1) Numerical Derivative :-

a) $f(x,y) = x^2 + y^2$

$\frac{\partial f}{\partial x} = 2x$; $\frac{\partial f}{\partial y} = 2y$ [Analytical Derivative]

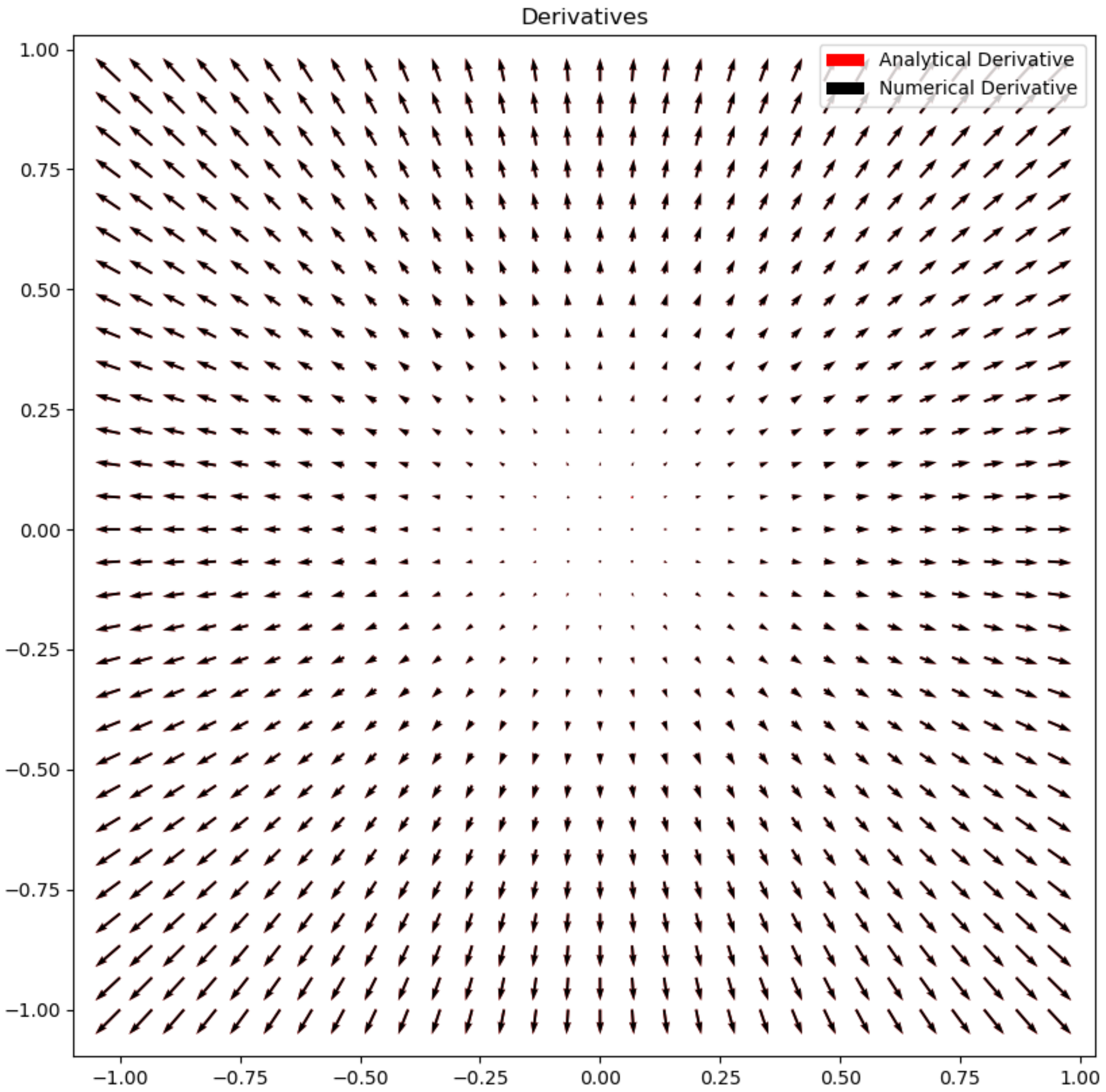
Numerical Derivative :- The numerical partial Derivative can be found in the submitted code.

