andrewyk.leung@mail.utoronto.ca
https://snoopysnipe.github.io/ta/c37f20/

1. Division $\frac{x}{y}$

computer gives us $fl(fl(x)/fl(y))$

$$= (x(1-\delta_1)/y(1-\delta_2))(1-\delta_3)$$

$$= \frac{x}{y}\frac{(1-\delta_1)(1-\delta_3)}{(1-\delta_2)}$$

$$= \frac{x}{y}\frac{(1-\delta_1-\delta_3+\cancel{\delta_1\delta_3})}{(1-\delta_2)}$$

$$= \frac{x}{y}\frac{(1-\delta_1-\delta_3)}{(1-\delta_2)}\frac{(1+\delta_2)}{(1+\delta_2)}$$

$$= \frac{x}{y}\frac{(1-\delta_1-\delta_3+\delta_2-\cancel{\delta_1\delta_2}-\cancel{\delta_3\delta_2})}{(1-\cancel{\delta_2\delta_2})}$$

$$= \frac{x}{y}(1-(\delta_1+\delta_3-\delta_2))$$

$$= \frac{x}{y}(1-\delta_1)$$

where $|\delta_1| \le 3\varepsilon$

2. a) We have $L_4 P_4 L_3 P_3 L_2 P_2 L_1 P_1 A = U$

We need $PA = LU$ where

$$P = P_4 P_3 P_2 P_1$$

$$L = \tilde{L}_1^{-1} \tilde{L}_2^{-1} \tilde{L}_3^{-1} L_4^{-1}$$

$$L_4 P_4 L_3 P_3 L_2 P_2 L_1 P_1 A$$

$$= L_4 P_4 L_3 P_3 L_2 (P_2 L_1 P_2)(P_2 P_1) A$$

$$= L_4 P_4 L_3 (P_3 L_2 P_3)(P_3 P_2 L_1 P_2 P_3)(P_3 P_2 P_1) A$$

since $P_i^{-1} = P_i$
$P_i P_i = I$

$$= L_4 (P_4 L_3 P_4)(P_4 P_3 L_2 P_3 P_4)(P_4 P_3 P_2 L_1 P_2 P_3 P_4)(P_4 P_3 P_2 P_1) A$$

let $\tilde{L}_3 = (P_4 L_3 P_4)$ $\tilde{L}_2 = P_4 P_3 L_2 P_3 P_4$

$\tilde{L}_1 = P_4 P_3 P_2 L_1 P_2 P_3 P_4$ ad $P = P_4 P_3 P_2 P_1$

We have $L_4 \tilde{L}_3 \tilde{L}_2 \tilde{L}_1 PA = U$

$$\Rightarrow PA = \tilde{L}_1^{-1} \tilde{L}_2^{-1} \tilde{L}_3^{-1} L_4^{-1} U$$

let $L = \tilde{L}_1^{-1} \tilde{L}_2^{-1} \tilde{L}_3^{-1} L_4^{-1}$

$$\Rightarrow PA = LU$$

$$\tilde{L}_1^{-1} = (P_4 P_3 P_2 L_1 P_2 P_3 P_4)^{-1}$$

$= P_4 P_3 P_2 L_1^{-1} P_2 P_3 P_4$ since $P_i = P_i^{-1}$

b) $\tilde{L}_1^{-1} = P_4 P_3 P_2 L_1^{-1} P_2 P_3 P_4$

$$L_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 & 0 \\ m_{31} & 0 & 1 & 0 & 0 \\ m_{41} & 0 & 0 & 1 & 0 \\ m_{51} & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$L_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -m_{21} & 1 & 0 & 0 & 0 \\ -m_{31} & 0 & 1 & 0 & 0 \\ -m_{41} & 0 & 0 & 1 & 0 \\ -m_{51} & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P_2 L_1^{-1} P_2 = \begin{bmatrix} 1 & & & & \\ -m_{51} & 1 & & & \\ -m_{31} & & 1 & & \\ -m_{41} & & & 1 & \\ -m_{21} & & & & 1 \end{bmatrix}$$

$$P_3 P_2 L_1^{-1} P_2 P_3 = \begin{bmatrix} 1 & & & & \\ -m_{51} & 1 & & & \\ & & 1 & & \\ -m_{41} & & & 1 & \\ -m_{31} & & & & 1 \\ -m_{21} & & & & \end{bmatrix}$$

$$P_4 P_3 P_2 L_1^{-1} P_2 P_3 P_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -m_{51} & 1 & 0 & 0 & 0 \\ -m_{41} & 0 & 1 & 0 & 0 \\ -m_{21} & 0 & 0 & 1 & 0 \\ -m_{31} & 0 & 0 & 0 & 1 \end{bmatrix} = \tilde{L}_1^{-1}$$

3, a) $P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $P_1 A = \begin{bmatrix} 4 & 4 & 4 \\ 3 & 5 & 9 \\ 1 & 5 & 5 \end{bmatrix}$

$L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{4} & 1 & 0 \\ -\frac{1}{4} & 0 & 1 \end{bmatrix}$ $L_1 P_1 A = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 2 & 6 \\ 0 & 4 & 4 \end{bmatrix}$

$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ $P_2 L_1 P_1 A = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 2 & 6 \end{bmatrix}$

$L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}$ $L_2 P_2 L_1 P_1 A = \begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}$

