Bernoulli Random Variable

used for indicator random variables

$$X = \begin{cases} 1 & \text{if some event } E \text{ has occurred.} \\ 0 & \text{if not} \end{cases}$$

with pmf

probulity of E occurring

$$p_X(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

Then

$$\mu = E(X) = \sum x \, p_X(x)$$

$$= 0 \cdot (1-p) + 1 \cdot p = p$$

$$Var(X) = E((X-\mu)^2)$$

$$= \sum (x-p)^2 p_X(x)$$

$$= (-p)^2 (1-p) + (1-p)^2 p$$

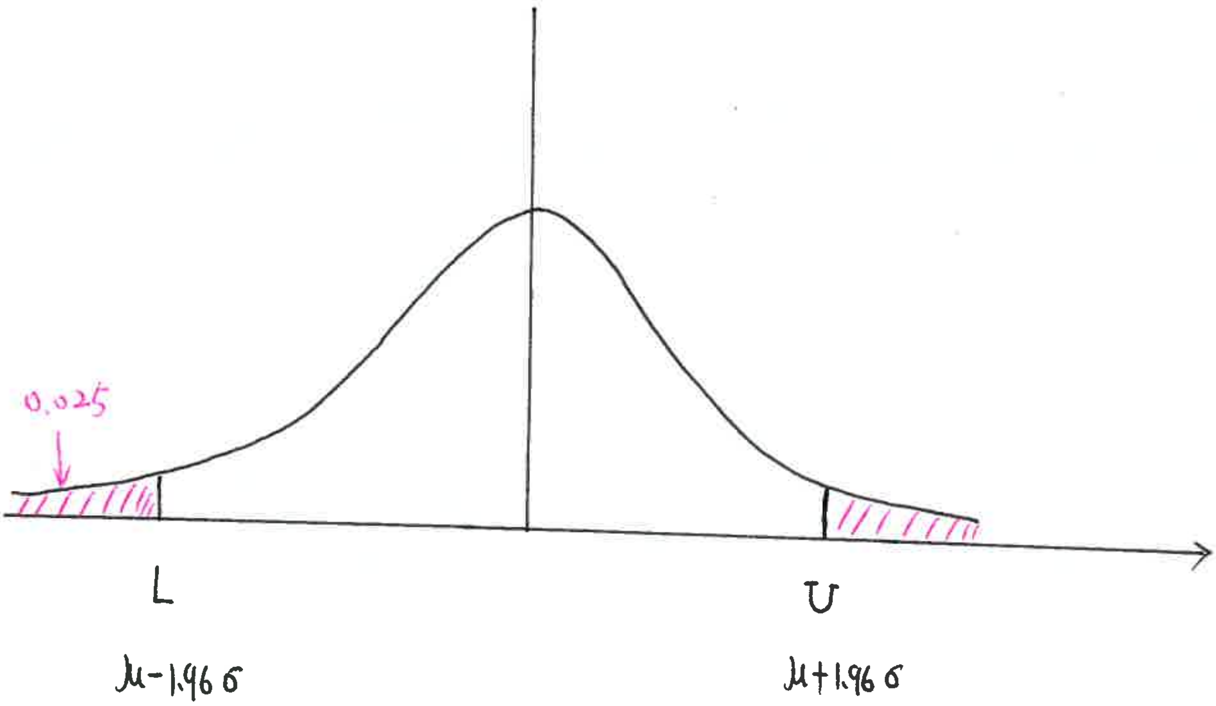
$$= p(1-p)[p + (1-p)]$$

$$= p(1-p)$$

$$\Rightarrow \sigma_X = \sqrt{p(1-p)}$$

For lecture 6, slides 18 & 19

Normal Distribution



For lecture 6, slide 23

$$\left|\frac{(x \otimes y) \otimes z - xyz}{xyz}\right|$$

$$= \left|\frac{xyz\overbrace{(1+\delta_1)(1+\delta_2)}^{1+\theta_2} - xyz}{xyz}\right|$$

$$= \left|\frac{xyz(1+\theta_2) - xyz}{xyz}\right|$$

$$= |(1+\theta_2) - 1|$$

$$= |\theta_2|$$

$$\leq \frac{2u}{1-2u}$$

$$\lesssim 2u$$

For lecture 8, slide 1

$$\left|\frac{fl(x)\oplus fl(y)-(x+y)}{x+y}\right|$$

$$=\left|\frac{[x(1+\delta_1)+y(1+\delta_2)](1+\delta_3)-(x+y)}{x+y}\right|$$

$$=\left|\frac{x\overbrace{(1+\delta_1)(1+\delta_3)}^{1+\theta_2}+y\overbrace{(1+\delta_2)(1+\delta_3)}^{1+\bar{\theta}_2}-x-y}{x+y}\right|$$

