

11/9/24 midterm rewrites

Caitlyn Jones

$$2. a) \quad A = \begin{bmatrix} 90 & 60 & 120 \\ 60 & 30 & 90 \\ 30 & 90 & 120 \end{bmatrix} = I$$

$$Ax = b \Rightarrow \begin{bmatrix} 90 & 60 & 120 \\ 60 & 30 & 90 \\ 30 & 60 & 120 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 390 \\ 240 \\ 330 \end{bmatrix}$$

$$b) \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & -1/4 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 90 & 60 & 120 \\ 0 & 40 & 80 \\ 0 & 0 & 30 \end{bmatrix}$$

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

$$Ly = Pb$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & -1/4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 390 \\ 240 \\ 330 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & -1/4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 390 \\ 330 \\ 240 \end{bmatrix}$$

$$y_1 = 390$$

$$1/3 y_1 + y_2 = 330$$

$$1/3(390) + y_2 = 330$$

$$130 + y_2 = 330$$

$$y_2 = 200$$

$$2/3 y_1 - 1/4 y_2 + y_3 = 240$$

$$2/3(390) - 1/4(200) + y_3 = 240$$

$$260 - 50 + y_3 = 240$$

$$210 + y_3 = 240$$

$$y_3 = 30$$



$$y = \begin{bmatrix} 390 \\ 200 \\ 30 \end{bmatrix}$$

$$Ax = y$$

$$\begin{bmatrix} 90 & 100 & 120 \\ 0 & 40 & 80 \\ 0 & 0 & 30 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 390 \\ 200 \\ 30 \end{bmatrix}$$

$$30x_3 = 30$$

$$x_3 = 1$$

$$40x_2 + 80x_3 = 200$$

$$40x_2 + 80(1) = 200$$

$$40x_2 + 80 = 200$$

$$40x_2 = 120$$

$$x_2 = 3$$

$$90x_1 + 100x_2 + 120x_3 = 390$$

$$90x_1 + 100(3) + 120(1) = 390$$

$$90x_1 + 180 + 120 = 390$$

$$90x_1 + 300 = 390$$

$$90x_1 = 90$$

$$x_1 = 1$$

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



$$c) \frac{\|\Delta X\|}{\|X\|} = \frac{\|\hat{X} - X\|}{\|X\|}$$

$$\hat{X} - X = \begin{bmatrix} 1.0405678 - 1.0000000 \\ 0.0001345 - 0.0000000 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0405678 \\ 0.0001345 \end{bmatrix}$$

$$\|\Delta X\| = \sqrt{(0.0405678)^2 + (0.0001345)^2}$$

$$\approx 0.0400000$$

$$\frac{\|\Delta X\|}{\|X\|} = \frac{0.0400000}{1}$$

$$\frac{\|\Delta X\|}{\|X\|} \leq \text{cond}(M) \cdot \epsilon_{\text{mach}}$$

$$0.0400000 \leq \text{cond}(M) \cdot 10^{-8}$$

$$\frac{4.00 \times 10^{-2}}{10^{-8}} \leq \text{cond}(M)$$

$$4.00 \times 10^6 \leq \text{cond}(M)$$



3. a) it is a bad idea because it can lead to the loss of precision. this happens because the system might not have enough storage to represent the arithmetic between two drastically different floating point numbers with drastically different magnitudes, which results in the loss of many floating point numbers. for example,

$$a = 1.000000000001 \text{ and } b = 1.000000000000$$

$$a - b = .000000000001 \text{ or } 1.0 \times 10^{-12}$$

some systems may write the answer as 0 instead of  $1.0 \times 10^{-12}$  due to the fact that these numbers are represented by floating point numbers with such different magnitudes.

b) as  $N \rightarrow \infty$ ,  $S_N$  diverges or gets closer and closer to infinity. however, since the numbers being added to the final sum get smaller and smaller, we will end up having to add numbers with drastically different magnitudes in IEEE, which will result in the loss of precision. So, instead of the series diverging to  $\infty$ , the series will stop adding some values of  $S/N$  once their numbers in floating point systems stop contributing to the overall sum due to their miniscule magnitude compared to the converging sum.



