Recap: Lecture 11 2-21-2017

$$A\underline{x} = \underline{b} \qquad A \in \mathbb{R}^{n\times n}$$

iterative

Direct: GE

$$A = \begin{pmatrix} x & x & x \\ & x & x \\ & & x \end{pmatrix} \quad \begin{pmatrix} x & & \\ x & x & \\ x & x & x \end{pmatrix}$$

$$A = \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \end{pmatrix} \to \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \end{pmatrix} \to \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \end{pmatrix}$$

$$A = LU$$

$$\approx \frac{2}{3}n^3 \qquad \begin{pmatrix} 1 & & 0 \\ & 1 & \\ * & & 1 \;\; 1 \end{pmatrix}$$

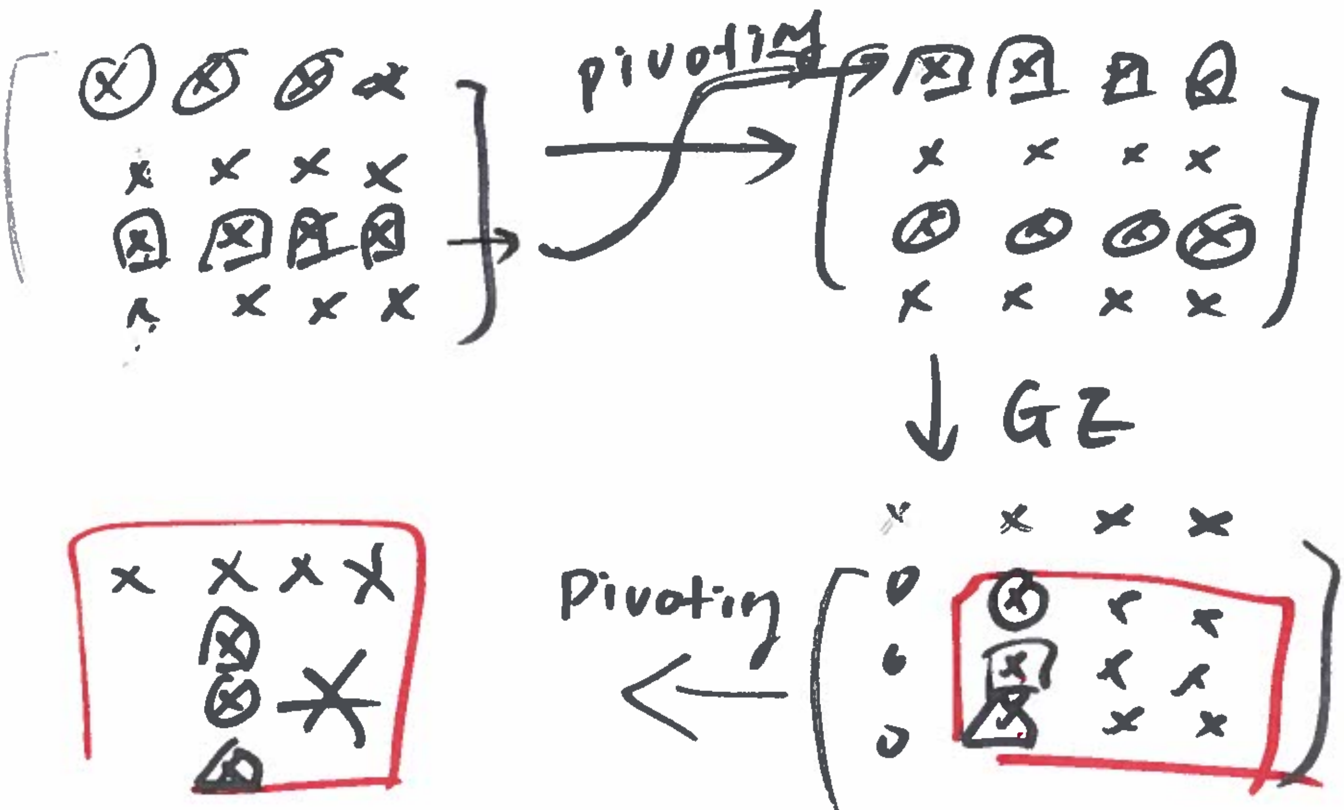
$$A\underline{x} = \underline{b} \iff L\underbrace{U\underline{x}}_{\underline{y}} = \underline{b} \iff \begin{cases} \text{Solve} \\ L\underline{y} = \underline{b} \\ U\underline{x} = \underline{y} \end{cases}$$

