Sec. 2.5 Sensitivity analysis of linear system

Goal: 1. Condition number

2. sensitivity analysis

1. Condition number.

eg. Consider $\underbrace{\begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}}_{\vec{b}}$.

approximate solutions: $\vec{x}^{(1)} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}$, $\vec{x}^{(2)} = \begin{pmatrix} 0.999 \\ 1.00 \end{pmatrix}$.

Which one is preferred?

$A\vec{x}^{(1)} - \vec{b} = \begin{pmatrix} 0.000001 \\ 0 \end{pmatrix}$, $A\vec{x}^{(2)} - \vec{b} = \begin{pmatrix} 0.000780 \\ 0.000913 \end{pmatrix}$.

If we want small residual, then $\vec{x}^{(1)}$ is preferred.

However, exact solution $\vec{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. $\|\vec{x}^{(1)} - \vec{x}\| \ll \|\vec{x}^{(2)} - \vec{x}\|$, $\vec{x}^{(2)}$ is more accurate.

Some understanding behind the problem:

$\hat{A} = \begin{pmatrix} 0.780 & 0.563\,0001095\ldots \\ 0.913 & 0.659 \end{pmatrix}$ is singular

So A is nearly singular.

An $O(10^{-6})$ perturbation of the data will render the problem $A\vec{x} = \vec{b}$ insolvable.

Def: Condition number of A is $K_p(A) = \|A\|_p \cdot \|A^{-1}\|_p$.

Recall: For $\vec{x} = (x_1, \cdots, x_n)^T$, $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$, $1 \le p < \infty$, $\|x\|_\infty = \max_{1 \le i \le n} |x_i|$

For $A \in \mathbb{C}^{n \times n}$, $\|A\|_p = \sup_{\|\vec{x}\|=1} \|A\vec{x}\|_p$, $1 \le p < \infty$

$\|A\|_1 = \max_{1 \le j \le n} \|A(:,j)\|_1$, $\|A\|_\infty = \max_{1 \le i \le n} \|A(i,:)\|_1$, $\|A\|_2 \le \|A\|_F := \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}}$

eg. $A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \Rightarrow A^{-1} = 10^6 \begin{pmatrix} 0.659 & -0.563 \\ -0.913 & 0.780 \end{pmatrix}$

$\Rightarrow K_1(A) = \|A\|_1 \cdot \|A^{-1}\|_1 \approx 10^6$

Theorem 1:
1. $K(A) \geq 1$. (If A is unitary, $K(A) = 1$)
2. $K(A)$ is large if A is close to singular
3. $K(\alpha A) = K(A)$, (scaling invariant).
4. $K_2(A) = \dfrac{\sigma_{max}(A)}{\sigma_{min}(A)}$, where $\sigma_{max}(A)$ and $\sigma_{min}(A)$ are the maximal and minimal singular values of A, i.e. $\sigma(A) = \sqrt{\lambda(A^* A)}$

Remark:
1. Small eigenvalues or determinant does NOT imply "near singular".
   eg. αI_n, all eigenvalues $= \alpha$, $\det(\alpha I_n) = \alpha^n$.
   If $\alpha = 10^{-10}$. small, but $K(\alpha I_n) = K(I_n) = 1$.
2. If A is normal, i.e. $A^* A = A \cdot A^*$, then $K(A) = \dfrac{|\lambda_{max}(A)|}{|\lambda_{min}(A)|}$
3. All eigenvalues are 1 does NOT imply small condition number.
   eg, $A = \begin{pmatrix} 1 & & \alpha \\ & \ddots & \\ & & 1 \end{pmatrix}$, $A^{-1} = \begin{pmatrix} 1 & & -\alpha \\ & \ddots & \\ & & 1 \end{pmatrix}$, $K_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = (1+|\alpha|)^2$

   If $\alpha = 10^8$, $K_\infty(A) \approx 10^{16}$

2. <u>Sensitivity analysis</u>

linear system: $A\vec{x} = \vec{b}$

perturbed linear system: $\hat{A}\hat{x} = \hat{b}$

$\|\hat{x} - x\|$: absolute error, $\dfrac{\|\hat{x} - \vec{x}\|}{\|\vec{x}\|}$: relative error ✓

(The sensitivity of the linear system has nothing to do with the numerical algorithm used).

<u>Theorem 2</u>: Suppose $A \in \mathbb{R}^{n \times n}$ is nonsingular and $A\vec{x} = \vec{b}$. If eps. $K(A) < 1$, then the stored linear system $\hat{A}\hat{x} = \hat{b}$ is nonsingular, and

$$\frac{\|\hat{x} - \vec{x}\|}{\|\vec{x}\|} \leq \frac{2 \cdot \text{eps} \cdot K(A)}{1 - \text{eps} \cdot K(A)}$$

To prove Theorem 2, we need the following Lemmas.

Lemma 1: If $\hat{A}$ is the stored version of any $A \in \mathbb{R}^{m\times n}$, then $\hat{A} = A + E$, where $E \in \mathbb{R}^{m\times n}$ and $\|E\| \leq eps \cdot \|A\|$. eps: machine precision

proof: Let $\hat{A} = (\hat{a}_{ij})$. Then $\hat{a}_{ij} = a_{ij}(1+\varepsilon_{ij})$, where $|\varepsilon_{ij}| \leq eps$

$$\|E\|_1 = \|\hat{A} - A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |\hat{a}_{ij} - a_{ij}| = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}\cdot\varepsilon_{ij}| \leq eps \cdot \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| = eps \cdot \|A\|_1.$$

The proof for $\|\cdot\|_\infty$ is similar.

