SC ASSIGNMENT - 1

Name : Meet Popat
Roll No. : 2020320

Q7.

a. Output of the part a is:

0.5822073316515288
0.5797135815734098
0.5788814056433012
0.5784651440685238
0.5782153315683285
0.5780487667534508
0.5779297805478292
0.5778405346932214
0.577771117576444
0.5777155815682065
0.5776701414855578
0.5776322736978301
0.5776002309764809
0.5775727652416682
0.5775489611978291
0.5775281323494514
0.5775097537135299
0.5774934169591495
0.5774787997122512
0.5774656440682016
0.577453741243179
0.5774429204111771
0.5774330404528918
0.5774239837672805
0.5774156515681996
0.5774079602664202
0.5774008386555254
0.5773942257008384
0.5773880687857824
0.5773823223089227
0.5773769465525795
0.5773719067634975
0.5773671724007521
0.5773627165162711
0.5773585152416505
0.5773545473603523
0.5773507939494777
0.5773472380778735
0.5773438645508655

0.5773406596931707
0.5773376111636459
0.5773347077964406
0.577331939464333
0.5773292969607322
0.5773267718973916
0.5773243566154296
0.5773220441077846
0.5773198279512748
0.5773177022470506
0.577315661568166

b. Output of part b :
0.5772197901404903
0.5772167013748222
0.5772161263242399
0.5772159246680912
0.5772158312352449
0.577215780449559
0.5772157498141715
0.5772157299243794
0.5772157162847442
0.5772157065265553
0.5772156993055031
0.5772156938126143
0.577215689537403
0.5772156861448545
0.5772156834077089
0.5772156811674058
0.5772156793105871
0.5772156777544755
0.5772156764374801
0.5772156753129938
0.5772156743452568
0.577215673506438
0.5772156727746101
0.5772156721323229
0.5772156715655328
0.5772156710628664
0.5772156706149998
0.5772156702142466
0.5772156698542288
0.5772156695296005
0.5772156692358834

0.5772156689692576
0.5772156687264989
0.5772156685048309
0.5772156683019034
0.5772156681156311
0.577215667944273
0.5772156677862554
0.5772156676402354
0.5772156675050191
0.5772156673795781
0.5772156672629993
0.5772156671544533
0.5772156670532187
0.5772156669586632
0.5772156668702006
0.5772156667873283
0.5772156667095789
0.5772156666365351
0.5772156665678327

Q4)

a. Output of snippet :
   0.5
   0.25
   0.125
   0.0625
   0.03125
   0.015625
   0.0078125
   0.00390625
   ….
   ….
   ….
   4e-323
   2e-323
   1e-323
   5e-324
   0.0

   Explanation : **(Underflow Limit)**
   This code snippet first sets the variable "a" to equal 1 before repeatedly halving it and reporting the result. The smallest number > 0 that IEEE double precision has is the 2nd last integer that it prints. Additionally known as UFL (Underflow limit). As it dives by two more, it prints 0 when it reaches the UFL i.e it starts from
   1.000..
   0.10000...
   0.01000...
   And so on till the first one comes on the last possible digit in IEEE.

b. Output of the snippet :
   1.1102230246251565e-16

   Explanation:**(Machine Epsilon)**
   The above value is equal to half of the IEEE double precision machine epsilon Em.
   It's because the answer should still be 1, even if we must print eps at the end, which according to the code is the value that must be added to 1. We repeatedly cut the eps in half to discover this. The loop ends when eps equals Em/2 because the shortest leap between two successive numbers is Em (epsilon machine). Since 1. + Em/2 ==== at this point, only 1. will result from rounding.

c. Output of snippet :
   2

4
8
16
32
64
.
.
.
.
2.247116418577895e+307
4.49423283715579e+307
8.98846567431158e+307
inf

Explanation: **(Overflow)**
This code sample first sets the variable "a" to equal 1 before repeatedly doubling it and printing it. Up until the value of a<inf in IEEE double point precision. As 'a' is the special number known as the overflow limit or OFL, which is the greatest number that can be represented in IEEE, it leaves the loop when 'a' becomes >= inf. this is due to the fact that the 2nd last printed number times 2 >= inf, making an equal to inf and breaking the loop.