$$=\left|\frac{x\theta_2+y\bar{\theta}_2}{x+y}\right|$$

$$\leq \frac{|x|}{|x+y|}|\theta_2|+\frac{|y|}{|x+y|}|\bar{\theta}_2|$$

$$< \frac{|x|}{|x+y|}\frac{2u}{1-2u}+\frac{|y|}{|x+y|}\frac{2u}{1-2u}$$

$$=\frac{|x|+|y|}{|x+y|}\frac{2u}{1-2u}$$

$$\lesssim \frac{|x|+|y|}{|x+y|}2u$$

For lecture 8, slide 3

Quadratic equ

$ax^2+bx+c=0 \quad (a\neq 0).$

$$\Rightarrow \quad x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}.$$

WLOG, assume $b>0$

If $b^2 \gg 4ac$

$$x_1=\frac{-b-\sqrt{b^2-4ac}}{2a}$$

$$x_2=\frac{-b+\sqrt{b^2-4ac}}{2a}$$ not good since $\sqrt{b^2-4ac} \sim b$

$$a(x-x_1)(x-x_2)=ax^2+bx+c=0$$

$$\Rightarrow \quad ax^2-a(x_1+x_2)x+ax_1x_2=ax^2+bx+c=0$$

$$\Rightarrow \quad ax_1x_2=c$$

$$\Rightarrow \quad x_2=\frac{c}{ax_1}$$

For lecture 8, slide 4

$$I_n = \int_0^1 \frac{x^n}{x+2}\,dx$$

$$= \int_0^1 x^{n-1}\left[\frac{x}{x+2}\right]dx$$

$$= \int_0^1 x^{n-1}\left[\frac{x+2-2}{x+2}\right]dx$$

$$= \int_0^1 x^{n-1}\left[1-\frac{2}{x+2}\right]dx$$

$$= \int_0^1 x^{n-1}\,dx - 2\int_0^1 \frac{x^{n-1}}{x+2}\,dx$$

$$= \frac{1}{n} - 2I_{n-1}$$

$< \frac{1}{n}$ since $I_n > 0$ for all n.

$$I_n = \frac{1}{n} - 2I_{n-1}$$

$$\Rightarrow I_{n-1} = \frac{1}{2n} - \frac{1}{2}I_n$$

For lecture 8, slides 8 & 9

Roundoff error

$$RE = \frac{fl(f(x+h)) \ominus fl(f(x))}{h} - \frac{f(x+h)-f(x)}{h}$$

$$= \frac{[f(x+h)(1+\delta_1) - f(x)(1+\delta_2)](1+\delta_3)}{h} - \frac{f(x+h)-f(x)}{h}$$

$$= \frac{f(x+h)\overbrace{(1+\delta_1)(1+\delta_3)}^{1+\theta_2} - f(x)\overbrace{(1+\delta_2)(1+\delta_3)}^{1+\bar{\theta}_2}}{h} - \frac{f(x+h)-f(x)}{h}$$

$$= \frac{f(x+h)(1+\theta_2) - f(x)(1+\bar{\theta}_2)}{h} - \frac{f(x+h)-f(x)}{h}$$

$$= \frac{f(x+h)\theta_2 + f(x)\bar{\theta}_2}{h}$$

$$\Rightarrow |RE| \le \frac{|f(x+h)|+|f(x)|}{h}\,\frac{2u}{1-2u}$$

Use Taylor series, we have

$$f(x+h) = f(x) + h f'(\xi), \quad \xi \in (x, x+h)$$

So

$$|RE| \le \frac{2|f(x)| + h|f'(\xi)|}{h}\,\frac{2u}{1-2u}$$

$$= \frac{2|f(x)|}{1-2u}\,\frac{2u}{h} + |f'(\xi)|\frac{2u}{1-2u}$$

If $f \in C^1$, then $|f'(\xi)| \le M_1$, and $|f(x)| \le M_0$

Hence,

$$|RE| \le \frac{4|f(x)|}{1-2u}\,\frac{u}{h} + \frac{2|f'(\xi)|}{1-2u}\,u$$

$$\le \frac{4M_0}{1-2u}\,\frac{u}{h} + \frac{2M_1}{1-2u}\,u$$

$$\approx 4M_0\frac{u}{h} + 2M_1 u = k_2\frac{u}{h} + k_3 u$$

For lecture 8, slide 11.

The truncation error

Expand $f(x+h)$ in a Taylor series about x.

$$f(x+h) = f(x) + h f'(x) + \frac{1}{2}h^2 f''(\xi) \qquad \xi \in (x, x+h)$$

Then the truncation error is

$$TE = f'(x) - \frac{f(x+h) - f(x)}{h}$$

$$= f'(x) - \frac{f(x) + hf'(x) + \frac{1}{2}h^2 f''(\xi) - f(x)}{h}$$

$$= -\frac{1}{2}h f''(\xi)$$

$$\Rightarrow \quad |TE| = \frac{1}{2}h|f''(\xi)|$$

If $f \in C^2$, so $|f''(\xi)| \le M_2$. Hence

$$|TE| \le \frac{1}{2}hM_2 = k_1 h$$

$$|\text{Total Error}| \le |TE| + |RE| = k_1 h + k_2 \frac{u}{h} + k_3 u$$

We can consider it is a fun of h. Denote $g(h) = k_1 h + k_2 \frac{u}{h} + k_3 u$.

mininum total error $\Rightarrow g'(h) = 0 \Rightarrow k_1 - k_2 \frac{u}{h^2} = 0 \Rightarrow h \sim u^{\frac{1}{2}}$

For lecture 8, slide 11

Example of fixed point iteration

$f(x) = x^3 + 4x^2 - 10 = 0$

$\Rightarrow$ 1. $x + x^3 + 4x^2 - 10 = x$

$\Rightarrow x = x - x^3 - 4x^2 + 10 = g_1(x)$

$\Rightarrow$ 2. $x^2 + 4x - 10/x = 0$

$\Rightarrow x^2 = 10/x - 4x$

$\Rightarrow x = \sqrt{10/x - 4x} = g_2(x)$

$\Rightarrow$ 3. $4x^2 = 10 - x^3$

$\Rightarrow 2x = \sqrt{10 - x^3}$

$\Rightarrow x = \sqrt{10 - x^3}/2 = g_3(x)$

$\Rightarrow$ 4. $x^2(x+4) = 10$

$\Rightarrow x^2 = \frac{10}{x+4}$

$\Rightarrow x = \sqrt{\frac{10}{x+4}} = g_4(x)$

5. $x = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$

check it gives the same f

$\Rightarrow x = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$

$\Rightarrow \frac{x^3 + 4x^2 - 10}{3x^2 + 8x} = 0$

$\Rightarrow x^3 + 4x^2 - 10 = 0$

For Lecture 9, slide 4

Explanation by Taylor Series

$x_{n+1} = g(x_n)$ where $x_n = e_n + x^*$

$\Rightarrow e_{n+1} + x^* = g(x^* + e_n)$

$= g(x^*) + e_n g'(x^*) + O(e_n^2)$

but $x^* = g(x^*)$

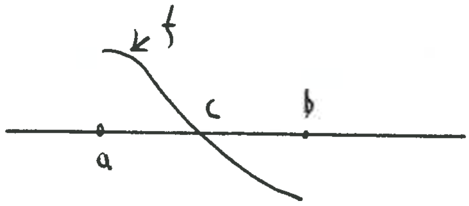
$\Rightarrow e_{n+1} = e_n g'(x^*) + O(e_n^2)$