b) $f(x,y,z) = \sin x + xz$

$\frac{\partial f}{\partial x} = \cos x + z$; $\frac{\partial f}{\partial y} = 0$; $\frac{\partial f}{\partial z} = x$

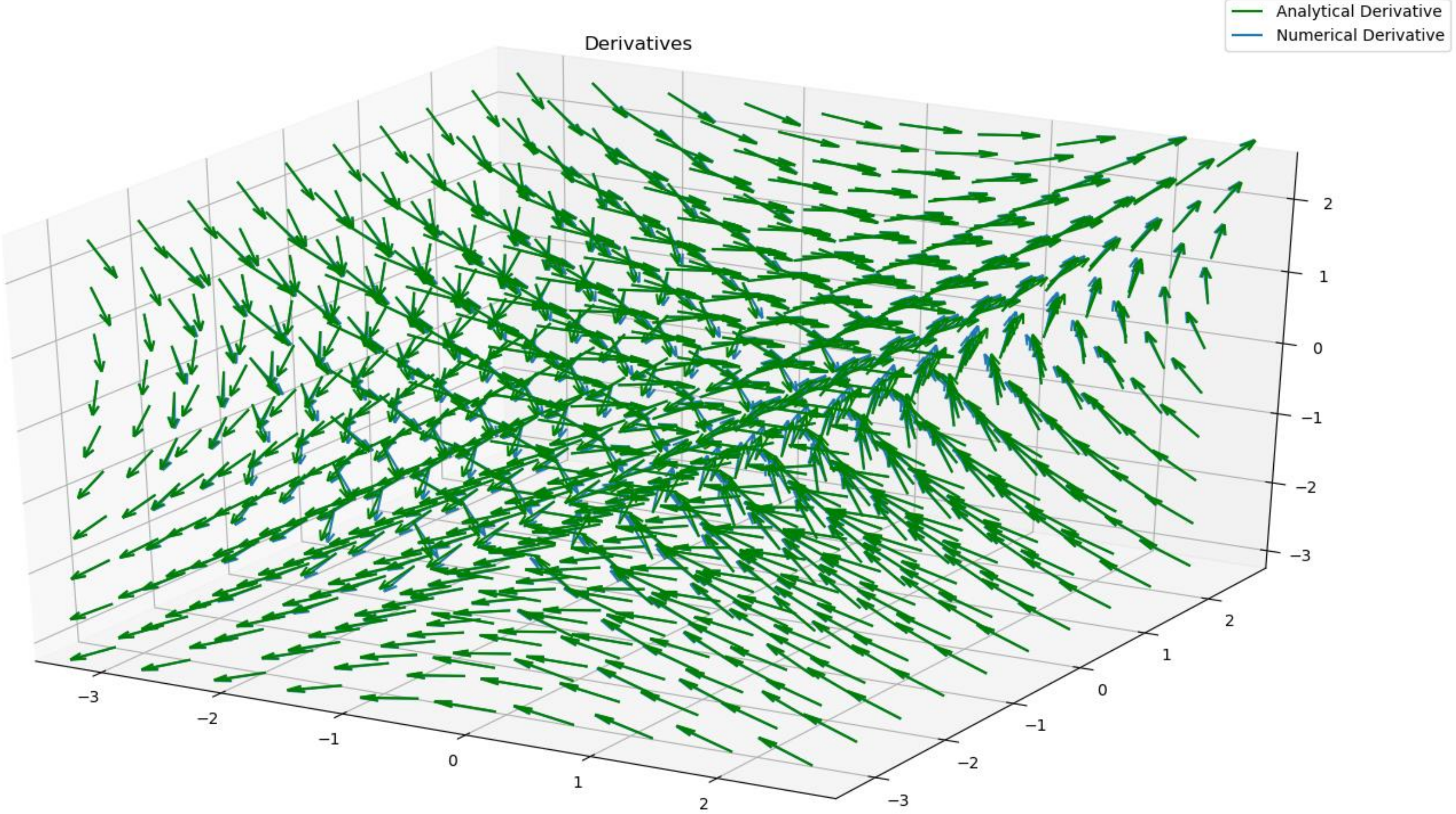
Numerical Derivative can be found in the submitted code.