Screenshot of the output:

In [1]:
```
a = 1
while a != 0:
    a /= 2
    print(a)
```

```
0.5
0.25
0.125
0.0625
0.03125
0.015625
0.0078125
0.00390625
0.001953125
0.0009765625
0.00048828125
0.000244140625
0.0001220703125
6.103515625e-05
3.0517578125e-05
1.52587890625e-05
7.62939453125e-06
3.814697265625e-06
1.9073486328125e-06
```

In [2]:
```
a = 1; eps = 1; b = a + eps
while a != b:
    eps /= 2
    b = a + eps
print(eps)
```

```
1.1102230246251565e-16
```

In [*]:
```
from numpy import inf
a=1
while a != inf:
    a *= 2
    print(a)
```

```
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
```

SC- Assignment 1

MEET POPAT
2020320

Q:1). Linear equation is $ax = b$ where $a, x, b \in \mathbb{R}$

Let x be the true soln to $ax = b$ &
x_0 be the approximate soln to $ax = b$

$\therefore$ Absolute forward error = | True Value - Approximate value |
$= |(x - x_0)|$ —— (1)

For backward error $y = ax - b$
or $f(\hat{y}) = a\hat{x}$
$\therefore$ for $x_0 \Rightarrow \hat{y} = ax_0$
for $x \Rightarrow y = ax$.

$\therefore$ Absolute Backward error $= |y - \hat{y}|$
$= |ax - ax_0|$
$= |a(x - x_0)|$ —— (2)

Similarly,
Relative Forward error $= \dfrac{|(x - x_0)|}{|x|}$ —— (3)

Relative Backward error $= \dfrac{|y - \hat{y}|}{y}$
$= \dfrac{|a(x - x_0)|}{|ax|}$
$= \dfrac{|(x - x_0)|}{|x|}$ —— (4)

Q1.....

Absolute condition number $(\hat{K})$ = $\dfrac{\text{Absolute Forward error}}{\text{Absolute Backward error}}$

$$= \frac{|x - x_0|}{|a||x - x_0|}$$

$$\hat{K} = 1/|a|$$

Relative condition number (K) = $\dfrac{\text{Relative Forward error}}{\text{Relative Backward error}}$

$$= \frac{|x - x_0|/|x|}{|a(x - x_0)|/|ax|}$$

$$= \frac{\cancel{|x - x_0|}\,\cancel{|a|}\,\cancel{|x|}}{|a|\,\cancel{|x - x_0|}\,\cancel{|x|}}$$

$$= 1.$$

$\therefore$ Absolute condition number is $\dfrac{1}{|a|}$.

Relative condition number is $\underline{1}$.

Q:2). We have

Absolute condition number $(\hat{K}) = \frac{|f(x+\delta x) - f(x)|}{|x+\delta x - x|} = \frac{|\delta f|}{|\delta x|}$

Absol Frwd Error

Abs Backwd Error $= \frac{|\delta f|}{|\delta x|}$

Relative condition number $(K) = \left|\frac{x f'(x)}{f(x)}\right|$

a]. $(x-1)^\alpha$

$\therefore f(x) = (x-1)^\alpha$

$\hat{K} = \left|\frac{\delta (x-1)^\alpha}{\delta x}\right| = |\alpha (x-1)^{\alpha-1}|$ (Abs. condⁿ no.)

& $K = \left|\frac{x f'(x)}{f(x)}\right| = \left|\frac{x\, \alpha (x-1)^{\alpha-1}}{(x-1)^\alpha}\right|$

$= \left|\frac{\alpha x}{(x-1)}\right|$ (Rel. condⁿ no)

$\therefore$ The absolute condⁿ number is $(\alpha (x-1)^{\alpha-1})$ and will be large for larger value of x & with dependence on α

$\therefore$ for 1). $\alpha < 0$ & $x \to 1$

2). $0 < \alpha < 1$ & $x \to 1$

3). $\alpha > 1$ & $x \to \infty$ or $x \to -\infty$.

The relative condⁿ number is $\left|\frac{\alpha x}{(x-1)}\right|$ & will be larger for $x \to 1$ & $\alpha \neq 0$

b]. $\ln(x)$

* ∴ Abs. condition number (K). $= \left| \frac{d\ln(x)}{dx} \right| = \frac{1}{|x|}$

* Rel. condition number $= \left| \frac{x\, f'(x)}{f(x)} \right| = \left| \frac{x \frac{1}{x}}{x \times \ln(x)} \right| = \frac{1}{|\ln x|}$

* ∴ The abs condition number is $\frac{1}{|x|}$ & is larger for: $\underline{\underline{x \to 0}}$

• The relative cond^n no is $1/|\ln x|$ & is larger for $\underline{\underline{x \to 1}}$

c]. $x^{-1} e^x$

* Abs. condition number $= \left| \frac{d(x^{-1}e^x)}{dx} \right| = \left| \frac{e^x}{x} - \frac{e^x}{x^2} \right|$

