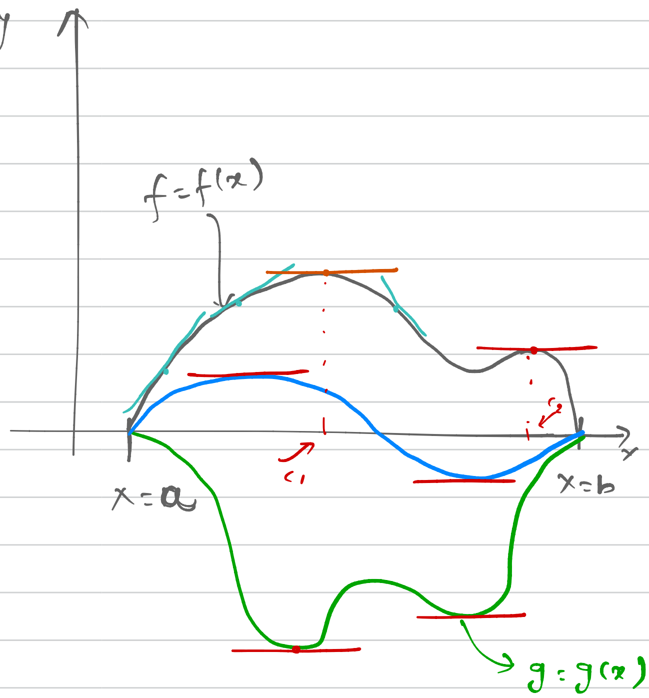
Lecture 13

Taylor's series
- Rolle's theorem
- Mean value theorem
- Taylor's theorem or Taylor's series

Roots problem
- Error in fixed point iteration method
- Error in Newton-Raphson method

vector and matrix
- Errors in $Ax = b$ problem
- Condition number of matrix A
- Norm of vector and matrix

- ## Rolle's theorem



If a continuous & differentiable function $f : [a, b] \longrightarrow \mathbb{R}$ is such that

$f(a) = 0, \quad f(b) = 0.$

Then, there exist at least one point c such that $a < c < b$ and

$$\frac{df}{dx}(c) = 0 .$$

[f needs to be continuous at all points $x \in [a, b]$
f needs to be differentiable at all points $x \in (a, b)$]