1a



The following is the quiver plot demonstrates that both numerical and actual partial derivatives are the same and their quiver plots overlap

1b



The following is the quiver plot demonstrates that both numerical and actual partial derivatives are the same and their quiver plots overlap

Hessian

a) $f(x,y) = x^2 + y^3$

$$\frac{\partial f}{\partial x} = 2x \quad ; \quad \frac{\partial f}{\partial y} = 3y^2$$

$$\frac{\partial^2 f}{\partial x^2} = 2 \quad ; \quad \frac{\partial^2 f}{\partial y^2} = 6y \quad ; \quad \frac{\partial^2 f}{\partial x \partial y} = 0$$

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 6y \end{bmatrix}$$

at point $(1,2)$ we get $= \begin{bmatrix} 2 & 0 \\ 0 & 12 \end{bmatrix}$

Numerical result :-

$$= \begin{bmatrix} 2.000 & 0 \\ 0 & 12.006 \end{bmatrix}$$

b) $f(x,y) = \dfrac{1}{1+e^{3x+2y}}$

$$\frac{\partial f}{\partial x} = \frac{-3e^{3x+2y}}{(1+e^{3x+2y})^2} \quad ; \quad \frac{\partial f}{\partial y} = \frac{-2e^{3x+2y}}{(1+e^{3x+2y})^2}$$

$$\frac{-1}{3}\frac{\partial f}{\partial x} = \frac{e^{3x+2y}}{(1+e^{3x+2y})^2}$$

$$-\frac{1}{3}\frac{\partial^2 f}{\partial x^2} = \frac{e^{3x+2y}(3)(1+e^{3x+2y})^2 - \left[2(1+e^{3x+2y})e^{3x+2y}(3)\, e^{3x+2y}\right]}{(1+e^{3x+2y})^4}$$

$$= \frac{e^{3x+2y}(3)}{(1+e^{3x+2y})^2} - \frac{6(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$= \frac{3e^{3x+2y}(1+e^{3x+2y}) - 6(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$= \frac{3e^{3x+2y} + 3(e^{3x+2y})^2 - 6(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$-\frac{1}{3}\frac{\partial^2 f}{\partial x^2} = \frac{3e^{3x+2y} - 3(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$\frac{\partial f}{\partial x^2} = \frac{9(e^{3x+2y})^2 - 9e^{3x+2y}}{(1+e^{3x+2y})^3} \rightarrow 1$$

$$\frac{\partial f}{\partial y} = \frac{-2e^{3x+2y}}{(1+e^{3x+2y})^2} \Rightarrow -\frac{1}{2}\frac{\partial f}{\partial y} = \frac{e^{3x+2y}}{(1+e^{3x+2y})^2}$$

$$-\frac{1}{2}\frac{\partial^2 f}{\partial y^2} = \frac{e^{3x+2y}(2)(1+e^{3x+2y})^2 - \left[2(1+e^{3x+2y})e^{3x+2y}(2)\cdot e^{3x+2y}\right]}{(1+e^{3x+2y})^4}$$

$$= \frac{2e^{3x+2y}}{(1+e^{3x+2y})^2} - \frac{4(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$= \frac{2e^{3x+2y}(1+e^{3x+2y}) - 4(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$-\frac{1}{2}\frac{\partial^2 f}{\partial y^2} = \frac{2e^{3x+2y} - 2(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{4(e^{3x+2y})^2 - 4(e^{3x+2y})}{(1+e^{3x+2y})^3} \rightarrow 2$$

$$\frac{\partial f}{\partial y} = \frac{-2e^{3x+2y}}{(1+e^{3x+2y})^2} \Rightarrow -\frac{1}{2}\frac{\partial f}{\partial y} = \frac{e^{3x+2y}}{(1+e^{3x+2y})^2}$$

$$-\frac{1}{2}\frac{\partial^2 f}{\partial x \partial y} = \frac{e^{3x+2y}(3)(1+e^{3x+2y})^2 - \left[2(1+e^{3x+2y})e^{3x+2y}(3)\, e^{3x+2y}\right]}{(1+e^{3x+2y})^4}$$

$$= \frac{3e^{3x+2y}}{(1+e^{3x+2y})^2} - \frac{6(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$= \frac{3e^{3x+2y}(1+e^{3x+2y}) - 6(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$-\frac{1}{2}\frac{\partial^2 f}{\partial x \partial y} = \frac{3e^{3x+2y} - 3(e^{3x+2y})^2}{(1+e^{3x+2y})^3}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{6(e^{3x+2y})^2 - 6e^{3x+2y}}{(1+e^{3x+2y})^3} \longrightarrow 3$$

$$\Rightarrow \frac{\partial^2 f}{\partial x^2} = \frac{9(e^{3x+2y})^2 - 9e^{3x+2y}}{(1+e^{3x+2y})^3} \rightarrow ①$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{4(e^{3x+2y})^2 - 4e^{3x+2y}}{(1+e^{3x+2y})^3} \rightarrow ②$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{6(e^{3x+2y})^2 - 6e^{3x+2y}}{(1+e^{3x+2y})^3} \rightarrow ③$$