$\Rightarrow$ linear convergence since $\lim \frac{e_{n+1}}{e_n} = k = g'(x^*)$

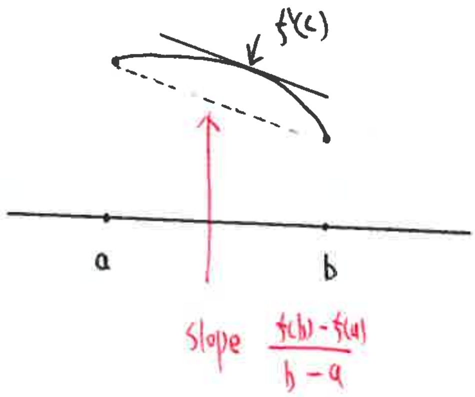
When $g'(x^*) = 0$, then we expect quadratic convergence.

So convergence iff $|g'(x^*)| < 1$.

For lecture 9 slide 6



Intermediate value theorem



Mean value theorem

For lecture 9, slide 8

Fixed point iteration: $x_{n+1} = g(x_n)$

Newton's method is a special fixed point iteration with

$$g(x) = x - \frac{f(x)}{f'(x)}$$

Note that

$$g'(x^*) = 1 - \frac{f'(x^*)^2 - f(x^*) f''(x^*)}{f'(x^*)^2} = 1 - \frac{f'(x^*)^2}{f'(x^*)^2} = 0$$

$\Rightarrow$ quadratic convergence

For lecture 10, slide 2

Explanation by Taylor Series for Newton's method

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

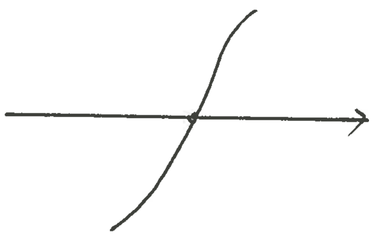
$$x^* + e_{n+1} = x^* + e_n - \frac{f(x^*+e_n)}{f'(x^*+e_n)}$$

$$\Rightarrow e_{n+1} = e_n - \frac{f(x^*+e_n)}{f'(x^*+e_n)} = \frac{e_n f'(x^*+e_n) - f(x^*+e_n)}{f'(x^*+e_n)}$$

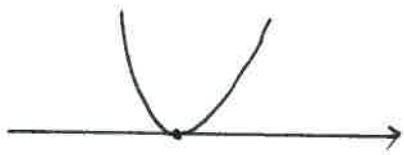
$$= \frac{e_n\left[f'(x^*) + e_n f''(x^*) + \frac{1}{2}e_n^2 f'''(x^*) + O(e_n^3)\right] - \left[\overset{\overset{0}{\|}}{f(x^*)} + e_n f'(x^*) + \frac{1}{2}e_n^2 f''(x^*) + O(e_n^3)\right]}{f'(x^*) + e_n f''(e_n) + O(e_n^2)}$$

$$= \frac{\frac{1}{2}e_n^2 f''(x^*) + O(e_n^3)}{f'(x^*) + e_n f''(x^*) + O(e_n^2)}$$

$$= \begin{cases} k e_n^2 & \text{if } f'(x^*) \neq 0 \quad \text{simple root} \\ \frac{1}{2} e_n & \text{if } f'(x^*) = 0 \quad \text{double root.} \end{cases}$$



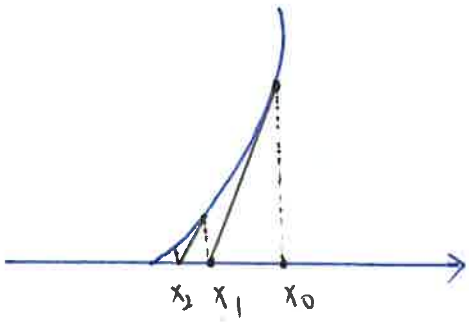
simple root



double root.

For lecture 10, slide 5.

Theorem: if $f'>0$, $f''>0$ (convex), in $[x^*, x]$, then Newton's method converges for all $x_0 \in [x^*, x]$



Similarly, $f'>0$, $f''<0$ in $[x, x^*]$, then NM converges for all $x_0 \in [x, x^*]$

For lecture 10, slide 9

Secant Method

$$x_{n+1} = x_n - f(x_n)\left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right]$$

$$\Rightarrow x^* + e_{n+1} = x^* + e_n - f(x_n)\frac{(x^*+e_n)-(x^*+e_{n-1})}{f(x_n) - f(x_{n-1})}$$

$$\Rightarrow e_{n+1} = e_n - f(x_n)\frac{e_n - e_{n-1}}{f(x_n) - f(x_{n-1})}$$

$$= \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})} = \frac{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}}{f(x_n) - f(x_{n-1})} e_n e_{n-1}$$

$$= \frac{\frac{f(x^*+e_n)}{e_n} - \frac{f(x^*+e_{n-1})}{e_{n-1}}}{f(x^*+e_n) - f(x^*+e_{n-1})} e_n e_{n-1}$$

$$= \frac{\frac{f(x^*) + e_n f'(x^*) + \frac{1}{2}e_n^2 f''(x^*)}{e_n} - \frac{f(x^*) + e_{n-1}f'(x^*) + \frac{1}{2}e_{n-1}^2 f''(x^*)}{e_{n-1}}}{f(x^*) + e_n f'(x^*) - f(x^*) - e_{n-1}f'(x^*)} e_n e_{n-1}$$

$$= \frac{f''(x^*)}{2f'(x^*)} e_n e_{n-1}$$

ignore higher order terms

$$\Rightarrow e_{n+1} \sim C e_n e_{n-1}$$

Note, $e_n \sim C_1 e_{n-1}^p$ and $e_{n+1} \sim C_1 e_n^p = C_1(C_1 e_{n-1}^p)^p \sim \tilde{C} e_{n-1}^{p^2}$

$e_{n+1} \sim C e_n e_{n-1} \sim C C_1 e_{n-1}^p e_{n-1} \sim \tilde{C} e_{n-1}^{p+1}$

$$\Rightarrow p^2 = p+1 \Rightarrow p = \frac{1+\sqrt{5}}{2} \quad \text{golden ratio}$$

For lecture 10, slide 14.

