Numerical Solution

Goal: to find $\underline{\underline{A}}\,\underline{x} = \underline{b}$

$\underline{x}$ ⇓ to be found

If $\det(\underline{\underline{A}}) \neq 0$, you could do

$$\underline{x} = \underline{\underline{A}}^{-1}\underline{b}$$

You never actually find $\underline{\underline{A}}^{-1}$

Two Methods

① Iterative Method.

⇒ Matrix-Vector Projection

⇒ a sequence of vectors

⇒ converge to $\underline{x}$

GMRES, GMR etc.

→ Krylov Subspace Method.

② Decomposition Method.

Given $\underline{\underline{A}}$ from $\underline{\underline{B}}$ and $\underline{\underline{C}}$

such that $\underline{\underline{A}} = \underline{\underline{C}}\ \underline{\underline{B}}$

with finding $\underline{y}$ so that $\underline{\underline{C}}\ \underline{y} = \underline{b}$

and $\underline{x}$ such that $\underline{\underline{B}}\ \underline{x} = \underline{y}$ is

easier than $\underline{\underline{A}}\ \underline{x} = \underline{b}$?

$$\underline{y} = \underline{\underline{C}}^{-1}\ \underline{b}$$

$$\underline{\underline{B}}\ \underline{x} = \underline{y} = \underline{\underline{C}}^{-1}\ \underline{b}$$

$$\Rightarrow \underline{x} = \underline{\underline{B}}^{-1}\ \underline{\underline{C}}^{-1}\ \underline{b}$$

$$\underline{x} = \underline{\underline{A}}^{-1}\ \underline{b}$$

$$\boxed{\underline{\underline{A}} = \underline{\underline{C}}\ \underline{\underline{B}}}$$

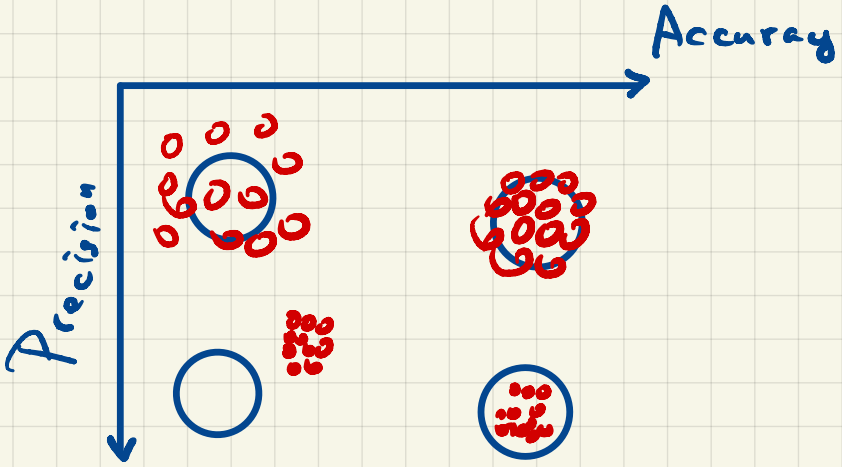
LU
QR
SVD

] exact method

Accuracy v.s. Precision

Accuracy: How close are the approximations to the truth

Precision: How close are the approximations to each other



Numerical Errors come from

- choice of the model
- numerical approximation
- limits of data representation
- Errors in implementation
- fluctuation in hardware

Truth = Approximation + Error

$$x^* = x + \varepsilon$$

$$\Rightarrow \varepsilon = x^* - x$$ ← usually not a good error

Relative error

$$\varepsilon_{rel} = \frac{x^* - x}{x^*}$$

Note: In the iterative methods, you have a sequence of x_i: $x_0, x_1, \ldots \to \underline{x}$

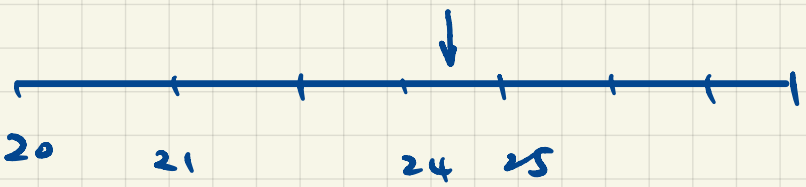
$$\underline{r}_i = \underline{\underline{A}}\,\underline{x}_i - \underline{b}$$

↑

residual

not the error

Significant Figures



$24 < x < 25$

$x = 24.75$?

24.525 ?

24.5 ?

Sig. figure tell you how much <u>useful</u> information you have

Typically the known digit + one more

ex).

53800 ← 5 sig. digit.

5.38×10^4 ← 3 "

5.380×10^4 ← 4 "

0.01234 ←

$= 1.234 \times 10^{-2}$ 4 # sig. digits.

Trailing zeros count, leading zero do not

Computer store # in binary:

0 . 1.

%	%	%	%	%	%

Finite # of bits = Finite # of digits.

32 bits = 7 sig. digits

64 bits = 15 sig. digits

Round - off Errors.

type

integer : -1, 1, 0, 1, 2

float / double / long double : decimal

π.

character : 'a', 'c' ...