Hessian at point (1,1)

$$\frac{\partial^2 f}{\partial x^2} = \frac{9e^{10} - 9e^5}{(1+e^5)^3} = \underline{\underline{0.059}}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{4e^{10} - 4e^5}{(1+e^5)^3} = \underline{\underline{0.0262}}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{6e^{10} - 6e^5}{(1+e^5)^3} = \underline{\underline{0.0393}}$$

Hessian :-

$$\begin{bmatrix} \partial^2 f/\partial x^2 & \partial^2 f/\partial x \partial y \\ \partial^2 f/\partial y \partial x & \partial^2 f/\partial y^2 \end{bmatrix} = \underline{\underline{\begin{bmatrix} 0.059 & 0.0393 \\ 0.0393 & 0.0262 \end{bmatrix}}}$$

Numerical Derivative result :

$$= \begin{bmatrix} 0.05885956 & 0.0392588 \\ 0.0392588 & 0.02618526 \end{bmatrix}$$

c) Logistic Regression :-

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} y^i \log h_\theta(x^i) + (1-y^i) \log (1-h_\theta(x^i))$$

Where $h_\theta(x) = g(\theta^T x)$, $g(z) = \frac{1}{1+e^{-z}}$ $[z = \theta^T x]$

$$g(z) = \frac{1}{1+e^{-z}} \Rightarrow \frac{\partial g(z)}{\partial z} = -\frac{e^{-z}(-1)}{(1+e^{-z})^2}$$

$$= \frac{e^{-z}}{(1+e^{-z})^2} \rightarrow (1)$$

$$= \frac{1}{1+e^{-z}}\left(1 - \frac{1}{1+e^{-z}}\right) = \frac{dg}{dz} = g(z)(1-g(z))$$

$\hookrightarrow (1)$

$$J(\theta) = \frac{-1}{m} \sum_{i=1}^{m} y^i \log h_\theta(x^i) + (1-y^i) \log (1 - h_\theta(x^i) \quad (1)$$

$$\frac{\partial J\theta}{\partial \theta_j} = \frac{-1}{m} \sum_{i=1}^{m} \frac{y_i}{h_\theta(x^i)} \frac{\partial h_\theta(x^i)}{\partial \theta_j} + \frac{(1-y^i)}{1-h_\theta(x^i)} \cdot \left(\frac{-\partial h_\theta(x^i)}{\partial \theta_j} \right) \hookrightarrow (2)$$

From (1)

$$h_\theta(x^i) = g(\theta^T x^i) = g(z)$$

$$\frac{\partial h_\theta(x^i)}{\partial \theta_j} = \cancel{\frac{\partial g(\theta^T x^i)}{\partial \theta^T x^i}} = \frac{\partial g(z)}{\partial z} \cdot \frac{\partial z}{\partial \theta_j}$$

$$= g(z)(1-g(z) \cdot \left[\frac{\partial(\theta^T x^i)}{\partial \theta_j} \right]$$

$$= \left[g(z)(1-g(z) \; x^i_{(j)} \right] \rightarrow (3)$$

$$\left[h_\theta(x^i) = g(z) = g(\theta^T x^i) \right]$$

Substitute (3) in (2)

$$= \frac{-1}{m} \sum_{i=1}^{m} \frac{y_i}{\cancel{h_\theta(x^i)}} \cdot \left[\cancel{h_\theta(x^i)} (1-h_\theta(x^i)) \, x^i_{(j)} \right]$$

$$+ \frac{(1-y^i)}{\cancel{1-h_\theta(x^i)}} (-1) \left[h_\theta(x^i) \left(\cancel{1-h_\theta(x^i)} \right) x^i_{(j)} \right]$$