Gaussian Elimination

Ex:

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ 2x_1 + 4x_2 - x_3 &= 2 \\ 3x_1 + 2x_2 + 2x_3 &= 3 \end{aligned}$$

$$[A;b] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & 4 & -1 & 2 \\ 3 & 2 & 2 & 3 \end{array}\right]$$

$R_2 \leftarrow R_2 - \frac{2}{1} R_1 \qquad l_{21} = \frac{2}{1}$

$R_3 \leftarrow R_3 - \frac{3}{1} R_1 \qquad l_{31} = \frac{3}{1}$

> multipliers

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & -1 & 5 & 0 \end{array}\right]$$

$R_3 \leftarrow R_3 - \left(\frac{-1}{2}\right) R_2 \qquad l_{32} = \frac{-1}{2} \rightarrow$ multiplier

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & \frac{11}{2} & 0 \end{array}\right]$$

upper triangular matrix

$$x_3 = \frac{0}{11/2} = 0$$

$$x_2 = \frac{0 - 1 \cdot 0}{2} = 0$$

$$x_1 = \frac{1 - x_2 + x_3}{1} = \frac{1}{1} = 1$$

For lecture 11, slide 6

Extreme 2x2 case

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1+\varepsilon \\ 2 \end{bmatrix} \qquad \text{Exact sol } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$R_2 \leftarrow R_2 - \frac{1}{\varepsilon} R_1$$

$$\begin{bmatrix} \varepsilon & 1 \\ 0 & 1-\frac{1}{\varepsilon} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1+\varepsilon \\ 2-\frac{1}{\varepsilon}(1+\varepsilon) \end{bmatrix}$$ only consider exact arthemetic here.

suppose $\varepsilon < u \simeq 10^{-16}$. On a computer, we will do floating point arithemetic. which implies

$$\begin{bmatrix} \varepsilon & 1 \\ 0 & -\frac{1}{\varepsilon} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{\varepsilon} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since $1+\varepsilon$ rounds to 1 and $1-\frac{1}{\varepsilon}$ rounds to $-\frac{1}{\varepsilon}$

For lecture 11, slide 12.

Extreme 2x2 case with partial pivoting

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1+\varepsilon \\ 2 \end{bmatrix}$$

Partial pivoting

$$\begin{bmatrix} 1 & 1 \\ \varepsilon & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1+\varepsilon \end{bmatrix}$$

$R_2 \leftarrow R_2 - \varepsilon \times R_1$ $\quad l_{21} = \frac{\varepsilon}{1} = \varepsilon$ multiplier

$$\begin{bmatrix} 1 & 1 \\ 0 & 1-\varepsilon \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1+\varepsilon-2\varepsilon \end{bmatrix}$$

using floating point arithmetic

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For lecture 11, slide 14

LU Factorization

Consider

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 4 & -1 \\ 3 & 2 & 2 \end{bmatrix}$$

$R_2 \leftarrow R_2 - \frac{2}{1} R_1 \quad \ell_{21} = 2$

$R_3 \leftarrow R_3 - \frac{3}{1} R_1 \quad \ell_{31} = 3$

> multipliers

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & -1 & 5 \end{bmatrix}$$

$R_3 \leftarrow R_3 - \left(\frac{-1}{2}\right) R_2 \quad \ell_{32} = \frac{-1}{2}$ multiplier

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & \frac{11}{2} \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 4 & -1 \\ 3 & 2 & 2 \end{bmatrix}$$

Define $L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\ell_{21} & 1 & 0 \\ -\ell_{31} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$

$$L_1 A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & -1 & 5 \end{bmatrix}$$

check it.

Define $L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\ell_{32} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix}$

$$L_2 L_1 A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & \frac{11}{2} \end{bmatrix} = U$$

$$\Rightarrow A = \underbrace{L_1^{-1} L_2^{-1}}_{\text{denote it by } L} U = LU$$

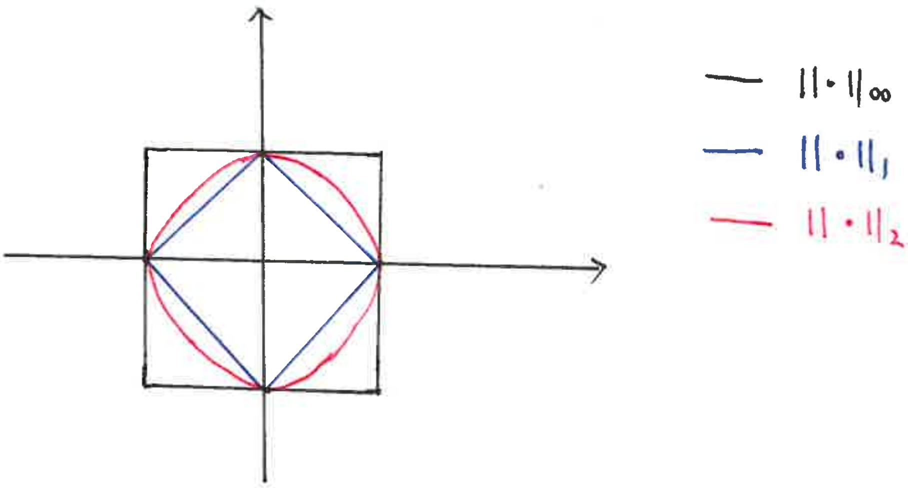
$$L_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

$$L_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \ell_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}$$

$$L_2^{-1} L_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \ell_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}$$