GE may fail

$$A=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

pivot = 0

remedy : pivoting



pivoting : two purposes

① GE can continue

② the scheme will be less sensitive to rounding error

to demonstrate this next

Example:

To solve

$$\begin{cases} 10^{-20}u + v = 1 \\ u + 2v = 4. \end{cases}$$

- GE without rounding

$$\left[\begin{array}{cc|c} 10^{-20} & 1 & 1 \\ 1 & 2 & 4 \end{array}\right] \xrightarrow{R_2 - 10^{20}R_1 \to R_2} \left[\begin{array}{cc|c} 10^{-20} & 1 & 1 \\ 0 & 2-10^{20} & 4-10^{20} \end{array}\right]$$

Solve for $v = \dfrac{4-10^{20}}{2-10^{20}}$

Solve for $u = \dfrac{1-v}{10^{-20}} = \dfrac{2\cdot 10^{20}}{10^{20}-2}$

$$x = \begin{bmatrix} u \\ v \end{bmatrix} = \left[\frac{2\cdot 10^{20}}{10^{20}-2}, \frac{4-10^{20}}{2-10^{20}}\right]^T \approx \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

- GE with rounding.

$fl(2-10^{20}) = -10^{20}$

$fl(4-10^{20}) = -10^{20}$

After one step of GE, we have (with rounding)

$$\left[\begin{array}{cc|c} 10^{-20} & 1 & 1 \\ & -10^{20} & -10^{20} \end{array}\right]$$

Back solve $\begin{cases} v = 1 \\ u = 0 \end{cases}$

$\Rightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \underline{x}_c$

$\Rightarrow \underline{x} - \underline{x}_c \approx \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ error: large.

pivot $= 10^{-20} = a_{11}$ (with respect to ε_{mach})

- GE with pivoting, with rounding

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 10^{-20} & 1 & 1 \end{array}\right] \xrightarrow{R_2 - 10^{-20} R_1 \to R_2}$$

$$\Downarrow$$

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & 1-2\cdot 10^{-20} & 1-4\cdot 10^{-20} \end{array}\right]$$

Take into account of rounding

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & 1 & 1 \end{array}\right]$$

back solve this $v = 1$.

$u = 2$

$$\underline{x}_{cc} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

error $\underline{x} - \underline{x}_{cc} = \frac{1}{2^{20}-2}\begin{bmatrix} 4 \\ -2 \end{bmatrix} \approx 10^{-20}$

pivot $=$ $\underline{1}$

§ 2-3 Sensitivity of $A\underline{x}=\underline{b}$ and condition number of A

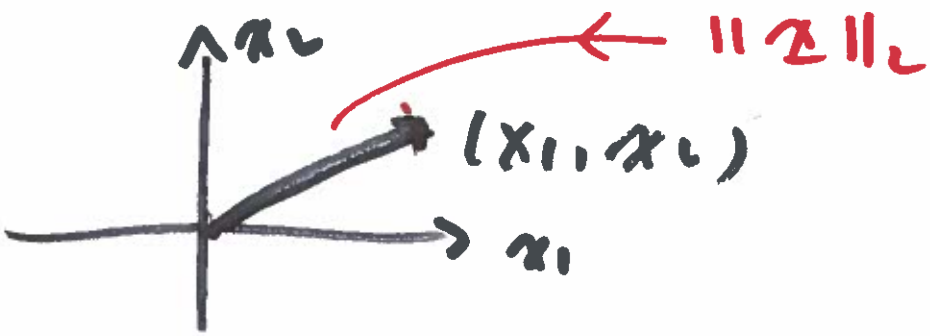
As some preparation.

how to measure the size of a vector

and how to measure the difference?

$$\underline{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\in\mathbb{R}^n \qquad \underline{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}\in\mathbb{R}^n.$$

$$\|\underline{x}\|_2=\left(\sum_{j=1}^{n}|x_j|^2\right)^{\frac{1}{2}} \qquad \text{2-norm}$$



$$\|\underline{x}\|_\infty=\max_{1\le j\le n}|x_j| \qquad \infty\text{-norm}$$

Difference:

$$\|\underline{x}-\underline{y}\|_2=\left(\sum_{j=1}^{n}|x_j-y_j|^2\right)^{\frac{1}{2}}$$

$$\|\underline{x}-\underline{y}\|_\infty=\max_{1\le j\le n}|x_j-y_j|$$

$\|\cdot\|_2$, $\|\cdot\|_\infty$ are two examples of norms of $\mathbb{R}^n$ (measurement).