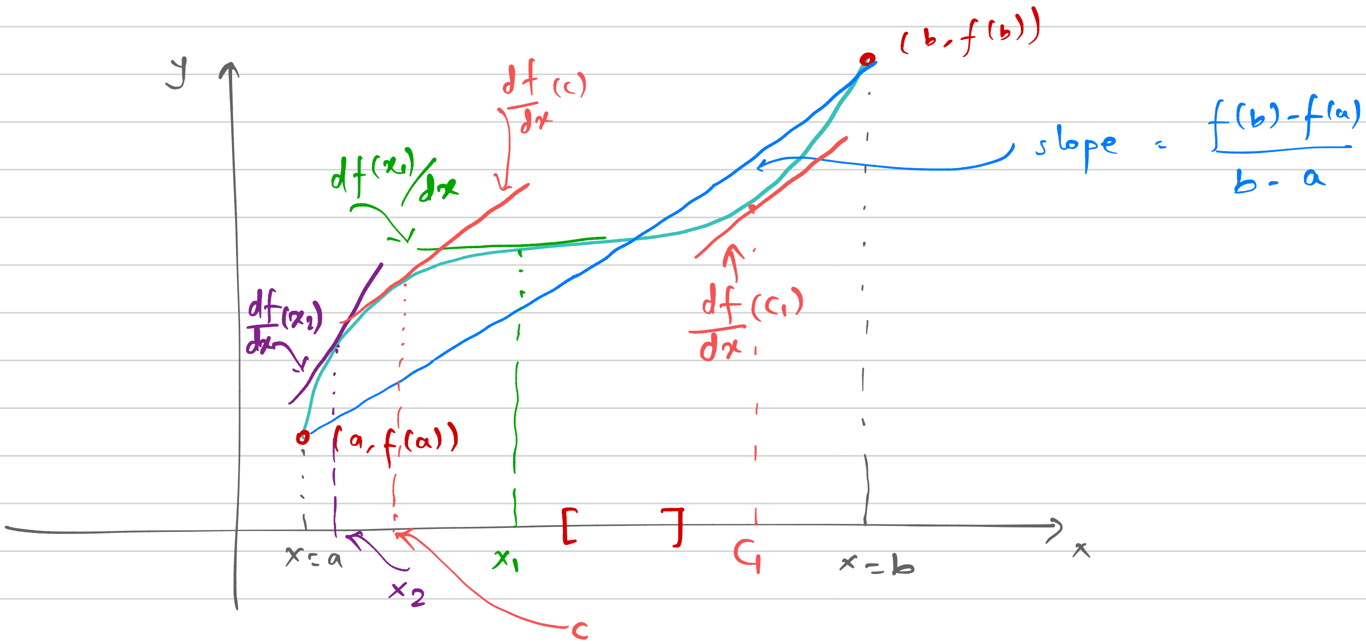
- Mean value theorem

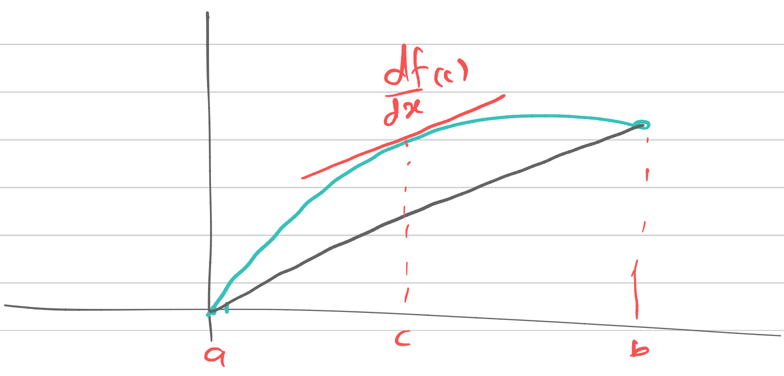
If function $f: [a,b] \longrightarrow \mathbb{R}$, $a<b$, is continuous and differentiable at all points in interval $[a,b]$. Then there exist atleast one point $c \in (a,b)$ $(a<c<b)$ such that

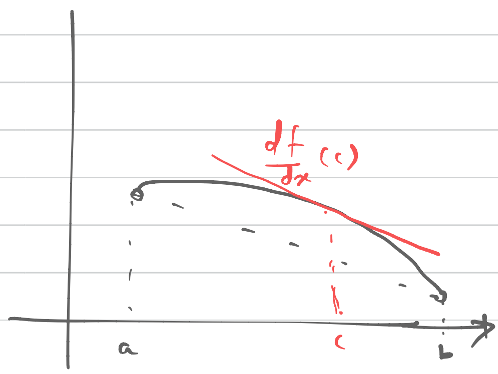
$$\frac{df}{dx}(c) = \underbrace{\frac{f(b)-f(a)}{b-a}}.$$

consider a straight line in (x,y) plane that connects $(a, f(a))$ and $(b, f(b))$ then

slope of that line $= \dfrac{f(b)-f(a)}{b-a}$







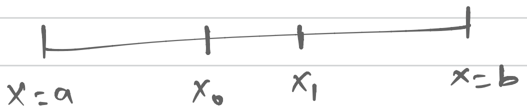
From <u>mean value theorem</u>

$$\boxed{f(b) = f(a) + \frac{df}{dx}(c)\,(b-a)}$$

there is one such $c \in (a,b)$

c depends on f, a, b.

Instead of applying mean value theorem in interval $[a,b]$, I will



apply in interval $[x_0, x_1]$

$$\boxed{f(x_1) = f(x_0) + f'(d)(x_1 - x_0)}$$

$d \in (x_0, x_1)$

$f : [a,b] \to \mathbb{R}$, continuous & differentiable

MVT $\rightarrow$ $f'(c) = \dfrac{f(b) - f(a)}{b-a}$ $\qquad f' = \dfrac{df}{dx}$

