ERROR AND SENSITIVITY ANALYSIS FOR SYSTEMS OF LINEAR EQUATIONS

• Conditioning of linear systems.
• Estimating errors for solutions of linear systems
• Backward error analysis
• Relative element-wise error analysis

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.

Analysis I: Asymptotic First Order Analysis

Let $E$ be an $n \times n$ matrix and $e_b$ be an $n$-vector.

“Perturb” $A$ into $A(\epsilon) = A + \epsilon E$ and $b$ into $b + \epsilon e_b$.

Note: $A + \epsilon E$ is nonsingular for $\epsilon$ small enough. Why?

The solution $x(\epsilon)$ of the perturbed system is s.t.

$$(A + \epsilon E)x(\epsilon) = b + \epsilon e_b.$$

Let $\delta(\epsilon) = x(\epsilon) - x$. Then,

$$(A + \epsilon E)\delta(\epsilon) = (b + \epsilon e_b) - (A + \epsilon E)x = \epsilon (e_b - Ex)$$

$\delta(\epsilon) = \epsilon (A + \epsilon E)^{-1}(e_b - Ex)$.

$x(\epsilon)$ is differentiable at $\epsilon = 0$ and its derivative is

$$x'(0) = \lim_{\epsilon \to 0} \frac{\delta(\epsilon)}{\epsilon} = A^{-1} (e_b - E x).$$

A small variation $[\epsilon E, \epsilon e_b]$ will cause the solution to vary by roughly $\epsilon x'(0) = \epsilon A^{-1}(e_b - Ex)$.

The relative variation is such that

$$\frac{\|x(\epsilon) - x\|}{\|x\|} \leq \epsilon \|A^{-1}\| \left( \frac{\|e_b\|}{\|b\|} + \|E\| \right) + O(\epsilon^2).$$

Since $\|b\| \leq \|A\||x|$:

$$\frac{\|x(\epsilon) - x\|}{\|x\|} \leq \epsilon \|A\||A^{-1}| \left( \frac{\|e_b\|}{\|b\|} + \|E\| \right) + O(\epsilon^2).$$
The quantity $\kappa(A) = \|A\| \|A^{-1}\|$ is called the condition number of the linear system with respect to the norm $\|\cdot\|$. When using the $p$-norms we write:

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

Note: $\kappa_2(A) = \sigma_{\text{max}}(A)/\sigma_{\text{min}}(A) = \text{ratio of largest to smallest singular values of } A$. 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 $\alpha$, 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 $\infty$-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 $\alpha$ – but all the eigenvalues of $A$ are equal to one.

---

**Rigorous norm-based error bounds**

Previous bound is valid only when perturbation is “small enough,” where “small” is not precisely defined.

New bound valid within an explicitly given neighborhood.

**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

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

To prove, first need to show that $A + E$ is nonsingular if $A$ is nonsingular and $E$ is small.

---

Begin with simple case:

**LEMMA:** If $\|E\| < 1$ then $I - E$ is nonsingular and

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

**Proof** is based on following 5 steps

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

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

c) From which we get:

$$(I - E)^{-1} = \sum_{i=0}^{k} E^i + (I - E)^{-1} E^{k+1} \rightarrow$$
d) \((I - E)^{-1} = \lim_{k \to \infty} \sum_{i=0}^{k} E^i\). We write this as
\[
(I - E)^{-1} = \sum_{i=0}^{\infty} E^i
\]
e) Finally:
\[
\| (I - E)^{-1} \| = \left\| \lim_{k \to \infty} \sum_{i=0}^{k} E^i \right\| = \lim_{k \to \infty} \left\| \sum_{i=0}^{k} E^i \right\|
\]
\[
\leq \lim_{k \to \infty} \sum_{i=0}^{k} \| E^i \| \leq \lim_{k \to \infty} \sum_{i=0}^{k} \| E \|^i
\]
\[
\leq \frac{1}{1 - \| E \|}
\]