$$c) S_N \approx \log(N) + \lambda, \text{ where } \lambda = 0.57721$$

$$\frac{\frac{1}{N}}{S_N} < \epsilon_{\text{main}}$$

$$\Rightarrow \frac{\frac{1}{N}}{\log(N) + \lambda} < 10^{-16}$$

$$\Rightarrow \frac{1}{N} < 10^{-16} \cdot (\log(N) + \lambda)$$

$$\Rightarrow N > \frac{1}{10^{-16} \cdot (\log(N) + \lambda)}$$

$$\Rightarrow N > \frac{10^{16}}{\log(N) + \lambda}$$

We can estimate that  $N = 10^{16}$ , so

$$\begin{aligned} \log(N) + \lambda &= \log(10^{16}) + 0.57721 \\ &= 16.57721 \end{aligned}$$

$$\text{so, } \frac{1}{10^{16}} < \epsilon_{\text{main}}$$

$$\Rightarrow \frac{1}{10^{17}} < \epsilon_{\text{main}}$$

so, the sum will stop changing around  $N = 16$



$$4. a) Ac = y \quad A = \begin{bmatrix} 1 & 20 \\ 1 & 30 \\ 1 & 10 \end{bmatrix} \quad c = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \quad y = \begin{bmatrix} 35 \\ 70 \\ 30 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 20 \\ 1 & 30 \\ 1 & 10 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 35 \\ 70 \\ 30 \end{bmatrix}$$

$$b) c = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

c) 205,000 absences

$$d) y_i = a_0 + a_1 x_i + a_2 x_i^2$$

$$A = \begin{bmatrix} 1 & 20 & 400 \\ 1 & 30 & 900 \\ 1 & 10 & 100 \end{bmatrix} \quad c = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \quad y = \begin{bmatrix} 35 \\ 70 \\ 30 \end{bmatrix}$$

$$c = \begin{bmatrix} 55 \\ -4 \\ 0.15 \end{bmatrix}$$

1,155,000 absences

the prediction gets worse. because the data points are so close to each other, it is hard to accurately predict large scale absences. the linear model is more accurate for a real life example, as it demonstrates a steady and highly proportional relationship between the flu cases and employee absences, instead of the quadratic model, which causes extreme predictions due to the squared variable.



5. a)  $\hat{r}_+ = 9.9 \times 10^4$ ,  $\hat{r}_- = 1.0000003385357559 \times 10^{-5}$

relative error  $r_+ = 0.0$

relative error  $r_- = 3.384 \times 10^{-7}$

b)  $r_-$  has a significantly greater relative error compared to  $r_+$ ; however, the error is still relatively small. This is due to catastrophic cancellation in the floating point arithmetic for the values of  $r_-$ . When we compute  $r_-$ , we are subtracting nearly equal numbers. Because of this, the floating point representation of this subtraction results in the loss of precision in the computed root.


$$c) r_+ \cdot r_- = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \cdot \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

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

$$= \frac{b^2 - (b^2 - 4ac)}{4a^2}$$

$$= \frac{0 - 4ac}{4a^2}$$

$$= \frac{c}{a}$$

therefore  $r_+ \cdot r_- = \frac{c}{a}$  



$$\hat{r}_- = \frac{c}{a \cdot \hat{r}_+}$$

$$\hat{r}_- = 1.0000000001 \times 10^{-5}$$

relative error  $r_- = 0.0$

$$d) r_- = \frac{b}{2} (-1 - \sqrt{1 - 4c/b^2})$$

$$\approx \frac{b}{2} (-1 - (1 - 2c/b^2))$$

for  $1 - \frac{2c}{b^2}$  to give no correct digits,

then

$$\frac{2c}{b^2} < \epsilon_{\text{mach}}$$