I can pick two points x_0, x_1 s.t. $a \leq x_0 < x_1 \leq b$

then MVT is still applicable to interval $[x_0, x_1]$

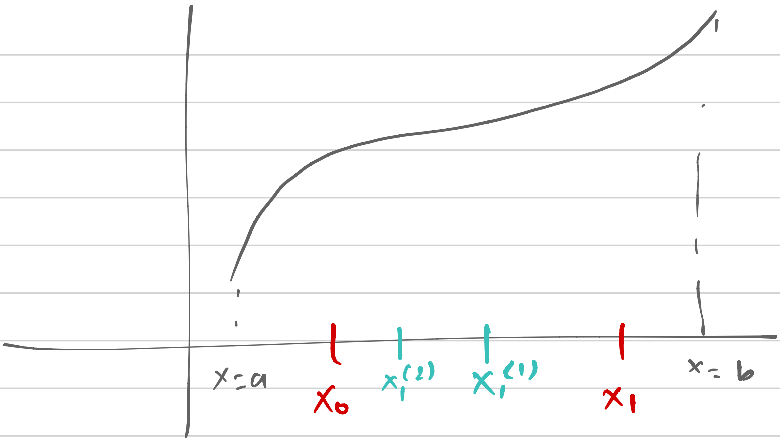
$$f'(d) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

if $|x_0 - x_1|$ is very small and since d is such that

$x_0 < d < x_1$

we can approximately write

$$\boxed{f(x_1) = f(x_0) + f'(d)(x_1 - x_0) \approx f(x_0) + f'(x_0)(x_1 - x_0)}$$



- <u>Taylor's theorem / Taylor's series</u>

$f : [a,b] \longrightarrow \mathbb{R}$ is continuous at all $x \in [a,b]$

and <u>derivatives upto n^{th} order of function f exist</u>

at all points $x \in (a,b)$,

$$\frac{df}{dx}, \frac{d^2f}{dx^2}, \ldots, \frac{d^nf}{dx^n} \quad \text{exist for all } x \in (a,b).$$

Then <u>there is some $c \in (a,b)$</u> such that

$$\boxed{f(b) = f(a) + \frac{df}{dx}(a)(b-a) + \frac{1}{2!}\frac{d^2f}{dx^2}(a)(b-a)^2 + \ldots + \frac{1}{(n-1)!}\frac{d^{(n-1)}f}{dx^{(n-1)}}(a)(b-a)^{n-1} + \frac{1}{n!}\frac{d^nf}{dx^n}(c)(b-a)^n}$$

c exist and depends on $f, \frac{df}{dx}, \frac{d^2f}{dx^2}, \ldots, \frac{d^nf}{dx^n}$

a, b

<u>Observations</u>.

(•) put $n=1 \longrightarrow$ you get <u>MVT</u>

(·) Suppose upto N^{th} derivative exist

for any $1 \le n \le N$, Taylor's theorem holds

$$f(b) = \sum_{k=0}^{n-1} \frac{1}{k!} \frac{d^k f}{dx^k}(a)(b-a)^k + \frac{1}{n!} \frac{d^n f}{dx^n}(c)\,(b-a)^n$$

point c exist but you don't know

where at $k=0$ (i) $\frac{d^0 f}{dx^0} = f$

(ii) $0! = 1$

$$f(b) = \sum_{k=0}^{n-1} \frac{1}{k!} \frac{d^k f}{dx^k}(a)(b-a)^k + \underbrace{\frac{1}{n!} \frac{d^n f}{dx^n}(c)}_{\text{Remainder } r_n(a,b)}(b-a)^n$$

$$f(b) = \sum_{k=0}^{n-1} \frac{1}{k!} \frac{d^k f}{dx^k}(a)(b-a)^k + r_n(a,b)\,(b-a)^n$$

if b and a are such that $|b-a|$ is small

then $(b-a)^n$ is smaller and smaller with higher n

$$\boxed{f(b) \approx \sum_{k=0}^{n-1} \frac{1}{k!} \frac{d^k f}{dx^k}(a)(b-a)^k}$$

Sketch of Taylor's theorem proof:

Consider a function $F : (a,b) \to \mathbb{R}$ s.t.