Lemma 2: Suppose $A\vec{x} = \vec{b}$ is perturbed to $(A+E)\hat{x} = (\vec{b}+\vec{e})$.

Then $$\frac{\|\hat{x}-\vec{x}\|}{\|\vec{x}\|} \leq \frac{\|A^{-1}\|\cdot\|A\|}{1-\|A^{-1}\|\cdot\|E\|}\left(\frac{\|\vec{e}\|}{\|\vec{b}\|} + \frac{\|E\|}{\|A\|}\right)$$

proof: $$A(\hat{x}-\vec{x}) = A\hat{x} - A\vec{x} = (\vec{b}+\vec{e}-E\hat{x}) - \vec{b} = \vec{e} - E\hat{x} = (\vec{e}+E\vec{x}) - E(\hat{x}-\vec{x})$$

$$\hat{x}-\vec{x} = A^{-1}[(\vec{e}+E\vec{x}) - E(\hat{x}-\vec{x})]$$

By Cauchy inequality and triangle inequality,

$$\|\hat{x}-\vec{x}\| \leq \|A^{-1}\|\cdot\|\vec{e}+E\vec{x} - E(\vec{x}-\hat{x})\| \leq \|A^{-1}\|\cdot(\|\vec{e}\| + \|E\vec{x}\| + \|E(\vec{x}-\hat{x})\|)$$

$$\leq \|A^{-1}\|(\|\vec{e}\| + \|E\|\cdot\|\vec{x}\| + \|E\|\cdot\|\hat{x}-\vec{x}\|)$$

$$\Rightarrow (1-\|A^{-1}\|\cdot\|E\|)(\|\hat{x}-\vec{x}\|) \leq \|A^{-1}\|\cdot(\|\vec{e}\| + \|E\|\cdot\|\vec{x}\|)$$

$$\|\hat{x}-\vec{x}\| \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|E\|}(\|\vec{e}\| + \|E\|\|\vec{x}\|)$$

$$\frac{\|\hat{x}-\vec{x}\|}{\|\vec{x}\|} \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|E\|}\left(\frac{\|\vec{e}\|}{\|\vec{x}\|} + \|E\|\right) = \frac{\|A^{-1}\|\cdot\|A\|}{1-\|A^{-1}\|\cdot\|E\|}\cdot\left(\frac{\|\vec{e}\|}{\|A\|\cdot\|\vec{x}\|} + \frac{\|E\|}{\|A\|}\right)$$

Note that $\|\vec{b}\| = \|A\vec{x}\| \leq \|A\|\cdot\|\vec{x}\| \Rightarrow \frac{1}{\|A\|\cdot\|\vec{x}\|} \leq \frac{1}{\|\vec{b}\|}$.

$$\frac{\|\hat{x}-\vec{x}\|}{\|\vec{x}\|} \leq \frac{\|A^{-1}\|\cdot\|A\|}{1-\|A^{-1}\|\cdot\|E\|}\left(\frac{\|\vec{e}\|}{\|\vec{b}\|} + \frac{\|E\|}{\|A\|}\right)$$

proof of Theorem 2:

① If $\hat{A}$ is singular, then there is a nonzero vector $\vec{z}$ such that $\hat{A}\vec{z} = \vec{0}$.

$$\hat{A} = A+E \Rightarrow (A+E)\vec{z} = 0 \Rightarrow A\vec{z} = -E\vec{z} \Rightarrow \vec{z} = -A^{-1}E\vec{z}$$

$$\|\vec{z}\| \leq \|A^{-1}\|\cdot\|E\|\cdot\|\vec{z}\| \leq \|A^{-1}\|\cdot eps\cdot\|A\|\cdot\|\vec{z}\| = \underbrace{eps\cdot K(A)}_{<1}\cdot\|\vec{z}\| < \|\vec{z}\|.$$ Contradiction!

So $\hat{A}$ is nonsingular.

② By Lemma 1, $\|E\| \leq eps \cdot \|A\|$, $\|\vec{e}\| \leq eps \cdot \|\vec{b}\|$.

Using Lemma 2, we get

$$\frac{\|\hat{x} - \vec{x}\|}{\|\vec{x}\|} \leq \frac{\|A^{-1}\| \cdot \|A\|}{1 - \|A^{-1}\| \cdot \|E\|} \left(\frac{\|\vec{e}\|}{\|\vec{b}\|} + \frac{\|E\|}{\|A\|} \right) \leq \frac{K(A)}{1 - \|A^{-1}\| \cdot \|E\|} (eps + eps) \leq \frac{2 \cdot eps \cdot K(A)}{1 - \|A^{-1}\| \cdot \|E\|}$$

$$\|E\| \leq eps \cdot \|A\| \implies - \|A^{-1}\| \cdot \|E\| \geq - \|A^{-1}\| \cdot eps \cdot \|A\| = - eps \cdot K(A)$$

$$\implies 1 - \|A^{-1}\| \cdot \|E\| \geq 1 - eps \cdot K(A) \implies \frac{1}{1 - \|A^{-1}\| \cdot \|E\|} \leq \frac{1}{1 - eps \cdot K(A)}$$

So $$\frac{\|\hat{x} - \vec{x}\|}{\|\vec{x}\|} \leq \frac{2 \cdot eps \cdot K(A)}{1 - eps \cdot K(A)}.$$