For lecture 12, slide 1

Unit vectors in 1, 2 and ∞ norms

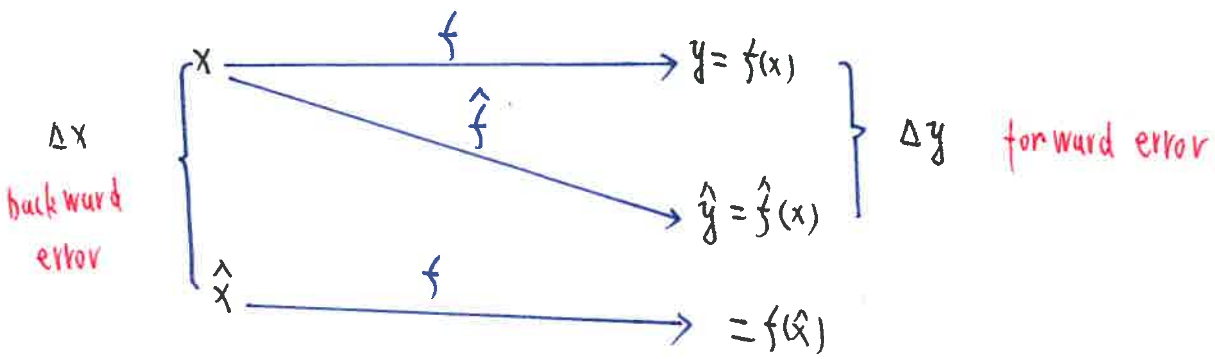


Example of matrix norm

$$A = \begin{bmatrix} 1 & -8 & 6 \\ 2 & 2 & -3 \\ 5 & 1 & 3 \end{bmatrix} \qquad \|A\|_1 = 12 \qquad \|A\|_\infty = 15$$

For lecture 13, slides 5 & 10

Diagram for forward and backward errors



Ask a different question: Is the thing we computer the result of exact solution on a different input?

Approximate cos(1) by two term Taylor series $\hat{f}(x) = 1 - \frac{1}{2}x^2$

$$y = f(1) = \cos(1) \simeq 0.5403$$

$$\hat{y} = \hat{f}(1) = 1 - \frac{1}{2} \cdot 1^2 = 0.5$$

So forward error $= \hat{y} - y \simeq -0.0403$

$$\hat{y} = 0.5 = f(x + \Delta x) = \cos(x + \Delta x)$$

$$\Rightarrow 1 + \Delta x = a\cos(0.5) \simeq 1.0472$$

$$\Rightarrow \Delta x \simeq 0.0472$$

Hence Δy and Δx are similar in size.

For lecture 13, slide 13

Hilbert matrix

$$H_{ij} = \frac{1}{i+j-1} \quad 1 \le i,j \le n.$$

$\text{Cond}(H) \simeq e^{7n/2}$ so rapidly becomes ill-conditioned as n increases

Arise from naive least-squares fitting of polynomials

For lecture 13, slide 21

Example of "nearly singular" matrix

$$A = \begin{bmatrix} 1001 & 1000 \\ 1 & 1 \end{bmatrix} \quad \text{add} \quad \Delta A = \begin{bmatrix} 0 & 0 \\ 0.001 & 0 \end{bmatrix}$$

$\Rightarrow A + \Delta A$ is singular.

$\|\Delta A\|_1 = 0.001 \quad \|A\|_1 = 1002$

So $\dfrac{1}{K_1(A)} \leq \dfrac{\|\Delta A\|_1}{\|A\|_1} = \dfrac{0.001}{1002}$

$\Rightarrow K_1(A) \geq 1.002 \times 10^6$

For lecture 14, slide 1.

Determinant does not help

$$A = \begin{bmatrix} 1001 & 1000 \\ 1 & 1 \end{bmatrix} \qquad \det(A) = 1 \qquad \text{cond}(A) > 10^6$$

$$B = \begin{bmatrix} 0.1 & & & \\ & 0.1 & & \\ & & \ddots & \\ & & & 0.1 \end{bmatrix}_{100 \times 100}$$

$$\det(B) = 0.1^{100} = 10^{-100}$$

$$\text{cond}(B) = \|B\| \, \|B^{-1}\| = 0.1 \cdot 10 = 1$$

For lecture 14, slide 2.

Geometric interpretation

$a_{11}x_1 + a_{12}x_2 = b_1$ presents a line

$a_{21}x_1 + a_{22}x_2 = b_2$ presents another line

ill-conditioned means those two lines are almost parallel

For lecture 14, slide 3.

Monotone norm

monotone norm means $\|A\| = \||A|\|$

norm of A norm of $|A|$

All norms except 2-norm.

For lecture 14, slide 8

Illustration of Backward error for triangular systems

Ex: $\begin{bmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

$$\Rightarrow \hat{x}_1 = b_1 \oslash T_{11} = \frac{b_1}{T_{11}}(1+\delta_1) = \frac{b_1}{\hat{T}_{11}} \quad \text{where } \hat{T}_{11} = \frac{T_{11}}{1+\delta_1}$$

$$T_{21}\hat{x}_1 + T_{22}x_2 = b_2$$

$$\Rightarrow \hat{x}_2 = (b_2 \ominus (T_{21} \otimes \hat{x}_1)) \oslash T_{22}$$

$$= \frac{(b_2 - T_{21}\hat{x}_1(1+\delta_2))(1+\delta_3)}{T_{22}}(1+\delta_4)$$

$$= \frac{b_2 - \hat{T}_{21}\hat{x}_1}{\hat{T}_{22}} \quad \text{where } \hat{T}_{21} = T_{21}(1+\delta_2) \text{ and } \hat{T}_{22} = \frac{T_{22}}{(1+\delta_3)(1+\delta_4)}$$