- $\frac{dF}{dx}, \ldots, \frac{d^nF}{dx^n}$ exist

- $F(a) = 0, \quad F(b) = 0$

- $\frac{dF}{dx}(a) = \frac{d^2F}{dx^2}(a) = \cdots = \frac{d^{n-1}F}{dx^{n-1}}(a) = 0$

Then there is $c \in (a,b)$ s.t

$$\frac{d^nF}{dx^n}(c) = 0.$$

Proof: We have $F(a) = 0, \quad F(b) = 0$

then from Rolle's theorem there is

point $c_1 \in (a,b)$ s.t

$$\frac{dF}{dx}(c_1) = 0$$

Take a new function $g = \frac{dF}{dx}$

Note that

$$g(a) = 0, \quad g(c_1) = 0$$

So $g : (a, c_1] \to \mathbb{R}$ st $g(a) = 0, \ g(c_1) = 0$

there is $c_2 \in (a, c_1)$ s.t

(s.t. "such that")

$$\frac{dg}{dx}(c_2) = 0 \Rightarrow \frac{d^2F}{dx^2}(c_2) = 0$$

⋮

You show by repeating above procedure, following

- $g = \frac{d^{n-2}F}{dx^{n-2}}$

- $g(a) = 0$
- $g(c_{n-2}) = 0$

& there is $c_{n-1} \in (a, c_{n-2})$ s.t

$$\frac{dg}{dx}(c_{n-1}) = 0 \Rightarrow \underbrace{\frac{d^{n-1}F}{dx^{n-1}}(c_{n-1}) = 0}$$

$$h = \frac{d^{n-1}F}{dx^{n-1}}$$

- $h(a) = 0$
- $h(c_{n-1}) = 0$

further application of Rolle's theorem ensures there is a point $c_n \in (a, c_{n-1})$ s.t.

$$\frac{dh}{dx}(c_n) = 0$$

$$\Rightarrow \boxed{\frac{d^nF}{dx^n}(c_n) = 0}$$

proves our original theorem that there is a point $c \in (a,b)$ s.t $\frac{d^nF}{dx^n}(c) = 0$

Now using above result for function F, we prove Taylor's theorem: In Taylor's theorem we have function f

- f is continuous at all $x \in [a,b]$
- f is differentiable at all $x \in (a,b)$

First look at polynomial as follows

$$P(x) = \sum_{k=0}^{n} \alpha_k (x-a)^k$$

(·) $P(a) = \alpha_0$, (·) $\frac{dP}{dx}(a) = \cdots = \frac{d^{n-1}P}{dx^{n-1}}(a) = 0$

Define $F(x) = f(x) - P(x)$

· $\alpha_k = \frac{1}{k!} \frac{d^k f}{dx^k}(a)$ $k = 0, 1, 2, \ldots, n-1$

$$\boxed{\begin{gathered} \frac{d^0 f}{dx^0} = f \\ 0! = 1 \end{gathered}}$$

$$\boxed{\begin{gathered} \cdot\ F(a) = 0,\ F(b) = 0 \\ \cdot\ \frac{dF}{dx}(a) = 0 \\ \frac{d^2F}{dx^2}(a) = 0 \\ \vdots \\ \frac{d^{n-1}F}{dx^{n-1}}(a) = 0 \end{gathered}}$$

· $\alpha_n = \frac{1}{(b-a)^n}\left(f(b) - \sum_{k=0}^{n-1} \alpha_k (b-a)^k \right)$

Since F satisfies all the conditions, from previous theorem there is some $c \in (a,b)$ s.t.

$$\frac{d^n F}{dx^n}(c) = 0$$

$$\Rightarrow \frac{d^n f}{dx^n}(c) - \frac{d^n P}{dx^n}(c) = 0$$

$$\Rightarrow \frac{d^n f}{dx^n}(c) = \frac{d^n P}{dx^n}(c)$$

$$\Rightarrow \frac{d^n f}{dx^n}(c) = \frac{n!}{(b-a)^n}\left(f(b) - \sum_{k=0}^{n-1} \alpha_k (b-a)^k \right)$$

$$\Rightarrow \frac{1}{n!} \frac{d^n f}{dx^n}(c)\,(b-a)^n = f(b) - \sum_{k=0}^{n-1} \alpha_k (b-a)^k$$

$$\Rightarrow \boxed{f(b) = \sum_{k=0}^{n-1} \alpha_k (b-a)^k + \frac{1}{n!} \frac{d^n f}{dx^n}(c)\,(b-a)^n}$$

using the fact that $\alpha_k = \frac{1}{k!} \frac{d^k f}{dx^k}(a)$

Taylor's theorem is proved.

