Dividing by $\|x\|$ and using result of lemma

$$\begin{aligned}
\frac{\|y - x\|}{\|x\|} &\le \|(A + E)^{-1}\| \left[\|e_b\|/\|x\| + \|E\|\right] \\
&\le \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|E\|} \left[\|e_b\|/\|x\| + \|E\|\right] \\
&\le \frac{\|A^{-1}\|\|A\|}{1 - \|A^{-1}\|\|E\|} \left[\frac{\|e_b\|}{\|A\|\|x\|} + \frac{\|E\|}{\|A\|}\right]
\end{aligned}$$

Result follows by using inequality $\|A\|\|x\| \ge \|b\|$.... QED

The quantity $\boxed{\kappa(A) = \|A\| \, \|A^{-1}\|}$ is called the condition number of the linear system with respect to the norm $\|.\|$. When using the p-norms we write:

$$\kappa_p(A) = \|A\|_p \|A^{-1}\|_p$$

➤ Note: $\kappa_2(A) = \sigma_{max}(A)/\sigma_{min}(A)$ = ratio of largest to smallest singular values of A. Allows to define $\kappa_2(A)$ when A is not square.

➤ Determinant *is not* a good indication of sensitivity

➤ Small eigenvalues *do not* always give a good indication of poor conditioning.

Example: Consider, for a large α, the $n \times n$ matrix

$$A = I + \alpha e_1 e_n^T$$

➤ Inverse of A is : $A^{-1} = I - \alpha e_1 e_n^T$ ➤ For the ∞-norm we have

$$\|A\|_\infty = \|A^{-1}\|_\infty = 1 + |\alpha|$$

so that $$\kappa_\infty(A) = (1 + |\alpha|)^2.$$

➤ Can give a very large condition number for a large α – but all the eigenvalues of A are equal to one.

Simplification when $e_b = 0$:

$$\frac{\|x-y\|}{\|x\|} \le \frac{\|A^{-1}\| \ \|E\|}{1 - \|A^{-1}\| \ \|E\|}$$

Simplification when $E = 0$:

$$\frac{\|x-y\|}{\|x\|} \le \|A^{-1}\| \ \|A\| \frac{\|e_b\|}{\|b\|}$$

- Slightly less general form: Assume that $\|E\|/\|A\| \le \delta$ and $\|e_b\|/\|b\| \le \delta$ and $\delta\kappa(A) < 1$ then

$$\frac{\|x-y\|}{\|x\|} \le \frac{2\delta\kappa(A)}{1-\delta\kappa(A)}$$

✍ Show the above result

Another common form:

THEOREM 2: Let $(A + \Delta A)y = b + \Delta b$ and $Ax = b$ where $\|\Delta A\| \le \epsilon\|E\|$, $\|\Delta b\| \le \epsilon\|e_b\|$, and assume that $\epsilon\|A^{-1}\|\|E\| < 1$. Then

$$\frac{\|x - y\|}{\|x\|} \le \frac{\epsilon \, \|A^{-1}\| \, \|A\|}{1 - \epsilon\|A^{-1}\| \, \|E\|} \left(\frac{\|e_b\|}{\|b\|} + \frac{\|E\|}{\|A\|} \right)$$

- Results to be seen later are of this type.

Normwise backward error

- We solve $Ax = b$ and find an approximate solution y

Question: Find smallest perturbation to apply to A, b so that *exact* solution of perturbed system is y

Normwise backward error in just A or b

Suppose we model entire perturbation in RHS $\boldsymbol{b}$.

- Let $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{y}$ be the residual.
  Then $\boldsymbol{y}$ satisfies $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b} + \boldsymbol{\Delta b}$ with $\boldsymbol{\Delta b} = -\boldsymbol{r}$ exactly.
- The relative perturbation to the RHS is $\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$.

Suppose we model entire perturbation in matrix $\boldsymbol{A}$.