$$= -\frac{1}{m} \sum_{i=1}^{m} y_i \, x^i_{(j)} (1-h_\theta(x^i)) - (1-y^i) h_\theta(x^i) (x^i_{(j)})$$

$$= \frac{-1}{m} \sum_{i=1}^{m} y^i x^i_{(j)} - \cancel{y^i x^i_{(j)} h_\theta(x^i)} - h_\theta(x^i)(x^i_{(j)})$$
$$\cancel{+ y^i x^i_{(j)} h_\theta(x^i)}$$

$$= \frac{-1}{m} \sum_{i=1}^{m} y^i x^i_{(j)} - h_\theta(x^i)(x^i_{(j)})$$

$$\frac{\partial J}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^{m} \left[h_\theta(x^i) - y^i \right] x^i_{(j)}$$

$$\nabla J = \begin{bmatrix} \frac{\partial J}{\partial \theta_1} \\ \frac{\partial J}{\partial \theta_2} \\ \vdots \\ \frac{\partial J}{\partial \theta_n} \end{bmatrix} = \begin{bmatrix} \frac{1}{m} \sum_{i=1}^{m} \left[h_\theta(x^i) - y^i \right] x^i_1 \\ \frac{1}{m} \sum_{i=1}^{m} \left[h_\theta(x^i) - y^i \right] x^i_2 \\ \vdots \\ \frac{1}{m} \sum_{i=1}^{m} \left[h_\theta(x^i) - y^i \right] x^i_n \end{bmatrix}$$

2) a) Coordinates of either Maxima or minima :-

(i) <u>Maxima</u> :- $(1.5, 6)$, $(-4.75, -6)$, $(-4.75, 6)$, $(1.5, 0)$,

$(-4.75, 0)$, $(1.5, 6)$

<u>Minima</u> :- $(-1.5, -3)$, $(5, 3)$, $(-1.5, 3)$, $(5, -3)$

(ii) A function is perodic if it repeats itself at regular intervals i.e the ~~t~~ period of a function is "p" if $f(x+p) = f(x)$