Roots problem

- **Error in fixed point iteration method**

find x such that $x = g(x)$, g is some function

$g : [a,b] \to \mathbb{R}$

↳ x_0, $x_1 = g(x_0)$, $x_2 = g(x_1)$, ..., $x_i = g(x_{i-1})$, ...

Error at i^{th} iteration $= x_i - x_{i-1} = E_i$

$$E_i = x_i - x_{i-1} = g(x_{i-1}) - x_{i-1} = g(x_{i-1}) - g(x_{i-2})$$

from MVT: there is $\eta \in (a,b)$

$$g(x_{i-1}) = g(x_{i-2}) + g'(\eta)\underbrace{(x_{i-1} - x_{i-2})}_{E_{i-1}}$$

$$\boxed{E_i = g'(\eta)\, E_{i-1}}$$

$$e_i = |E_i| \Rightarrow \boxed{e_i \le |g'(\eta)|\, e_{i-1}} \le \underbrace{\left\{\max_{x \in (a,b)} |g'(x)|\right\} e_{i-1}}$$

$$\Downarrow$$

$$\boxed{e_i \le \left\{\max_{x \in (a,b)} |g'(x)|\right\} e_{i-1}}$$

if $\displaystyle\max_{x \in (a,b)} |g'(x)| < 1$

then $e_i \longrightarrow 0$

- **Error in Newton-Raphson method**

find x s.t. $f(x)=0$ $\qquad f:[a,b]\to\mathbb{R}$

$$x_0 \longrightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$\downarrow$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

$$\downarrow$$

$$\vdots$$

$$x_i = x_{i-1} - \frac{f(x_{i-1})}{f'(x_{i-1})}$$

Define "True error"

$$E_{t,i} = x_t - x_i$$

x_t → exact solution i.e. $f(x_t)=0$

x_i → is ith approximate solution

Consider

$$E_{t,i+1} = x_t - x_{i+1}$$

$$= x_t - \left(x_i - \frac{f(x_i)}{f'(x_i)}\right)$$

$$= \underbrace{x_t - x_i}_{E_{t,i}} + \frac{f(x_i)}{f'(x_i)}$$

$$\Rightarrow \quad E_{t,i+1} = E_{t,i} + \frac{f(x_i)}{f'(x_i)}$$

<u>because</u> $\quad E_{t,i} = x_t - x_i \Rightarrow x_i = x_t - E_{t,i}$

<u>then</u>

<u>from Taylor's series</u>

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2}f''(a)(b-a)^2 + r_1$$

$$f'(b) = f'(a) + f''(a)(b-a) + r_2$$

Substitute $\quad b = x_i, \quad a = x_t$

$$f(x_i) = \cancel{f(x_t)}^{0} + f'(x_t)(-E_{t,i}) + \frac{1}{2}f''(x_t)(E_{t,i})^2 + r_1$$