Similar to your assignment:

$$\frac{1}{\prod_{i=1}^{n}(1+\delta_i)} = 1+\theta_n \quad \text{where } |\theta_n| \le \frac{nu}{1-nu} = \gamma_n$$

$$\Rightarrow |\Delta T_{11}| = |\hat{T}_{11} - T_{11}| = |\theta_1||T_{11}| \le \gamma_1|T_{11}|$$

$$|\Delta T_{21}| = |\hat{T}_{21} - T_{21}| = |\delta_2||T_{21}| \le u|T_{21}| < \gamma_1|T_{21}|$$

$$|\Delta T_{22}| = |\hat{T}_{22} - T_{22}| = |\theta_2||T_{22}| < \gamma_2|T_{22}|$$

Overall, $|\Delta T_{ij}| \le \gamma_2 |T_{ij}|$ for linear 2x2 system

For lecture 14, slide 7

Growth factor

$$\begin{bmatrix} \delta & 1 \\ 1 & 1 \end{bmatrix} \qquad \ell_{21} = \frac{1}{\delta} \text{ multiplier.}$$

$$\to \begin{bmatrix} \delta & 1 \\ 0 & 1-\frac{1}{\delta} \end{bmatrix} \qquad \text{Then } L = \begin{bmatrix} 1 & 0 \\ \ell_{21} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{\delta} & 1 \end{bmatrix}$$

U

Suppose $\delta << 1$. Then

$\|L\|_M = \delta^{-1}$

$\|U\|_M = \delta^{-1} - 1$

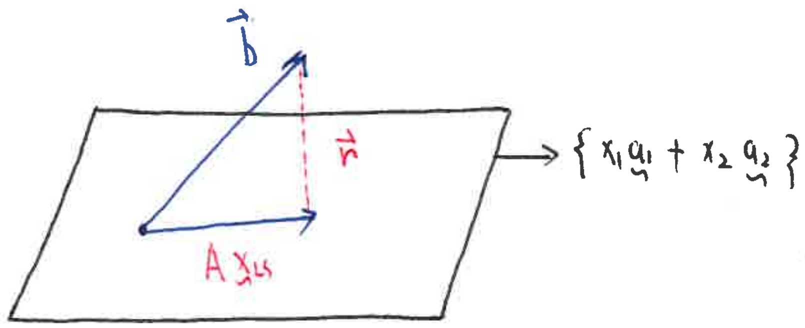
$\|A\|_M = 1$

$$\rho = \frac{\|L\|_M \|U\|_M}{\|A\|_M} = \frac{\delta^{-1}(\delta^{-1}-1)}{1} \sim \delta^{-2}$$

For lecture 14, slide 1

Algebra derivation of linear squares problem

$A = [\underset{\sim}{a_1}, \underset{\sim}{a_2}]$



$\Rightarrow \vec{r} \perp \text{plane} \Rightarrow \vec{r} \perp \underset{\sim}{a_1}$, and $\vec{r} \perp \underset{\sim}{a_2}$

$\Rightarrow A^T \underset{\sim}{r} = \underset{\sim}{0}$

$\Rightarrow A^T A \underset{\sim}{x} = A^T \underset{\sim}{b}$

For lecture 15, slide 9

Proof: $A\underset{\sim}{x} = 0 \Rightarrow x_1 \underset{\sim}{a_1} + x_2 \underset{\sim}{a_2} = \underset{\sim}{0}$

$\Rightarrow x_1 = x_2 = 0$ i.e $\underset{\sim}{x} = 0$ since $\underset{\sim}{a_1}, \underset{\sim}{a_2}$ are linearly independent.

$$A^T A \underset{\sim}{x} = \underset{\sim}{0} \Rightarrow \underset{\sim}{x}^T A^T A \underset{\sim}{x} = 0$$

$$\Rightarrow (A\underset{\sim}{x})^T A \underset{\sim}{x} = 0$$

$$\Rightarrow \|A\underset{\sim}{x}\|_2^2 = 0$$

$$\Rightarrow A\underset{\sim}{x} = \underset{\sim}{0}$$

$$\Rightarrow \underset{\sim}{x} = \underset{\sim}{0}.$$

For lecture 15, slide 10

Example

$$A = \underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}}_{\text{basis of } \mathbb{R}^2 \, (z=0)} \qquad \underset{\sim}{b} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

orthogonal projection of $\underset{\sim}{b}$ onto the column space of A is $\begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$

so expect $A \underset{\sim}{x}_{LS} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$$

$$(A^T A)^{-1} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$$

$$\Rightarrow \underset{\sim}{x}_{LS} = (A^T A)^{-1} A^T \underset{\sim}{b} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

$$= \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 9 \\ 23 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

so $A \underset{\sim}{x}_{LS} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$

For lecture 15, slide 12.

Polynomial Curve fitting

$\phi_1(t) = 1$, $\phi_2(t) = t$, ... $\phi_n(t) = t^{n-1}$

So $(A^TA)_{ij} = \sum_{k=1}^{m} x_k^{i+j-2}$

Then if $x_k = k \cdot \frac{1}{m}$.

We can show

$$(A^TA)_{ij} \sim m H_{ij}$$ Hilbert matrix

$\Rightarrow$ not the right way to do poly curve fitting.

Fix: use orthogonal poly basis.

$$L_n = \frac{1}{2^n n!} \frac{d}{dx^n} (x^2-1)^n$$ Legendre poly

For lecture 15, Slide 14.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1-\eta \end{bmatrix} \qquad u < \eta < \sqrt{u}$$

$$A^TA = \begin{bmatrix} 3 & 3-\eta \\ 3-\eta & 3-2\eta+\underbrace{\eta^2}_{\text{round to } 0} \end{bmatrix} = \begin{bmatrix} 3 & 3-\eta \\ 3-\eta & 3-2\eta \end{bmatrix} + \text{roundoff}$$

and

$$R = \begin{bmatrix} \sqrt{3} & (3-\eta)/\sqrt{3} \\ 0 & 0 \end{bmatrix}$$

Note if $A^T = A$, then $A^TA = A^2$. and $\lambda_{max}(A^2) = \lambda_{max}(A)^2$.

