ECEn 671: Mathematics of Signals and Systems

Randal W. Beard

Brigham Young University

September 1, 2023

Section 1

Matrix Condition Number

Matrix Condition Number

- Suppose that $A \in \mathbb{C}^{n \times n}$ is full rank and A^{-1} is to be computed numerically. How reliable is the computation?
- $Ax = b$ can be written as

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

- Therefore, the solution x is the intersection of $n-$ hyperplanes:

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

Matrix Condition Number, cont.

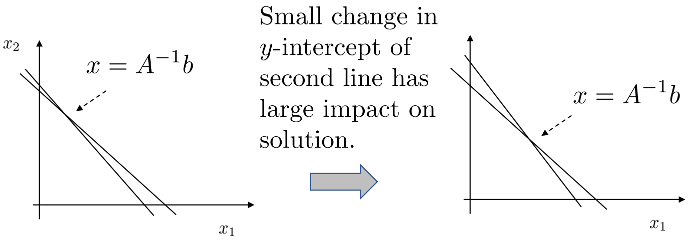
- The problem comes when these hyperplanes are almost parallel.
- In two dimensions we have two lines

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 &= b_1 \\
a_{21}x_1 + a_{22}x_2 &= b_2
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
x_2 &= -\frac{a_{11}}{a_{12}}x_1 + \frac{b_1}{a_{12}} \\
x_2 &= \underbrace{-\frac{a_{21}}{a_{22}}}_{\text{slope}}x_1 + \underbrace{\frac{b_2}{a_{22}}}_{\text{x-intercept}}
\end{aligned}
$$

Matrix Condition Number, cont.



If the two lines are almost parallel then small changes in the slope or x_2-intercept of either line will result in large changes in $x = A^{-1}b$.

Matrix Condition Number, cont.

- Since computers must represent numbers to finite precision, representation errors could significantly change the numerical solution to the equation $Ax = b$.
- The condition number quantifies this effect.

Definition

The condition number of a square matrix is defined to be

$$\mathcal{K}(A) = \|A\| \left\|A^{-1}\right\|$$

where $\|\cdot\|$ is an induced matrix norm usually taken to be the induced 2-norm.

Matrix Condition Number: Derivation

- Given the two equations $Ax = b$ and $(A + \epsilon E)x = b$ where ϵE is a “small” perturbation of A (introduced by finite machine precision of A)
- Let $x_0 = A^{-1}b$ and

$$\begin{aligned} x_E &= (A + \epsilon E)^{-1} b \\ &= [A(I + \epsilon A^{-1} E)]^{-1} b \\ &= (I + \epsilon A^{-1} E)^{-1} A^{-1} b \\ &= \underbrace{(I + \epsilon A^{-1} E)^{-1}}_{\text{perturbation}} x_0 \end{aligned}$$

Matrix Condition Number: Derivation, cont.

Using the Neumann expansion gives
$(I + \epsilon A^{-1}E)^{-1} = \sum_{i=0}^{\infty} (-\epsilon A^{-1}E)^i$. Therefore

$$\begin{aligned} x_E &= (I + \epsilon A^{-1}E)^{-1} A^{-1} b \\ &= (I - \epsilon A^{-1}E) A^{-1} b + O(\|\epsilon E\|^2 x_0) \\ &= A^{-1}b - \epsilon A^{-1} E A^{-1} b + O(\|\epsilon E\|^2 x_0) \\ &= x_0 - \epsilon A^{-1} E x_0 + O(\|\epsilon E\|^2 x_0) \end{aligned}$$

Therefore

$$\underbrace{\frac{\|x_E - x_0\|}{\|x_0\|}}_{\text{relative change in the solution}} \leq \underbrace{\epsilon \left\|A^{-1}\right\| \|E\|}_{\text{want to relate to relative change in } A} + O(\|\epsilon E\|^2)$$

Matrix Condition Number: Derivation, cont.

What is the relative change in A?

$$\frac{\|A - (A + \epsilon E)\|}{\|A\|} = \frac{\epsilon \|E\|}{\|A\|} \triangleq \rho$$

Therefore

$$\frac{\|x_E - x_0\|}{\|x_0\|} \leq \rho \underbrace{\|A^{-1}\| \, \|A\|}_{\mathcal{K}(A)} + O(\|\epsilon E\|^2)$$

The condition number $\mathcal{K}(A)$ relates (approximately) the relative change in A to the relative change in the solution x_0.

Matrix Condition Number: Implication

Rule of Thumb:
If the solution is computed to n digits then only

$$n - \log_{10} \mathcal{K}(A)$$

can be considered to be accurate.