$$f'(x_i) = f'(x_t) + f''(x_t)(-E_{t,i}) + r_2$$

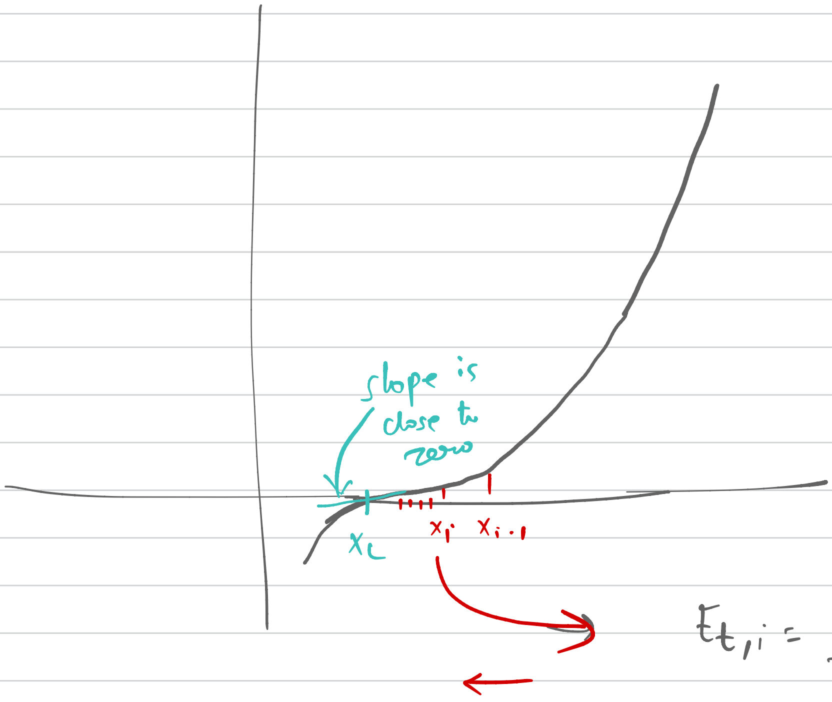
$$E_{t,i+1} = E_{t,i} + \frac{-f'(x_t)E_{t,i} + \frac{1}{2}f''(x_t)(E_{t,i})^2 + r_1}{f'(x_t) - f''(x_t)E_{t,i} + r_2}$$

$$= \frac{E_{t,i}\cancel{f'(x_t)} - f''(x_t)(E_{t,i})^2 + \cancel{r_2 E_{t,i}}^{0} - \cancel{f'(x_t)}E_{t,i} + \frac{1}{2}f''(x_t)(E_{t,i})^2 + \cancel{r_1}^{0}}{f'(x_t) - f''(x_t)E_{t,i} + r_2}$$

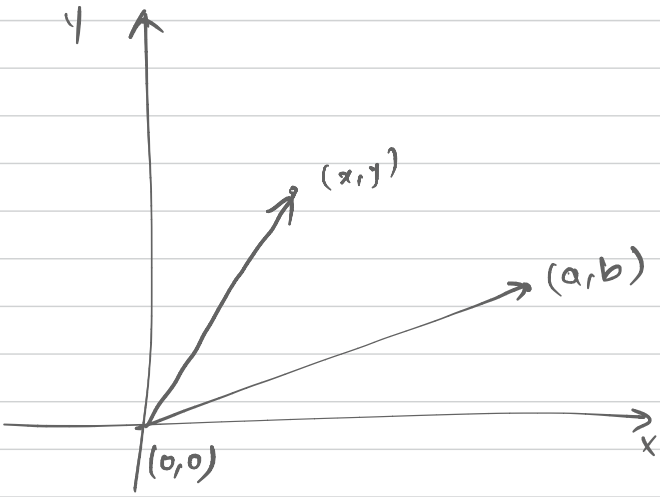
$$\boxed{\epsilon_{t,i+1} = \frac{-\frac{1}{2} f''(x_t)}{f'(x_t) - f''(x_t)\,\epsilon_{t,i}} (\epsilon_{t,i})^2}$$

$$\Rightarrow \boxed{\epsilon_{t,i+1} \approx \frac{-f''(x_t)}{2 f'(x_t)} (\epsilon_{t,i})^2}$$

Error reduces at quadratic speed.