* Relative condition number $= \left| \frac{x f'(x)}{f(x)} \right| = \left| \frac{x \frac{e^x}{x}\left(1 - \frac{1}{x}\right)}{e^x/x} \right|$

$= \underline{|x-1|}$

* The abs condition number is $\left| e^x [1 - 1/x] \right|$ & is larger for $\underline{x \to 0}$ & $\underline{x \to \infty}$

• The relative condition number is $|x-1|$ & is larger for $\underline{\underline{x \to \infty}}$ & $\underline{\underline{x \to -\infty}}$

$x = \quad \Rightarrow x$ tends to 0

d). $\frac{1}{(1+x^{-1})} = \frac{1}{(1+1/x)} = \frac{x}{(x+1)}$

* Absolute condⁿ number $= \left| \frac{d\left(\frac{x}{x+1}\right)}{dx} \right| = \left| \frac{1}{(x+1)^2} \right|$

* Relative condⁿ number $= \left| \frac{x f'(x)}{f(x)} \right| = \left| \frac{x \times 1 \times (1+x)}{(x+1)^2 \times x} \right| = \left| \frac{1}{(x+1)} \right|$

* ∴ The absoute condition number is $\left| \frac{1}{(1+x)^2} \right|$ is large for $\underline{x \to -1}$

* The relative condition number is $\left| \frac{1}{(1+x)} \right|$ & is large for $\underline{\underline{x \to -1}}$

Q:3).

a] Given: $\exists$ a function $x(\varepsilon)$ s.t.

$$f(x(\varepsilon)) + \varepsilon p(x(\varepsilon)) = 0 \quad \text{with } x(0) = x^*$$

$\therefore$ Differentiating w.r.t ε

$$f'(x(\varepsilon)).x'(\varepsilon) + (p(x(\varepsilon)) + \varepsilon p'(x(\varepsilon)) . \varepsilon x'(\varepsilon)) = 0$$

Now, $x'(\varepsilon).(f'(x(\varepsilon)) + \varepsilon p'(x(\varepsilon))) + p(x(\varepsilon)) = 0$

$$\Rightarrow x'(\varepsilon) = \frac{-p(x(\varepsilon))}{f'(x(\varepsilon)) + \varepsilon p'(x(\varepsilon))} = \frac{dx}{d\varepsilon}$$

$\therefore$ At $\varepsilon = 0$. to find $dx/d\varepsilon$

so, $$\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0} = \frac{-p(x(0))}{f'(x(0)) + 0 . p'(x(0))}$$

$$\therefore \boxed{\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0} = \frac{-p(x^*)}{f'(x^*)}}$$ as $x(0) = x^*$

Hence proved

Q:3 b] Wilkinson polynomial.

$$f(x) = (x-1)(x-2)(x-3)\ldots\ldots(x-20)$$

$$= a_0 + a_1 x + a_2 x^2 + \ldots + a_{20} x^{20}$$

Given that expressed coeff. inaccurately $= a_{19}$

so, $\boxed{p(x) = x^{19}}$

$x^* = j$

Using the differential result from part (a).

Now, $$\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0} = \frac{-p(x^*)}{f'(x^*)}$$

$\therefore f(x) = (x-1)(x-2)(x-3)\ldots\ldots(x-20)$

$$\begin{aligned} f'(x) = {} & (x-2)(x-3)\ldots\ldots(x-20) + \\ & + (x-1)(x-3)\ldots\ldots(x-20) + \\ & + (x-1)(x-2)(x-4)\ldots\ldots(x-20) + \\ & + \\ & \vdots \\ & (x-1)(x-2)\ldots\ldots(x-18)(x-20) + \\ & + (x-1)(x-2)\ldots\ldots(x-19) \end{aligned}$$

Now, $f'(x^*)$, where $x^* = j$ & $(1 \le j \le 20)$

$\Rightarrow f'(x^*) = (j-1)(j-2)\ldots\ldots(j-k)$ (All other terms are 0 $(k \neq j)$)

So, $$f'(x^*) = m \prod_{k=1,\, k\neq j}^{20} (j-k)$$

So, $$\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0,\, x^*=j} = \frac{-j^{19}}{\prod_{k=1,\, k\neq j}^{20} j-k}$$

$$\Rightarrow \boxed{\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0,\, x^*=j} = -\prod_{\substack{k=1 \\ (k\neq j)}}^{\infty} \frac{j}{j-k}}$$ Hence proved.