$\Rightarrow \quad \|A\|_2 = |\lambda_{max}(A)|$

For lecture 15, slide 17

Singular Value decomposition

$$\underset{m\times n}{\begin{bmatrix} & \\ & \end{bmatrix}} = \underset{m\times m}{\begin{bmatrix} & \\ & \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ & & 0 & \end{bmatrix}} \underset{n\times n}{\begin{bmatrix} & \\ & \end{bmatrix}}$$

$$\qquad\qquad U \qquad\qquad \Sigma \qquad\qquad V^T$$

U and V are orthogonal, i.e. $U^{-1} = U^T$ and $V^{-1} = V^T$.

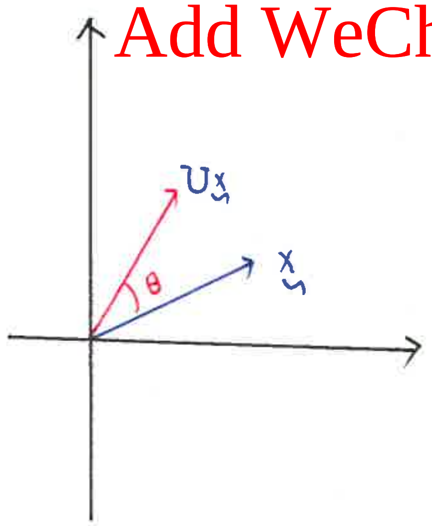
$\{\sigma_1 \geq \sigma_2 \cdots \geq \sigma_n \geq 0\}$ are singular value of A.

$\sigma_n > 0$ if A is full rank.

Example:

$$U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

orthogonal matrix $\longleftrightarrow$ rotation transformation



For lecture 16, slide 2.

Proof of $A^+ = (A^TA)^{-1}A^T = V\Sigma^+U^T$

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T)$$

$$= V\Sigma^T \underline{U^TU}\,\Sigma V^T$$

$$= V\Delta V^T \qquad \text{where } \Delta = \Sigma^T\Sigma = \mathrm{diag}(\sigma_1^2, \dots \sigma_n^2) = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}$$

So

$$(A^TA)^{-1} = (V\Delta V^T)^{-1}$$

$$= (V^T)^{-1}\Delta^{-1}V^{-1}$$

$$= V\Delta^{-1}V^T \qquad \text{where } \Delta^{-1} = \mathrm{diag}(\sigma_1^{-2}, \dots \sigma_n^{-2}) = \begin{bmatrix} 1/\sigma_1^2 & & & \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ & & & 1/\sigma_n^2 \end{bmatrix}$$

Then

$$A^+ = (A^TA)^{-1}A^T$$

$$= V\Delta^{-1}\underline{V^TV}\,\Sigma^TU^T$$

$$= V\underbrace{\Delta^{-1}\Sigma^T}_{\Sigma^+}U^T$$

$$\Sigma^+ = \begin{bmatrix} \sigma_1^{-2} & & & \\ & \sigma_2^{-2} & & \\ & & \ddots & \\ & & & \sigma_n^{-2} \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{bmatrix} = \begin{bmatrix} \sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & 0 \\ & & \ddots\ \sigma_n^{-1} & \end{bmatrix}$$

For lecture 16, slide 4.

Proof of $\|A^+\|_2 = 1/\sigma_n(A)$

$$\|A^+\|_2 = \|V\Sigma^+U^T\|_2$$

$$= \|\Sigma^+\|_2 \qquad \Sigma^+ = \begin{bmatrix} \sigma_1^{-1} & & & \\ & \ddots & & 0 \\ & & \sigma_n^{-1} & \end{bmatrix}$$

$$= \max_{\|x\|_2=1} \|\Sigma^+ x\|_2$$

$$= \max_{\|x\|_2=1} \sqrt{\sigma_1^{-2}x_1^2+\dots+\sigma_n^{-2}x_n^2} \qquad \Sigma^+\begin{bmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} \sigma_1^{-1}x_1 \\ \sigma_2^{-1}x_2 \\ \vdots \\ \sigma_n^{-1}x_n \end{bmatrix}$$

For any $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_m \end{bmatrix}$,

$$\sqrt{\sigma_1^{-2}x_1^2+\dots+\sigma_n^{-2}x_n^2} \le \sigma_n^{-1}\sqrt{x_1^2+\dots+x_n^2}$$

$$\le \sigma_n^{-1}\sqrt{x_1^2+\dots+x_n^2+x_{n+1}^2+\dots+x_m^2}$$

$$= \sigma_n^{-1}$$

If we let $x = \begin{array}{cl} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow \text{nth element} \\ \\ \\ \end{array} \end{array}$ then $\sqrt{\sigma_1^{-2}x_1^2+\dots+\sigma_n^{-2}x_n^2} = \sigma_n^{-1}$

Thus $\|A^+\|_2 = \sigma_n^{-1}(A)$

It implies. $\kappa_2(A) = \|A\|_2\|A^+\|_2 = \dfrac{\sigma_1(A)}{\sigma_n(A)}$

The condition number is the ratio of largest to smallest singular value.

For lecture 16, slide 5.

Proof of $k_2(A^TA) = k_2(A)^2$

$$k_2(A^TA) = \frac{\sigma_1(A^TA)}{\sigma_n(A^TA)}$$

but from earlier.

$$A^TA = V\Delta V^T \quad \text{where } \Delta = \text{diag}(\sigma_1^2 \cdots \sigma_n^2)$$

So $\sigma_1(A^TA) = \sigma_1^2 = \sigma_1(A)^2$.