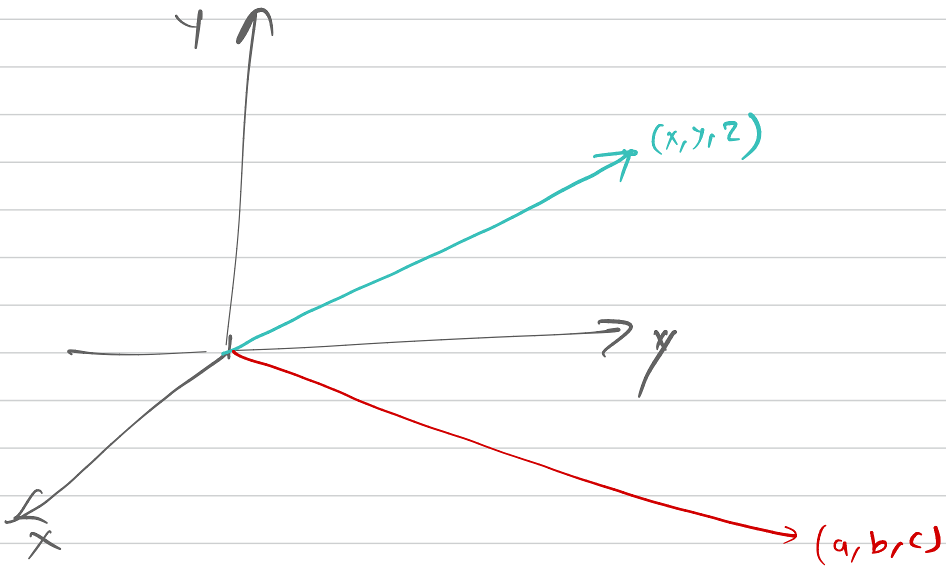


$$\epsilon_{t,i} = \frac{|f''(x_t)|}{2|f'(x_t)|} (\epsilon_{t,i-1})^2$$



(i) direction

(ii) length $l = \sqrt{x^2 + y^2}$



length $= \sqrt{x^2 + y^2 + z^2}$

length $= \sqrt{a^2 + b^2 + c^2}$

in n-dimension (meaning that vector x has "n elements")

length $= \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$

Norm : generalizes the notion of length/magnitude to vectors and matrices.

There are many ways to define a "norm" of a vector

(i) Euclidean norm of a vector

$$\|x\|_e = \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

coincides with our notion of length of vectors.

(ii) $$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

(iii) $$\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$$

$$= \max_{1 \le i \le n} \{|x_i|\}$$

(iv) $$\|x\|_p = \left(\sum_{i=1} |x_i|^p\right)^{1/p} \quad \text{for any } p \in [1, \infty)$$

p-th norm of vector x

Norm of a matrix :

$Ax = b$

$$\|Ax\|_p \le \|A\|_p \|x\|_p$$

$$\|A\|_p := \max_{\substack{x \text{ vectors} \\ \|x\|_p \neq 0}} \frac{\|Ax\|_p}{\|x\|_p} = \max_{\substack{x \text{ vectors} \\ \|x\|_p = 1}} \|Ax\|_p$$

if $p = 1$, $\|x\|_1$, $\|A\|_1$

$p = 2$, $\|x\|_2$, $\|A\|_2$ ← Frobenius norm

$p = \infty$, $\quad \|x\|_\infty$, $\quad \|A\|_\infty$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$\|A\|_p = \max_{\substack{\text{all } x \text{ vectors} \\ \|x\|_p = 1}} \|Ax\|_p$$

Representation of norm of A in terms of coefficients

(i) $p = 1$,

$$\|A\|_1 = \max_{j=1,2,\dots,n} \left(\sum_{i=1}^{n} |a_{ij}| \right)$$

$$\begin{bmatrix} a_{11} & & a_{1j} & & a_{1n} \\ a_{21} & & a_{2j} & & a_{2n} \\ a_{i1} & a_{i2} & a_{ij} & \cdots & a_{in} \\ a_{n1} & & a_{nj} & & a_{nn} \end{bmatrix}$$

(ii) $p = \infty$,

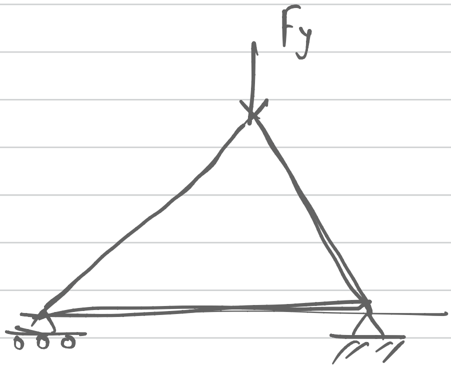
$$\|A\|_\infty = \max_{i=1,2,\dots,n} \left(\sum_{j=1}^{n} |a_{ij}| \right)$$