- Then $\boldsymbol{y}$ satisfies $\left(\boldsymbol{A} + \frac{\boldsymbol{r}\boldsymbol{y}^T}{\boldsymbol{y}^T\boldsymbol{y}}\right)\boldsymbol{y} = \boldsymbol{b}$
- The relative perturbation to the matrix is

$$\left\|\frac{\boldsymbol{r}\boldsymbol{y}^T}{\boldsymbol{y}^T\boldsymbol{y}}\right\|_2 / \|\boldsymbol{A}\|_2 = \frac{\|\boldsymbol{r}\|_2}{\|\boldsymbol{A}\|\,\|\boldsymbol{y}\|_2}$$

TB: 12; AB: 1.2.7; GvL 3.5 – PertA

Normwise backward error in both A & b

For a given y and given perturbation directions E, e_b, we define the Normwise backward error:

$$\eta_{E,e_b}(y) = \min\{\epsilon \mid (A + \Delta A)y = b + \Delta b; \text{ for all } \Delta A, \Delta b \text{ satisfying: } \|\Delta A\| \le \epsilon\|E\|; \text{ and } \|\Delta b\| \le \epsilon\|e_b\|\}$$

In other words $\eta_{E,e_b}(y)$ is the smallest ϵ for which

$$(1)\begin{cases} (A + \Delta A)y = & b + \Delta b; \\ \|\Delta A\| \le \epsilon\|E\|; & \|\Delta b\| \le \epsilon\|e_b\| \end{cases}$$

➤ $\boldsymbol{y}$ is given (a computed solution). $\boldsymbol{E}$ and $\boldsymbol{e_b}$ to be selected (most likely 'directions of perturbation for $\boldsymbol{A}$ and $\boldsymbol{b}$').

➤ Typical choice: $\boldsymbol{E} = \boldsymbol{A}$, $\boldsymbol{e_b} = \boldsymbol{b}$

✍ Explain why this is not unreasonable

Let $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{y}$. Then we have:

THEOREM 3: $\boldsymbol{\eta_{E,e_b}(y) = \frac{\|r\|}{\|E\|\,\|y\| + \|e_b\|}}$

Normwise backward error is for case $\boldsymbol{E = A, e_b = b}$:

$$\boldsymbol{\eta_{A,b}(y) = \frac{\|r\|}{\|A\|\,\|y\| + \|b\|}}$$

✍ Show how this can be used in practice as a means to stop some iterative method which computes a sequence of approximate solutions to $Ax = b$.

✍ Consider the 6×6 Vandermonde system $Ax = b$ where $a_{ij} = j^{2(i-1)}$, $b = A * [1, 1, \cdots, 1]^T$. We perturb A by E, with $|E| \leq 10^{-10}|A|$ and b similarly and solve the system. Evaluate the backward error for this case. Evaluate the forward bound provided by Theorem 2. Comment on the results.

Proof of Theorem 3

Let $D \equiv \|E\|\|y\| + \|e_b\|$ and $\eta \equiv \eta_{E,e_b}(y)$. The theorem states that $\eta = \|r\|/D$. Proof in 2 steps.

First: Any $\Delta A, \Delta b$ pair satisfying (1) is such that $\epsilon \geq \|r\|/D$. Indeed from (1) we have (recall that $r = b - Ay$)

$$Ay + \Delta Ay = b + \Delta b \rightarrow r = \Delta Ay - \Delta b \rightarrow$$

$$\|r\| \leq \|\Delta A\|\|y\| + \|\Delta b\| \leq \epsilon(\|E\|\|y\| + \|e_b\|) \rightarrow \epsilon \geq \frac{\|r\|}{D}$$

Second: We need to show an instance where the minimum value of $\|r\|/D$ is reached. Take the pair $\Delta A, \Delta b$:

$$\Delta A = \alpha r z^T; \quad \Delta b = \beta r \quad \text{with } \alpha = \frac{\|E\|\|y\|}{D}; \quad \beta = \frac{\|e_b\|}{D}$$

The vector z depends on the norm used - for the 2-norm: $z = y/\|y\|^2$. Here: Proof only for 2-norm

a) We need to verify that first part of (1) is satisfied:

$$(A + \Delta A)y = Ay + \alpha r \frac{y^T}{\|y\|^2} y = b - r + \alpha r$$
$$= b - (1 - \alpha)r = b - \left(1 - \frac{\|E\| \, \|y\|}{\|E\| \|y\| + \|e_b\|}\right) r$$
$$= b - \frac{\|e_b\|}{D} r = b + \beta r \quad \rightarrow$$

$(A + \Delta A)y = b + \Delta b \quad \leftarrow$ The desired result

Finally: b) Must now verify that $\|\Delta A\| = \eta\|E\|$ and $\|\Delta b\| = \eta\|e_b\|$. Exercise: Show that $\|uv^T\|_2 = \|u\|_2\|v\|_2$

$$\|\Delta A\| = \frac{|\alpha|}{\|y\|^2}\|ry^T\| = \frac{\|E\|\,\|y\|}{D}\frac{\|r\|\,\|y\|}{\|y\|^2} = \eta\|E\|$$

$$\|\Delta b\| = |\beta|\|r\| = \frac{\|e_b\|}{D}\|r\| = \eta\|e_b\| \qquad QED$$

TB: 12; AB: 1.2.7; GvL 3.5 – PertA

Estimating condition numbers.