$\sigma_n(A^TA) = \sigma_n^2 = \sigma_n(A)^2$.

Thus

$$k_2(A^TA) = \frac{\sigma_1(A)^2}{\sigma_n(A)^2} = \left(\frac{\sigma_1(A)}{\sigma_n(A)}\right)^2 = k_2(A)^2.$$

For lecture 16, slide 6.

Gram – Schmidt process

$\{a_1, a_2, \cdots a_n\}$ linearly independent

$q_1 = b_{11} a_1$ $\quad$ $q_1 \cdot q_1 = 1 \Rightarrow b_{11}$

$q_2 = b_{12} q_1 + b_{22} a_2$ $\quad$ $q_1 \cdot q_2 = 0, \; q_2 \cdot q_2 = 1 \Rightarrow b_{12}, b_{22}$

$\vdots$

$q_n = b_{1n} q_1 + b_{2n} q_2 + \cdots + b_{n-1,n} q_{n-1} + b_{nn} a_n$

$q_1 \cdot q_n = 0, \; q_2 \cdot q_n = 0, \ldots, \; q_{n-1} \cdot q_n = 0, \; q_n \cdot q_n = 1 \Rightarrow b_{1n}, b_{2n}, \ldots b_{n-1,n}, b_{nn}$

$\Rightarrow$

$a_1 = r_{11} q_1$

$a_2 = r_{12} q_1 + r_{22} q_2$

$\vdots$

$a_n = r_{1n} q_1 + r_{2n} q_2 + \cdots + r_{n-1,n} q_{n-1} + r_{nn} q_n$

or in matrix form

$$[a_1, a_2, \cdots a_n] = [q_1, q_2 \cdots q_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & & \\ & & & r_{nn} \end{bmatrix}$$

$$A = Q \cdot R$$

For lecture 16, slide 9.

Orthogonal transformation

- $k_2(Q) = 1$

Lemma: $\|Q\|_2 = 1$.

Proof: $\|Q\|_2 = \max\limits_{\|\underset{\sim}{x}\|_2=1} \|Q\underset{\sim}{x}\| = \max\limits_{\|\underset{\sim}{x}\|_2=1} \sqrt{\underset{\sim}{x}^T Q^T Q \underset{\sim}{x}} = \max\limits_{\|\underset{\sim}{x}\|_2=1} \|\underset{\sim}{x}\|_2 = 1$

Using the lemma, we have

$$k_2(Q) = \|Q\|_2 \|Q^{-1}\|_2 = \|Q\|_2 \|Q^T\|_2 = 1$$ since Q^T is also orthogonal.

- $k_2(QA) = k_2(A)$.

$k_2(QA) = \|QA\|_2 \|(QA)^+\|_2$ $\qquad (QA)^+ = A^+ Q^+ = A^+ Q^{-1} = A^+ Q^T$.

$= \|QA\|_2 \|A^+ Q^T\|_2$

$= \|A\|_2 \|A^+\|_2$

$= k_2(A)$

- $\underset{\sim}{x}_{LS}$ satisfies $\min \|b - Ax\|_2$

but on transformation $A\underset{\sim}{x} = \underset{\sim}{b} \longrightarrow QA\underset{\sim}{x} = Q\underset{\sim}{b}$

residual: $\underset{\sim}{r} \longrightarrow Q\underset{\sim}{r}$.

but $\|Q\underset{\sim}{r}\|_2 = \|\underset{\sim}{r}\|_2$

So orthogonal transform doesn't change the residual (2 norm).

$\Rightarrow$ doesn't change the solution.

For lecture 16, slide 11.

Example of Gauss-Newton Method

$y \simeq x_1 e^{x_2 t}$ $\qquad (T_i, Y_i),\ i=1,\cdots m$

$r_i = Y_i - x_1 e^{x_2 T_i}, \quad i=1,\cdots m$

Approximate by a tangent plane about the current guess

$$r_i(\underset{\sim}{x}) \simeq r_i(\underset{\sim}{x}_c) + \frac{\partial r_i}{\partial x_1}(\underset{\sim}{x}_c)(x_1 - x_{1c}) + \frac{\partial r_i}{\partial x_2}(\underset{\sim}{x}_c)(x_2 - x_{2c})$$

$$= r_i(\underset{\sim}{x}_c) - e^{x_{2c}T_i}(x_1 - x_{1c}) - x_{1c} T_i e^{x_{2c}T_i}(x_2 - x_{2c})$$

$$= r_i(\underset{\sim}{x}_c) + J_{i1}(\underset{\sim}{x}_c)\underbrace{(x_1 - x_{1c})}_{s_1} + J_{i2}(\underset{\sim}{x}_c)\underbrace{(x_2 - x_{2c})}_{s_2}$$

$$= r_i(\underset{\sim}{x}_c) + J_{i1}(\underset{\sim}{x}_c)s_1 + J_{i1}(\underset{\sim}{x}_c)s_2 \simeq 0$$

for $i=1,\cdots m$

We get m equs for s_1 & s_2

overdetermined system

$$\begin{bmatrix} J_{11}(\underset{\sim}{x}_c) & J_{12}(\underset{\sim}{x}_c) \\ J_{21}(\underset{\sim}{x}_c) & J_{22}(\underset{\sim}{x}_c) \\ \vdots & \\ J_{m1}(\underset{\sim}{x}_c) & J_{m2}(\underset{\sim}{x}_c) \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} -r_1(\underset{\sim}{x}_c) \\ -r_2(\underset{\sim}{x}_c) \\ \vdots \\ -r_m(\underset{\sim}{x}_c) \end{bmatrix}$$

or

$$J(\underset{\sim}{x}_c)\underset{\sim}{s} = -R(\underset{\sim}{x}_c)$$

Solve it to get $\underset{\sim}{s}$ and update the current guess

$\underset{\sim}{x}_+ = \underset{\sim}{x}_c + \underset{\sim}{s}$ $\qquad$ repeat to convergence

For lecture 17, slide 9