Quiz # 2: Optimization in Neural Networks and Newton's Method

Question 1

Using Newton's method, find an approximation recursive formula for $\sqrt{2}$. To help you, remember that $\sqrt{\mathbf{2}}$**.** is the positive solution for $x^2 - 2$, so you can use

$f(x) = x^2 - 2.$

Options :-

- $x_{k+1} = x_k - \frac{2x_k}{x_k^2 - 2}$
- $x_{k+1} = \frac{x_k^2 - 2}{2x_k}$
- $x_{k+1} = \frac{2x_k}{x_k^2 - 2}$
- $x_{k+1} = x_k - \frac{x_k^2 - 2}{2x_k}$

$f(x) = x^2 - 2$

NIM formula is:-

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \longrightarrow (1)$$

$$f(x_k) = x_k^2 - 2$$

$$f'(x_k) = 2x_k$$

$\therefore$ (1) becomes:-

$$\boxed{x_{k+1} = x_k - \frac{x_k^2 - 2}{2x_k}}$$

Question 2

Regarding the previous question, suppose you don't know any approximation for $\sqrt{2}$ and only that it is a positive real number such that $x^2 \ = 2$. Which value from the list below will result in the fastest convergence?

a):- 4

b):- 3

c):- 2

d):- The initial value does not impact in the Newton's method convergence.

Answer:- c

We know that $\sqrt{2}$ is a number between 1 and 2, so 2 is the closest value in this list of options, therefore is the value that will converge faster!

Question 3

Let's continue investigating the method we are developing to compute the $\sqrt{2}$. Remember that we used the fact that $\sqrt{2}$ is one of the roots of x^2 - 2. What would happen if we have chosen a negative value as initial point?

a):- The algorithm would not converge.
b):- The algorithm would converge to $\sqrt{2}$
c):- **The algorithm would converge to the negative root of** $x^2 - 2$
d):- **The algorithm would converge to** 0.

Answer:- d

Question 4

Did you know that it is possible to calculate the *reciprocal* of any number *without performing division?* (The reciprocal of a non-zero real number a is $1/a$).

Setting a non-zero real number a, use the function $\boldsymbol{f(x) = a - \frac{1}{x}} = a - x^{-1}$ to find such formula. This method was in fact used in older IBM computers to implement division in hardware! So, the iteration formula to find the reciprocal of a, in this case, is:

Options:-

- (selected) $x_{k+1} = 2x_k - ax_k^2$
- $x_{k+1} = 2x_k + ax_k^2$
- $x_{k+1} = 2x_k - x_k^2$
- $x_{k+1} = x_k - ax_k^2$

$$f(x) = a - x^{-1}$$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \rightarrow (1)$$

