COL 726 (HW1)

Name- Gulshan Jangid
Entry - 2014CS50285

Page No.
Date:

Ans 1. a) $\begin{bmatrix} 1 & 2.01 \\ 1.01 & 2.03 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.01 \\ 1.02 \end{bmatrix}$

$\Rightarrow x_1 + 2.01 x_2 = 1.01$ & $1.01x_1 + 2.03x_2 = 1.02$

on solving these two equations:

$x_2 = 1$ & $x_1 = -1$

taking $b + \Delta b$ we get

$x_1 + 2.01 x_2 = 1.01$ & $1.01x_1 + 2.03x_2 = 1.0201$

on solving these two we get

$x_2 = 0$ & $x_1 = 1.01$

b) We have $f(b) = [-1, 1]^T$

& $f(b+\Delta b) = [1.01, 0]^T$

Using ∞ form

$$K = \frac{\|\Delta y\|_\infty / \|y\|_\infty}{\|\Delta x\|_\infty / \|x\|_\infty}$$

here $\Delta y = f(b+\Delta b) - f(b) = [2.01, -1]^T$

& $\Delta x = [0, 0.0001]^T$

$$K = \frac{2.01/1}{0.0001/1.02} = 2.0502 \times 10^4$$

which is very large & should be because small perturbation in b changes f(b) to a large extent

Ans 2 a) Two situations in which loss of accuracy can occur are

1) when b^2 is very close to $4ac$
2) when $+\sqrt{b^2-4ac}$ is very close to $-b$, this can happen when $b>>ac$

The first condition in which b^2 is close to $4ac$ can affect both roots

The second condition will only affect the root $\frac{-b+\sqrt{b^2-4ac}}{2a}$ because of subtraction of two very close numbers

b) $$\frac{-b\pm\sqrt{b^2-4ac}}{2a} \times \frac{-b\mp\sqrt{b^2-4ac}}{-b\mp\sqrt{b^2-4ac}}$$

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

Since we want to avoid the 2^{nd} cancellation error we should use $\frac{-b-\sqrt{b^2-4ac}}{2a}$ for first root & $\frac{2c}{-b-\sqrt{b^2-4ac}}$ for 2^{nd} root

Ans 3 $f(x) = (x-1)*(x-1)$

$$\hat{f}(x) = (fl(x) \ominus 1) \circledast (fl(x) \ominus 1)$$

$$\hat{f}(x) = (x(1+\epsilon_1)-1)*(x(1+\epsilon_1)-1)*(1+\epsilon_2)^2*(1+\epsilon_3)$$

$$\Rightarrow |\hat{f}(x) - f(x)| = |(x(1+\epsilon_1)-1)^2(1+\epsilon_2)^2(1+\epsilon_3) - (x-1)^2|$$

Since $|\epsilon_1|, |\epsilon_2|, |\epsilon_3| \leq \epsilon_m$

& from the above equation we can see

$$|\hat{f}(x) - f(x)| \leq |(x(1+\epsilon_m)-1)^2(1+\epsilon_m)^3 - (x-1)^2|$$

on opening & re-arranging terms

$$|\hat{f}(x) - f(x)| \leq |\epsilon_m(5x^2-8x+3) + 3\epsilon_m^2((x-1)^2 + x^2\epsilon_m^2 + 2x\epsilon_m(x-1) + 2x(x-1) + \epsilon_m x^2)|$$

$$|\hat{f}(x) - f(x)| < \epsilon_m|5x^2-8x+3| + O(\epsilon_m^2 \cdot c_2)$$

Ans 4 a) when we multiply a matrix A with vector x we get linear combination of columns of A, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$$Ax = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

Take $C = AB$

$$\begin{bmatrix} c_1 | c_2 | \cdots \end{bmatrix} = \begin{bmatrix} A \end{bmatrix}\begin{bmatrix} b_1 | b_2 | \cdots \end{bmatrix}$$