---

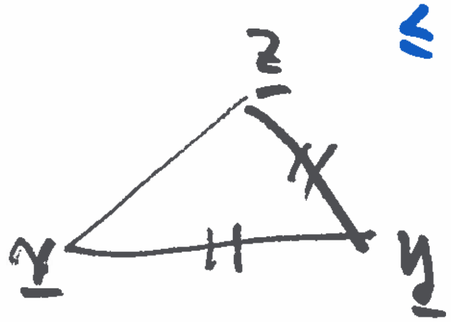
a norm $\|\cdot\|$ of $\mathbb{R}^n$ needs 3 properties:

1) $\|\underline{x}\| \geq 0$ and $\|\underline{x}\| = 0$ if and only if $\underline{x} = \underline{0}$

2) $\|a\underline{x}\| = |a|\,\|\underline{x}\|$ $\quad a \in \mathbb{R}$, $\underline{x} \in \mathbb{R}^n$

3). $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$ $\quad \forall\, \underline{x}, \underline{y} \in \mathbb{R}^n$

$\Leftrightarrow$ $\|\underline{x} - \underline{z}\| \leq \|\underline{x} - \underline{y}\| + \|\underline{y} - \underline{z}\|$.



motivation

How to measure a matrix?

Similar 3 properties are required.

One example:

$$\|A\|_\infty = \max_{1 \le i \le n} \boxed{\sum_{j=1}^{n} |a_{ij}|}$$

Sum of absolute value of ith row

Example:

$$\left\| \begin{bmatrix} -1 \\ 3 \\ -7 \end{bmatrix} \right\|_2 = \left((-1)^2 + 3^2 + (-7)^2 \right)^{\frac{1}{2}}$$

$$= \sqrt{59}$$

$$\| \cdot \|_\infty = 7$$

$$\left\| \begin{pmatrix} -1 & 2 & 1 \\ 3 & 7 & 1 \\ -7 & 10 & 0 \end{pmatrix} \right\|_\infty = 17.$$

Sensitivity of $A\underline{x} = \underline{b}$

is determined by the condition number of A

a chosen norm

$$\text{Cond}(A) = \|A\| \, \|A^{-1}\|$$

with ∞ - norm

$$\text{cond}(A, \infty) = \|A\|_\infty \, \|A^{-1}\|_\infty.$$

overall: the larger the condition number of A is, the more sensitive solving $A\underline{x} = \underline{b}$ is with respect to rounding error.

↳ regardless of $\|\cdot\|$.

[A rule of thumb]

Generally To solve $A\underline{x} = \underline{b}$

on a computer, and get

$\underline{x}_c$, then

$$\boxed{\frac{\|\underline{x} - \underline{x}_c\|}{\|\underline{x}\|} \approx \varepsilon_{mach} \, \mathrm{cond}(A)}$$

double-precision $\varepsilon_{mach} \approx 10^{-16}$

Assume we have examples, with

1) $\mathrm{cond}(A) \approx 10^{2}$, then

relative error $\approx 10^{-14}$

2) $\mathrm{cond}(A) \approx 10^{12}$

relative error $\approx 10^{-16} \cdot 10^{12}$

$\approx 10^{-4}$

Demo

Example: $\underset{\sim}{A} = \begin{bmatrix} 1 & 1 \\ 1+\varepsilon_p & 1 \end{bmatrix}$

$$0 < \varepsilon_p \ll 1$$

$$\text{cond}(A, \infty) = \|A\|_\infty \, \|A^{-1}\|_\infty$$

$$\|A\|_\infty = 2 + \varepsilon_p$$

$$A^{-1} = -\frac{1}{\varepsilon_p} \begin{pmatrix} 1 & -1 \\ -(1+\varepsilon_p) & 1 \end{pmatrix}$$

$$\|A^{-1}\|_\infty = \frac{2+\varepsilon_p}{\varepsilon_p}$$

$$\text{cond}(A, \infty) = \frac{(2+\varepsilon_p)^2}{\varepsilon_p} \approx \frac{4}{\varepsilon_p}$$

$\underset{\sim}{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad \underline{b} = \begin{bmatrix} 2 \\ 2+\varepsilon_p \end{bmatrix}$

exact

on computer, we get $\underset{\sim}{x}_c = A \backslash \underline{b}$

back slash

and examine $\dfrac{\|x - \underline{x}_c\|_\infty}{\|x\|_\infty}$

Example:

Hilbert matrix

$$a_{ij} = \frac{1}{i+j-1} \qquad \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

$H = \text{hilb}(n)$

$$\underline{x} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \qquad H\underline{x} = \underline{b}$$

exact

$$\underline{x_c} = H \setminus \underline{b}$$

$$\frac{\| \underline{x_c} - \underline{x} \|_\infty}{\| \underline{x} \|_\infty} =$$

Remark:

ill-conditioned problems are harder to solve with high precision. One technique to improve this: preconditioning

$$A\underline{x} = \underline{b}$$

$$\Rightarrow BA\underline{x} = B\underline{b} \qquad B \text{ invertible},$$

hopefully

$$\text{cond}(BA) \ll \text{cond}(A)$$

The "good" choice of B:

$B \approx A^{-1}$, easy to compute.