Q:3)

c]. For $x^* = 1$

$$\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0,\, x^*=1}$$

$$= -\prod_{k\neq 1} \frac{1}{1-k}, \quad k = 2, 3, \ldots 20$$

$$= \prod_{k\neq 1} \frac{1}{k-1}$$

$$= \frac{1}{(1)(2)(3)\ldots(19)} = \frac{1}{19!}$$

so, $\boxed{\frac{dx}{d\varepsilon} = \frac{1}{19!}}$ → for $x^* = 1$ & $\varepsilon = 0$.

↓

This is small ($<<<<1$)

For, $x^* = 20$

$$\left.\frac{dx}{d\varepsilon}\right|_{\varepsilon=0,\, x^*=20} = -\prod_{k\neq 20} \frac{20}{(20-k)}$$

$(k = 1, 2, \ldots \infty)$

$$= \frac{-(20)^{19}}{-(19)(18)\ldots(1)}$$

$$= \frac{(20)^{19}}{19!} >>> 1$$

Thus root $x^* = 1$ is more stable than $x^* = 20$

Because small change ε leads to small change in $x^* = 1$ but a huge difference in $x^* = 20$.

∴ Root is stable at $x^* = 1$ & unstable at $x^* = 20$

Q: 6]

a] $Ax = B$ where $B \in \mathbb{R}^m$

$\|x\|_A = \|Ax\|$

$\text{Rank}(A) = n$

So, $\|Ax\| = \|B\| = \|x\|_A$

i) First property of norm

$\|B\| \geq 0$

$\Rightarrow \|Ax\| \geq 0$

$\Rightarrow \|x\|_A \geq 0$

if ~~columns~~ $B = 0$ then $Ax = B$ have

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

where (a_i are columns of A)

since $\text{rank}(A) = n$.

then all the a_i's are independent.

$\therefore$ only trivial soln that exists is

$x_1 = x_2 = x_3 = \cdots = x_n = 0$

$\therefore \|x_A\| = 0 \Rightarrow x = 0$

ii) Homogeneous property.

$\|\alpha x\|_A = \|A(\alpha x)\|$

$= |\alpha| \|Ax\|$ (as α is const

So, $\boxed{\|\alpha x_A\| = |\alpha| \|x\|_A}$

Hence proved

iii). Let $Ax_1 = B_1$ and $Ax_2 = B_2$
where $x_1, x_2 \in \mathbb{R}^n$

By Triangle inequality.

$$\|B_1\| + \|B_2\| \geq \|B_1 + B_2\|$$

$$\Rightarrow \quad \|Ax_1\| + \|Ax_2\| \geq \|Ax_1 + Ax_2\|$$

$$\Rightarrow \quad \|x_1\|_A + \|x_2\|_A \geq \|x_1 + x_2\|_A$$

Hence D inequality also holds for x_1, x_2

$\therefore$ All properties of a norm holds.
Thus $\|x\|_A = \|Ax\|$ is a vector norm on $\mathbb{R}^n$.

Q:5) Relative condition number

$(\hat{K}) = \frac{|x|}{|\sin(x)\cos(x)|}$

To Error is made while rounding it to 16 decimal places

∴ while computing

$i = 0, \ldots 20$ the $(\pi - \pi')$ may increase

as $\pi' = |\pi/4 + 2\pi * 10^i - (\pi'/4 + 2\pi' * 10^i)|$

$= ((\pi - \pi')(1/4 + 2 * 10^i)$

As x increases decreases more it will affect the result $\tan(x)$.

Q:6).

b). $\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$ (Frobenius norm)

$A \in \mathbb{R}^{m \times n}$

Property 1). $\Rightarrow$ Since $|a_{ij}|^2 \geq 0$

$$\Rightarrow \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \geq 0.$$

$\Rightarrow \|A\|_F \geq 0.$ Hence proved

Property 2). $\|A\|_F = 0$

$$\Rightarrow \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2} = 0$$

This is true only if all $a_{ij} = 0$.

So, $\|A\|_F = 0$ i.e $a_{ij} = 0 \ \forall$ $i = 1 \ldots m$, $j = 1 \ldots n$

$\therefore$ only if A is 0 matrix

Hence proved

Property 3). $\|\alpha A\|_F$

$$\Rightarrow \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |\alpha a_{ij}|^2 \right)^{1/2}$$

$$\Rightarrow \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |\alpha|^2 |a_{ij}|^2}$$

$$= \alpha \sqrt{\sum_{i=1} \sum_{j=1} |a_{ij}|^2}$$

$$= \alpha \|A\|_F$$

Hence proved

Thus $\|A\|_F$ follows all 3 properties of norms.