From columnwise interpretation

$$c_k = Ab_k$$

thus each column of C is linear combination of columns of A, thus $Rank(C) \leq Rank(A)$

Take $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ & $x^T = [x_1, x_2, \dots x_n]$

$$z = x^T B = x_1 b_1 + x_2 b_2 + \dots + x_n b_n$$

here z is linear combination of rows of B

C = AB can be shown as

$$\begin{bmatrix} R_{c_1} \\ R_{c_2} \\ \vdots \\ R_{c_n} \end{bmatrix} = \begin{bmatrix} R_{a_1} \\ R_{a_2} \\ \vdots \\ R_{a_n} \end{bmatrix} \begin{bmatrix} B \end{bmatrix}$$

$$R_{c_k} = R_{a_k} \cdot B$$

thus each row of C is linear combination of rows of B so $Rank(C) \leq Rank(B)$

Hence $Rank(C) \leq \min(Rank(A), Rank(B))$

also $Rank(A) \leq \min(m, n)$
$\leq n$

& $Rank B \leq \min(n, p)$
$\leq n$

So $Rank(C) = Rank(AB) \leq n$

Ans 4(b) Given A is full rank

Suppose A maps non-zero vectors to zero vectors

i.e. $Ax = 0$

let $A = \begin{bmatrix} a_1 | a_2 \cdots a_n \end{bmatrix}$ & $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

then $a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$

which says that a linear combination of columns of A is zero, but we know that A is full rank $\Rightarrow \nexists$ no linear combination of columns of A that ~~add up~~ is zero (with $x \neq 0$)

Thus its a contradiction.

Hence A maps non zero vectors to non zero vectors

For linear independent vectors:

Suppose $V_1, V_2, \ldots, V_k$ are linearly independent

thus $c_1 V_1 + c_2 V_2 + \cdots + c_k V_k = 0 \Rightarrow c_1 = c_2 = \cdots c_k = 0$

multiplying by A we get

$$c_1 A V_1 + c_2 A V_2 + \cdots + c_k A V_k = 0 \Rightarrow c_1 = c_2 = \cdots = c_k = 0$$

which says

$AV_1, AV_2, \ldots, AV_k$ are linearly ind.

Ans 5(a) $\|x\| = \|x+y-y\|$

using triangle inequality

$$\|x\| \le \|x+y\| + \|-y\|$$

Since $\|-y\| = \|y\|$

$$\|x\| \le \|x+y\| + \|y\|$$

Gives $\|x+y\| \ge \|x\| - \|y\|$

operating similarly with $\|y\|$ we get

$$\|y\| \le \|y+x\| + \|x\|$$

$$\Rightarrow \|x+y\| \ge \|y\| - \|x\|$$

Since we are finding lower bound

~~so $\|x+y\| \ge \max(\|x\|+y\|,$~~

so $\|x+y\| \ge \max\left(\|x\| - \|y\|, \|y\| - \|x\|\right)$

Ans 6 (b) Our dataset is

$[9.99999997, 9.99999966, 9.99999972]$

~~me~~ on hand calculation

mean $(\mu) = 9.99999978\dot{3}3\ldots$

variance $= 1.803 \times 10^{-14}$

Using functions var1 & var2

$var1 = 4.263 \times 10^{-14}$

$var2 = 1.802 \times 10^{-14}$

Error in var2 $\approxeq |(1.803 - 1.802) \times 10^{-14}|$

$\approxeq 0.001 \times 10^{-14} \approxeq 10^{-17}$

Error in var1 $\approxeq |(4.263 - 1.803) \times 10^{-14}|$

$\approxeq 2.46 \times 10^{-14}$

Clearly error in var1 $\approxeq 10^{3} \times$ (error in var2).

Ans (6 C) We can use Kahan's Summation method where we keep a seperate running compensation (a variable to accumulate small errors.

when n is large then calculating sum~~ma~~ by adding summands ~~can~~ ~~induc~~ will accumulate more & more error