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 \(\| y - x \| \leq \| (A + E)^{-1} \| \| e_b \| + \| E \| \| x \|\)
Dividing by \(\| x \|\) and using result of lemma
\[
\frac{\| y - x \|}{\| x \|} \leq \| (A + E)^{-1} \| \frac{\| e_b \|}{\| x \|} + \| E \|\]
\[
\leq \frac{\| A^{-1} \|}{1 - \| A^{-1} \| \| E \|} \frac{\| e_b \|}{\| x \|} + \| E \|
\]
\[
\leq \frac{\| A^{-1} \| \| A \|}{1 - \| A^{-1} \| \| E \|} \left[ \frac{\| e_b \|}{\| A \| \| x \|} + \| E \| \right]
\]
Result follows by using inequality \(\| A \| \| x \| \geq \| b \|\).... QED

Can generalize result:

**LEMMMA:** If \(A\) is nonsingular and \(\| A^{-1} \| \| E \| < 1\) then \(A + E\) is non-singular and
\[
\| (A + E)^{-1} \| \leq \frac{\| A^{-1} \|}{1 - \| A^{-1} \| \| E \|}
\]

Proof is based on relation \(A + E = A(I + A^{-1}E)\) and use of previous lemma.
Now we can prove the 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
\[
\frac{\| x - y \|}{\| x \|} \leq \frac{\| A^{-1} \| \| A \|}{1 - \| A^{-1} \| \| E \|} \left( \| E \| + \| e_b \| \right)
\]

Simplification when \(e_b = 0\):
\[
\frac{\| x - y \|}{\| x \|} \leq \frac{\| A^{-1} \| \| E \|}{1 - \| A^{-1} \| \| E \|}
\]

Simplification when \(E = 0\):
\[
\frac{\| x - y \|}{\| x \|} \leq \frac{\| A^{-1} \| \| A \| \| e_b \|}{\| b \|}
\]

Slightly less general form: Assume that \(\| E \| / \| A \| \leq \delta\) and \(\| e_b \| / \| b \| \leq \delta\) and \(\delta \kappa(A) < 1\) then
\[
\frac{\| x - y \|}{\| x \|} \leq \frac{2 \delta \kappa(A)}{1 - \delta \kappa(A)}
\]

Show the above result
Another common form:

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

\[
\frac{\|x - y\|}{\|x\|} \leq \frac{\epsilon \|A^{-1}\| \|A\| \left(\|eb\| + \|E\|\right)}{1 - \epsilon \|A^{-1}\| \|E\| \left(\|b\| + \|A\|\right)}
\]

- Result 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 \(b\).

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

Suppose we model entire perturbation in matrix \(A\).

- Then \(y\) satisfies \(\left(A + \frac{ry^T}{y^Ty}\right)y = b\)
- The relative perturbation to the matrix is
  \[
  \frac{\|ry^T\|}{\|y^Ty\|_2} \leq \frac{\|r\|}{\|A\| \|y\|_2}
  \]

---

**Normwise backward error in both \(A\) & \(b\)**

For a given \(y\) and given perturbation directions \(E, eb\), we define the Normwise backward error:

\[
\eta_{E,eb}(y) = \min\left\{ \epsilon \mid (A + \Delta A)y = b + \Delta b; \right. \\left. \|\Delta A\| \leq \epsilon \|E\|; \ \|\Delta b\| \leq \epsilon \|eb\| \right\}
\]

In other words \(\eta_{E,eb}(y)\) is the smallest \(\epsilon\) for which

\[
(1) \left\{ \begin{array}{l}
(A + \Delta A)y = b + \Delta b; \\
\|\Delta A\| \leq \epsilon \|E\|; \\
\|\Delta b\| \leq \epsilon \|eb\|
\end{array} \right.
\]
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 $6 \times 6$ Vandermonde system $Ax = b$ where $a_{ij} = j^{2(i-1)}$, $b = A \ast [1, 1, \ldots, 1]^T$. 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.

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:

$$(A + \Delta A)y = Ay + \alpha r \frac{y^T}{\|y\|^2}y = b - r + \alpha r$$

$$= b - (1 - \alpha)r = b - \left(1 - \frac{\|E\|\|y\|}{\|E\|\|y\| + \|e_b\|}\right)r$$

$$= b - \frac{\|e_b\|}{D}r = b + \beta r \quad \rightarrow$$

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

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 $\Delta A, \Delta b$:

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

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

Typical choice: $E = A, e_b = b$

Explain why this is not unreasonable.
b) Finally: Must now verify that $\| \Delta A \| = \eta \| E \|$ and $\| \Delta b \| = \eta \| e_b \|$. Exercise: Show that $\| uv^T \|_2 = \| u \|_2 \| v \|_2$

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

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