$$f'x = \frac{1}{x^2}$$

$$\Rightarrow f'(x_k) = 1/x_k^2$$

$\therefore$ (1) becomes,

$$x_{k+1} = x_k - \frac{(a - x_k^{-1})}{\frac{1}{x_k^2}}$$

$$x_{k+1} = x_k - \frac{\left(\frac{a x_k - 1}{x_k}\right)}{1/x_k^2}$$

$$= x_k - x_k\,(a x_k - 1)$$

$$= x_k - a x_k^2 + x_k.$$

$$\boxed{x_{k+1} = 2x_k - a x_k^2}$$

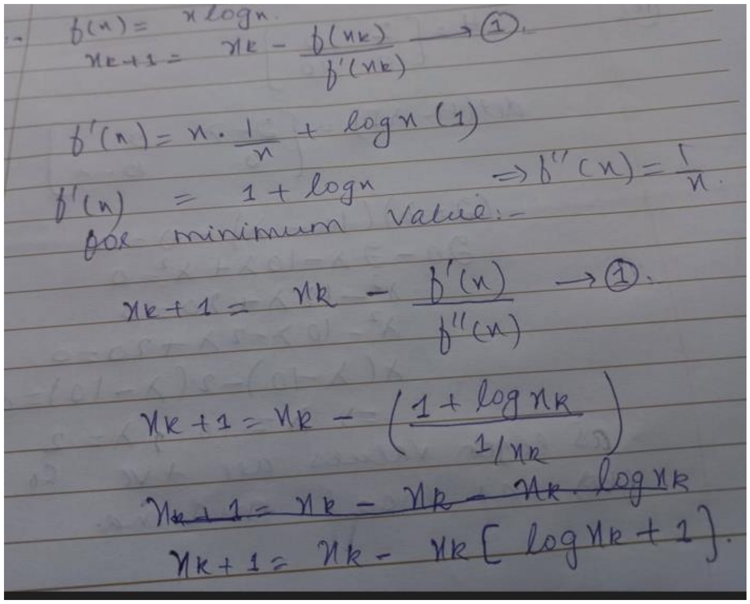
Question 5

Suppose we want to find the minimum value (suppose we already know that the minimum exists and is unique) of $f(x) = x \log x$, where $x \in (0, +\infty)$. Using Newton's method, what recursion formula we must use?

Hint :- $f(x) = x \log x,\ f'(x) = \log x + 1, f''(x) = 1/x$

Options:-

- ○ $x_{k+1} = x_k - \frac{x_k \log(x_k)}{\log(x_k)+1}$
- ○ $x_{k+1} = x_k - x_k^2 \log(x_k)$
- ○ $x_{k+1} = x_k - \log(x_k)$
- ◉ $x_{k+1} = x_k - x_k\left(\log(x_k) + 1\right)$



$f(x) = x \log x$

$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \rightarrow (1)$

$f'(x) = x \cdot \frac{1}{x} + \log x \, (1)$

$f'(x) = 1 + \log x \Rightarrow f''(x) = \frac{1}{x}$

For minimum value:-

$x_{k+1} = x_k - \frac{f'(x)}{f''(x)} \rightarrow (1)$

$x_{k+1} = x_k - \left(\frac{1 + \log x_k}{1/x_k}\right)$

$x_{k+1} = x_k - x_k - x_k \log x_k$

$x_{k+1} = x_k - x_k[\log x_k + 1]$

Question 6

Regarding the *Second Derivative Test* to decide whether a point with $f'(x)=0$ is a local minimum or local maximum, check all that apply.

Options: -

- [ ] If $f''(x) < 0$ then x is a local minimum.
- [x] If $f''(x) > 0$ then x is a local minimum.

  Correct

  Correct! If $f'(x) = 0$ and $f''(x) < 0$ then x is a local maximum!
- [ ] If $f''(x) = 0$ then x is an inflection point.
- [x] If $f''(x) = 0$ then the test is inconclusive.

  Correct

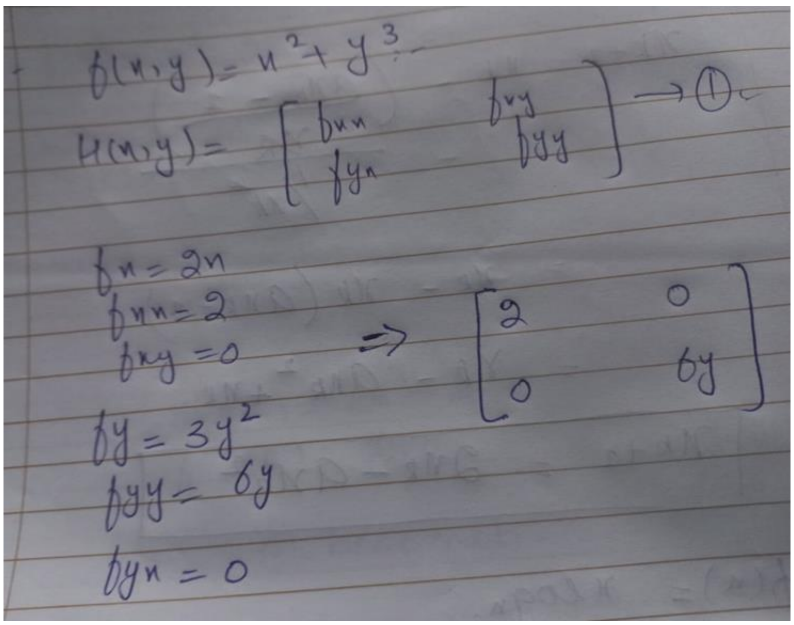
  Correct! If $f'(x) = f''(x) = 0$, then the test is inconclusive!