Consider the division of $\frac{7}{2} = 3.5$

Let 2, 7 and the answer to be integer

$\Rightarrow \frac{7}{2} = 3$

ROE. = 0.5

In double precision, you set a minimum error at

$$2^{-52} \approx 2.22 \times 10^{-16}$$

In double precision,

$$1 \neq 1 + 10^{-20}$$

eps : matlab for machine precision

float point

$a \neq b$

abs(a-b) < tol

↑

a # I can accept

Condition Number

The **condition number** of a matrix $\underline{\underline{A}}$ tells us how errors in $\underline{b}$ influences errors in $\underline{x}$ when solving $\underline{\underline{A}}\,\underline{x} = \underline{b}$

$$\underline{\underline{A}}\,(\underline{x} + \underline{\delta x}) = (\underline{b} + \underline{\delta b})$$

$\underline{\delta x}$: induced error in $\underline{x}$

$\underline{\delta b}$: error in $\underline{b}$

Goal: relate $\dfrac{\|\delta x\|}{\|x\|} \propto \dfrac{\|\delta b\|}{\|b\|}$

$$\cancel{\underline{\underline{A}}\,\underline{x}} + \underline{\underline{A}}\,\underline{\delta x} = \cancel{\underline{b}} + \underline{\delta b}$$

$$\Rightarrow \quad \underline{\underline{A}}\,\underline{\delta x} = \underline{\delta b}$$

$$\|\delta b\| = \|A\,\delta x\| \le \|A\|\,\|\delta x\|$$

Also, $\delta x = A^{-1}\,\delta b$

$$\Rightarrow \|\delta x\| = \|A^{-1}\,\delta b\| \le \|A^{-1}\|\,\|\delta b\|$$

$$\|b\| = \|A\,x\| \le \|A\|\,\|x\|$$

$$\frac{\|\delta x\|}{\|x\|} \le \|A^{-1}\|\,\|\delta b\|\,\frac{\|A\|}{\|b\|}$$

$$\frac{\|\delta x\|}{\|x\|} \le \boxed{\|A^{-1}\|\,\|A\|}\,\frac{\|\delta b\|}{\|b\|}$$

Define the condition number or $K(A)$

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

indicates how error grow

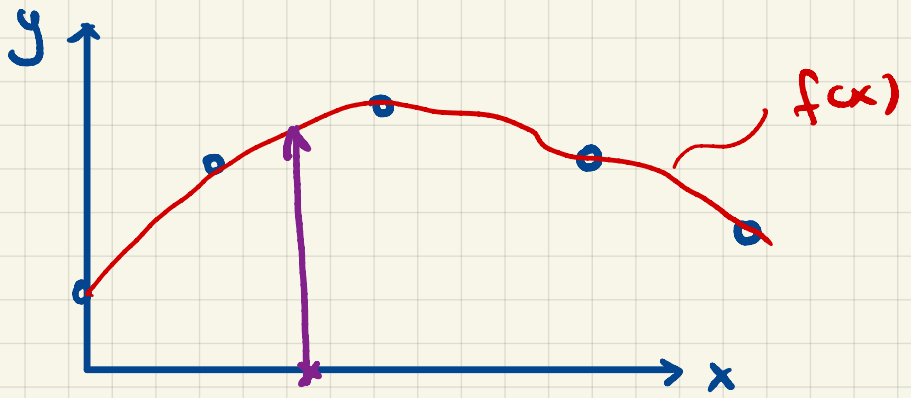
ex).

let $\frac{\|\delta \underline{b}\|}{\|\underline{b}\|} \sim 10^{-16}$ with $\kappa(\underline{\underline{A}}) \approx 10^{6}$

$$\frac{\|\delta \underline{x}\|}{\|\underline{x}\|} \leq \kappa(\underline{\underline{A}}) \frac{\|\delta \underline{b}\|}{\|\underline{b}\|} \sim 10^{-10}$$

Interpolation

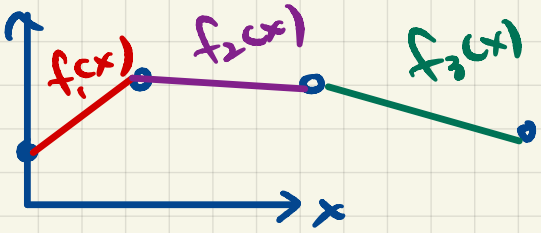
Given a data set, how do you determine values between the data.



Goal: Find f(x)

① Global interpolation: Use all data to construct one f(x) over the entire domain

② Piecewise interpolation: Use individual interpolants between data points



Polynominal Interpolation - Global

n data pt.

(x_1, y_1)

$\vdots$

(x_n, y_n)

$\Rightarrow$ n-1 polynominal

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{n-1} x^{n-1}$$

$a_0 \cdots a_{n-1} \Rightarrow$ n coefficients

$$f(x_1) = a_0 + a_1 x_1 + \cdots + a_{n-1} x_1^{n-1} = y_1$$

$\vdots$

$$f(x_n) = a_0 + a_1 x_n + \cdots + a_{n-1} x_{n-1}^{n-1} = y_n$$

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix}$$

Vandermonde Matrix

If $x_i \neq x_j$ & $i \neq j$, $\det(\underline{\underline{V}}) \neq 0$

$$\Rightarrow \quad \underline{a} = \underline{\underline{V}}^{-1} \underline{y}$$

but $K(\underline{\underline{V}})$ is huge