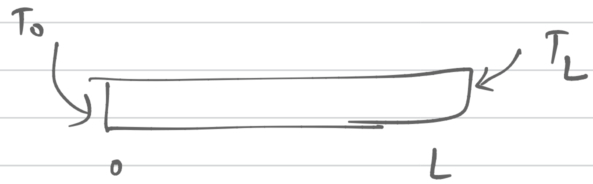
(iii) $p = 2$,

$$\|A\|_2 = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij})^2}$$

- Error in $Ax = b$ problem

Original problem $\quad Ax = b$, $\quad x$ is a solution





$$Ax = b \qquad x = \begin{bmatrix} T(x_1) \\ \vdots \\ \vdots \\ T(x_{n-1}) \end{bmatrix}$$

new problem $\quad Ax = \tilde{b}$, $\quad \tilde{x}$ is a solution

Our goal is to relate errors in data $\dfrac{\|b - \tilde{b}\|}{\|b\|}$

to errors in solution $\dfrac{\|x - \tilde{x}\|}{\|x\|}$

propogation of errors

error in data $\longrightarrow$ errors in solution

b

$Ax = b$, $\quad A\tilde{x} = \tilde{b}$, $\quad A(x - \tilde{x}) = b - \tilde{b}$

$\Rightarrow x = A^{-1}b$ $\quad \tilde{x} = A^{-1}\tilde{b}$, $\quad x - \tilde{x} = A^{-1}(b - \tilde{b})$

$$\boxed{\|x - \tilde{x}\|_p = \|A^{-1}(b - \tilde{b})\|_p \leq \|A^{-1}\|_p \, \|b - \tilde{b}\|_p}$$

we also have. $\|b\|_p = \|Ax\|_p \le \|A\|_p \|x\|_p$

$$\Rightarrow \frac{\|x-\tilde{x}\|_p}{\|x\|_p} \le \frac{\|A^{-1}\|_p \|b-\tilde{b}\|_p}{\|x\|_p}$$

$$\|x\|_p \ge \frac{\|b\|_p}{\|A\|_p}$$

$$\Rightarrow \frac{\|A\|_p}{\|b\|_p} \ge \frac{1}{\|x\|_p}$$

$$= \frac{1}{\|x\|_p}\left(\|A^{-1}\|_p \|b-\tilde{b}\|_p\right)$$

$$\le \frac{\|A\|_p}{\|b\|_p} \|A^{-1}\|_p \|b-\tilde{b}\|_p$$

$$= \left[\|A\|_p \|A^{-1}\|_p\right]\left[\frac{\|b-\tilde{b}\|_p}{\|b\|_p}\right]$$

$$\Rightarrow \boxed{\frac{\|x-\tilde{x}\|_p}{\|x\|_p} \le \underbrace{\left(\|A\|_p \|A^{-1}\|_p\right)}_{\text{depends only on } A} \frac{\|b-\tilde{b}\|_p}{\|b\|_p}}$$

condition number of matrix A

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

$$\boxed{\frac{\|x-\tilde{x}\|_p}{\|x\|_p} \le \mathrm{Cond}(A) \frac{\|b-\tilde{b}\|_p}{\|b\|_p}}$$

Roughly $\qquad \text{cond}(A) \propto \dfrac{1}{\det(A)}$

specially if A is singular i.e. A^{-1} does not exist

then $\qquad \text{cond}(A) = \infty$

$$\boxed{\text{cond}(A) \geq 1}$$

$$A = \begin{bmatrix} 1.01 & 0.99 \\ 0.99 & 1.01 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 25.25 & -24.75 \\ -27.75 & 25.25 \end{bmatrix}$$

$$\text{cond}(A) = \|A\|_\infty \, \|A^{-1}\|_\infty = 100$$

$$2 = \max\{2, 2\} \qquad\qquad \max\{50, 50\} = 50$$

$$\begin{array}{r} 25.25 \\ +\,24.75 \\ \hline = 50 \end{array}$$

$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} \leq 100 \, \frac{\|b - \tilde{b}\|_\infty}{\|b\|_\infty}$$

$$A = \begin{bmatrix} 1.001 & 0.999 \\ 0.999 & 1.001 \end{bmatrix} \longrightarrow A^{-1}\,?, \quad \text{cond}(A)$$