Question 7

Let $f(x, y) = x^2 + y^3$, then the Hessian matrix, $H(x,y)$ is:

Options:-

- ○ $$H(x,y) = \begin{bmatrix} 2x & 3y^2 \\ 3y^2 & 2x \end{bmatrix}$$
- ◉ $$H(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 6y \end{bmatrix}$$
- ○ $$H(x,y) = \begin{bmatrix} 0 & 2 \\ 6y & 0 \end{bmatrix}$$
- ○ $$H(x,y) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

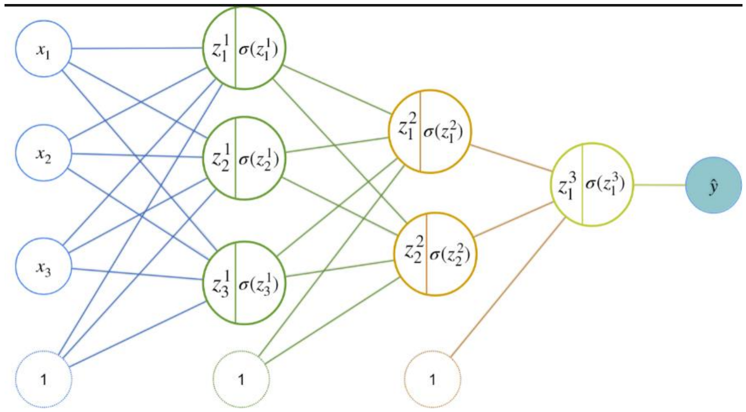


Question 8

How many parameters has a Neural Network with:

- **Input layer of size 3**
- **One hidden layer with 3 neurons**
- **One hidden layer with 2 neurons**
- **Output layer with size 1**

An image is provided below:



a):- 11

b):- 8

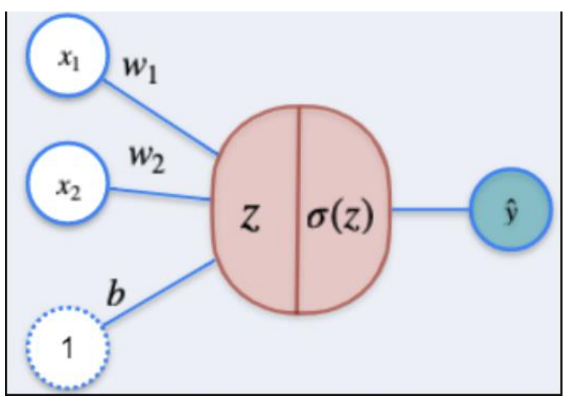
c):- 23

d):- 3

Answer:- c

Question 9

Given the following Single Layer Perceptron with Sigmoid function as activation function, and log-loss as Loss Function (L), the value for $\partial L/\partial w1$ is:



Options:-

- $-(y-\hat{y})$
- $-(y-\hat{y})x_1$
- $-(y-\hat{y})x_2$
- 1

Question 10:-

Suppose you have a function $f(x,y)$ **with** $\nabla f(x_0,y_0)=(0,0)$ **and such that**

$$H(x_0,y_0)=\begin{bmatrix} 2 & 0 \\ 0 & 10 \end{bmatrix}$$

Then the point (x_0, y_0) **is a:**

a):- local maximum

b):- local minimum

c):- saddle point

d):- we cant infer anything with given info.

Answer:- b

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 10 \end{bmatrix}$$

$$\det(H - \lambda I) = \begin{bmatrix} 2-\lambda & 0 \\ 0 & 10-\lambda \end{bmatrix}$$

$$(2-\lambda)(10-\lambda) - 0 = 0$$

$$20 - 2\lambda - 10\lambda + \lambda^2 = 0$$

$$\lambda^2 - 12\lambda + 20 = 0$$

$$\lambda^2 - 10\lambda - 2\lambda + 20 = 0$$

$$\lambda(\lambda - 10) - 2(\lambda - 10) = 0$$

$$\Rightarrow \lambda = 10,\ \lambda = 2$$

as eigen values are +ve so

$f(x)$ has local minima.