- Often we just want to get a lower bound for condition number [it is 'worse than ...']
- We want to estimate $\|A\| \, \|A^{-1}\|$.
- The norm $\|A\|$ is usually easy to compute but $\|A^{-1}\|$ is not.
- We want: Avoid the expense of computing A^{-1} explicitly.

Idea:

- Select a vector v so that $\|v\| = 1$ but $\|Av\| = \tau$ is small.
- Then: $\|A^{-1}\| \geq 1/\tau$ (show why) and:

$$\kappa(A) \geq \frac{\|A\|}{\tau}$$

- Condition number worse than $\|A\|/\tau$.

- Typical choice for v: choose $[\cdots \pm 1 \cdots]$ with signs chosen on the fly during back-substitution to maximize the next entry in the solution, based on the upper triangular factor from Gaussian Elimination.

- Similar techniques used to estimate condition numbers of large matrices in matlab.

Condition numbers and near-singularity

➤ $1/\kappa \approx$ relative distance to nearest singular matrix.

Let A, B be two $n \times n$ matrices with A nonsingular and B singular. Then

$$\frac{1}{\kappa(A)} \le \frac{\|A - B\|}{\|A\|}$$

Proof: B singular $\rightarrow$ $\exists\, x \neq 0$ such that $Bx = 0$.

$$\|x\| = \|A^{-1}Ax\| \le \|A^{-1}\|\,\|Ax\| = \|A^{-1}\|\|(A-B)x\| \\ \le \|A^{-1}\|\,\|A-B\|\|x\|$$

Divide both sides by $\|x\| \times \kappa(A) = \|x\|\|A\|\,\|A^{-1}\|$ ➤ result. QED.

 TB: 12; AB: 1.2.8 ;GvL 3.5; Ort 9.3-4 – PertBshort

Example:

$$\text{let } A = \begin{pmatrix} 1 & 1 \\ 1 & 0.99 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Then $\frac{1}{\kappa_1(A)} \le \frac{0.01}{2}$ ➤ $\kappa_1(A) \ge \frac{2}{0.01} = 200.$

➤ It can be shown that (Kahan)

$$\frac{1}{\kappa(A)} = \min_B \left\{ \frac{\|A - B\|}{\|A\|} \quad | \quad \det(B) = 0 \right\}$$

Estimating errors from residual norms

Let $\tilde{x}$ an approximate solution to system $Ax = b$ (e.g., computed from an iterative process). We can compute the residual norm:

$$\|r\| = \|b - A\tilde{x}\|$$

Question: How to estimate the error $\|x - \tilde{x}\|$ from $\|r\|$?

➤ One option is to use the inequality

$$\frac{\|x-\tilde{x}\|}{\|x\|} \leq \kappa(A) \frac{\|r\|}{\|b\|}.$$

➤ We must have an estimate of $\kappa(A)$.

Proof of inequality.

First, note that $A(x - \tilde{x}) = b - A\tilde{x} = r$. So:

$$\|x - \tilde{x}\| = \|A^{-1}r\| \le \|A^{-1}\| \; \|r\|$$

Also note that from the relation $b = Ax$, we get

$$\|b\| = \|Ax\| \le \|A\| \; \|x\| \quad \rightarrow \quad \|x\| \ge \frac{\|b\|}{\|A\|}$$

Therefore,

$$\frac{\|x - \tilde{x}\|}{\|x\|} \le \frac{\|A^{-1}\| \; \|r\|}{\|b\|/\|A\|} = \kappa(A)\frac{\|r\|}{\|b\|} \qquad \blacksquare$$

✍ Show that

$$\frac{\|x-\tilde{x}\|}{\|x\|} \ge \frac{1}{\kappa(A)} \frac{\|r\|}{\|b\|}.$$