$= U$

We have $L_2 P_2 L_1 P_1 A = U$

$\iff L_2 P_2 L_1 P_2 P_2 P_1 A = U$

$\iff L_2 \tilde{L}_1 P_2 P_1 A = U$

$\iff P_2 P_1 A = \tilde{L}_1^{-1} L_2^{-1} U$

$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix}$ $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

b)

We have $Ax=b \iff PAx=Pb \iff LUx=b$

Let $Ux=d$

Solve $Ld=Pb$

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 24 \\ 26 \\ 40 \end{bmatrix}$$

$d_1 = 24$

$d_2 = 20$

$d_3 = 12$

$$\begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 24 \\ 20 \\ 12 \end{bmatrix}$$

$x_3 = 3$

$x_2 = 2$

$x_1 = 1$

$$\therefore x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

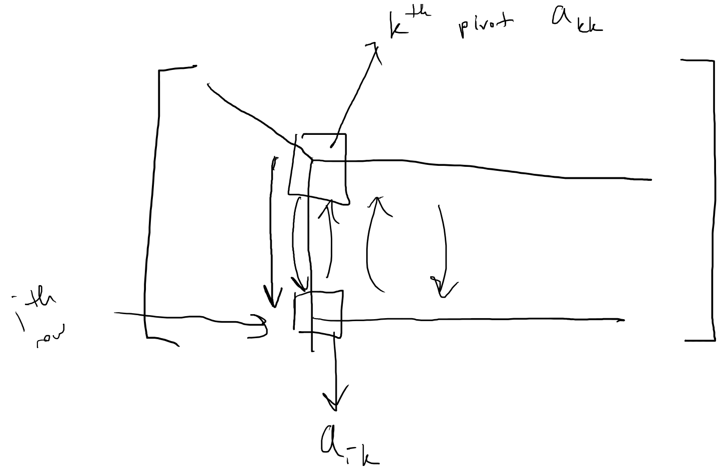
c) When you have several linear systems to solve which differ only in the right hand side, you can save much computation by factoring A once and then using the factorization to solve each of the systems. For example, this situation arises in the iterative improvement algorithm.

4. Remember: Premultiplying swaps rows, postmultiplying swaps columns

$$
\begin{aligned}
Ax = b &\iff PAx = Pb \\
&\iff PA\underbrace{QQ}_{I}x = Pb \\
&\iff PAQy = Pb \qquad \text{where } Qx = y \\
&\iff LUy = Pb \\
&\iff Ld = Pb \qquad \text{where } Uy = d
\end{aligned}
$$

Solve $Ld = Pb$ for d
Solve $Uy = d$ for y
Compute $x = Qy$

5.



- if dont pivot and pivot element is very small, multipliers become very large in this stage since pivot element is in the denominator of each multiplier
- if the multipliers are large, each element in lower right submatrix could get significantly amplified in this stage. Not good since round off error is directly proportional to the largest elements that occur during factorization
- if pivot, multipliers are guaranteed to be ≤ 1 in magnitude in this stage and elements in lower right submatrix will be damped. Thus, growth of round ~~off~~ error is reduced.

Assignment

1. $b^2 \approx 4ac$, $ac > 0$

$$x^2 - y^2 \Rightarrow (x-y)(x+y)$$

$$\frac{-b \pm \sqrt{(b-\sqrt{4ac})(b+\sqrt{4ac})}}{2a}$$

ex: $\sqrt{1+x} - \sqrt{1-x}$

cancellation when $x \to 0$

$$\sqrt{1+x} - \sqrt{1-x} \times \frac{\sqrt{1+x} + \sqrt{1-x}}{\sqrt{1+x} + \sqrt{1-x}}$$

$$= \frac{1+x-(1-x)}{\sqrt{1+x} + \sqrt{1-y}}$$

$$= \frac{2x}{\sqrt{1+x} + \sqrt{1-x}}$$

Assignment 5. b) $$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right|$$

$$\leq \sum_{i=1}^{m} \sum_{j=1}^{n} \left(|a_{ij}| |x_j| \right)$$

$$= \sum_{j=1}^{n} \left(|x_j| \sum_{i=1}^{m} |a_{ij}| \right)$$

$$\leq \left(\max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}| \right) \left(\sum_{j=1}^{n} |x_j| \right)$$

$$= \|A\|_1 \, \|x\|_1$$

$$\|A\|_1 \leq \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|$$

Since $\|x\|=1$ from definition, then

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}| = \sum_{i=1}^{m} |a_{ik}|$$

where the k^{th} column of A has max abs col sum

and $\|Ax\|_1 \leq \|A\|_1 \|x\|_1 = \sum_{i=1}^{m} |a_{ik}|$

Note that if we choose $x = e_k = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow k^{th}$ row ; k^{th} unit vector

note this is normalized

then $\|Ax\|_1$ is maximized

$$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right| = \sum_{i=1}^{m} |a_{ik}|$$

$$= \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$$

Therefore, $x = e_k$ is the vector among all the normalized vectors that maximizes $\|Ax\|$ as required in the definition which results in $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$

ex: $e^x - 1$ cancellation when $x \to 0$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$$e^x - 1 = x + \frac{x^2}{2!} + \frac{x^3}{3!} = \sum_{n=1}^{\infty} \frac{x^n}{n!}$$