§2-4 Symmetric definite matrix (SPD) and Cholesky factorization.

Review

- determinant of A : det(A)
  $|A|$

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\det A = ad - bc$.

$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$

$$\det(A) = a \underbrace{\begin{vmatrix} e & f \\ h & i \end{vmatrix}}_{\det} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$=$

- Given $A \in \mathbb{R}^{n \times n}$

1) If A is invertible, then

$A\underline{x} \neq \underline{0}$ for any nonzero $\underline{x} \in \mathbb{R}^n$

2) its contrapositive

if $A\underline{x} = \underline{0}$ for some nonzero $\underline{x} \in \mathbb{R}^n$

then A is $\underbrace{\text{singular}}_{\text{non invertible}}$

Pf: 1) by contradiction

$A^{-1}(A\underline{x} = \underline{0})$ for some nonzero $\underline{x} \in \mathbb{R}^n$

Left: $\underbrace{A^{-1}A}_{I}\underline{x} = I\underline{x} = \underline{x}$

Right: $A^{-1}\underline{0} = \underline{0}$ $\Rightarrow$ contradiction.

· Def: Given $A \in \mathbb{R}^{n \times n}$, then $\lambda \in \mathbb{C}$ is an eigenvalue of A, if

$A\underline{x} = \lambda \underline{x}$ for some $\underbrace{\text{nonzero}}$ vector $\underline{x} \in \mathbb{C}^n$.

This $\underline{x}$: eigenvector of A associated with λ.

$(\lambda, \underline{x})$ eigen-pair

How to find eigenvalues?

$A\underline{x} = \underbrace{\lambda \underline{x}}$ for some nonzero $\underline{x}$

$= \lambda I \underline{x}$

$\Leftrightarrow (A - \lambda I)\underline{x} = \underline{0} \quad \underline{x} \neq \underline{0}$

$\Leftrightarrow A - \lambda I$ is singular.

$\Leftrightarrow \boxed{\det(A - \lambda I) = 0}$ ← to solve

↳ polynomial of degree n.

"n eigenvalues"

Examples : find eigenvalues of A

$$A = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$$

To calculate :

1) $\det(A - \lambda I) = 0$

$\Leftrightarrow \det(\lambda I - A) = 0$

$\Leftrightarrow \det \begin{pmatrix} \lambda - 2 & -2 \\ -2 & \lambda - 5 \end{pmatrix}$

$= (\lambda - 2)(\lambda - 5) - (-2)^2$

$= \lambda^2 - 7\lambda + \underbrace{10 - 4}_{6} = 0$

$\lambda_1 = 1, \lambda_2 = 6$

2) $\det\left(\lambda I - \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\right)$

$= \begin{vmatrix} \lambda - 2 & 1 \\ -1 & \lambda - 2 \end{vmatrix}$

$= (\lambda-2)^2 - (-1) = (\lambda-2)^2 + 1 = 0$

$(\lambda - 2)^2 = -1$

$\lambda - 2 = \pm i \Rightarrow$

$\lambda_1 = 2 + i, \quad \lambda_2 = 2 - i$

3)
$\det\left(\lambda I - \begin{pmatrix} a & b & c \\ & d & e \\ & & f \end{pmatrix}\right)$

$= \begin{vmatrix} \lambda - a & -b & -c \\ & \lambda - d & -e \\ & & \lambda - f \end{vmatrix}$

$= (\lambda - a)\begin{vmatrix} \lambda - d & -e \\ 0 & x - f \end{vmatrix} + b\underbrace{\begin{vmatrix} 0 & x \\ 0 & x \end{vmatrix}}_{=0}$

$-c\underbrace{\begin{vmatrix} 0 & x \\ 0 & x \end{vmatrix}}_{=0} = (\lambda - a)(\lambda - d)(\lambda - f) = 0$

$\boxed{\begin{matrix}\lambda_1 = a \\ \lambda_2 = d \\ \lambda_3 = f.\end{matrix}}$