In the ~~above~~ Given plot the period ~~a~~ is approximately

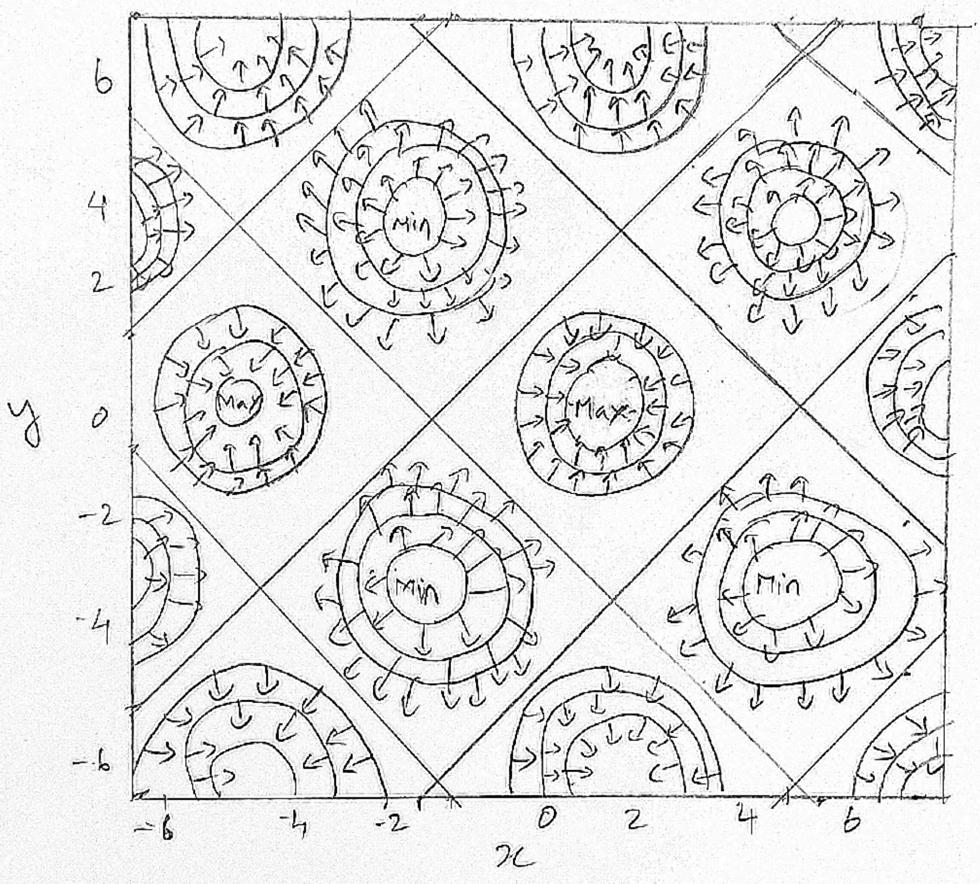
$p \simeq 6$ ~~units~~ units

(iii) The curve is not convex as it has multiple maxima and minima

(iv) ~~A point~~ ~~ca~~ A Good initial point is such that the ~~G~~ Gradient would quickly converge into a local minima.

One such point can be $(-1.2, -3)$

v) direction of Gradients



Direction of Gradients [Min = Minima, Max = Maxima] b)

(i) $\cancel{f(x)=}$ $x^2+y^2 = f(x,y)$

The function is of <u>bowl shape</u> with origin at (0,0) and has <u>1</u> minima and <u>4</u> maxima. in domain $[-1,1] \times [-1,1]$

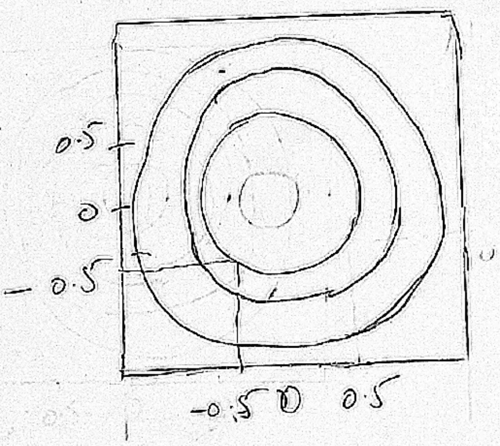
$\cancel{f(x)}$ $f(x,y) = (x-a)^2 + (y-b)^2$

This function is of <u>bowl shape</u> with origin at (a, b) and has <u>1</u> minma and <u>4</u> maxima. in domain $[-1,1] \times [-1,1]$

(ii) $f(x,y) = 10.99\left[(x-a)^2 + (y-b)^2\right]$

The function has <u>1</u> ~~mini~~ minima and <u>4</u> maxima. in domain $[-1,1] \times [-1,1]$

(iii) A good starting point is one where the gradient would ~~go~~ quickly lead us to the solution. In this case a good starting point can be <u>[-0.5, -0.5]</u>



(iv) The equation of ellipse is given by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

This can be written in a matrix form as

$$[x, y]\begin{bmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 1$$

For the ellipse shown in fig 3, $a > b$

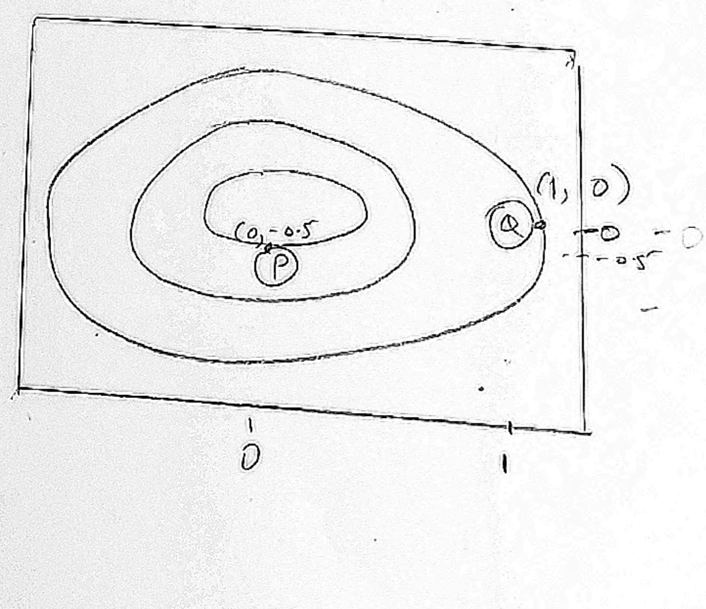
Let $a = 1$, $b = 0.4$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 6.25 \end{bmatrix}$$

Initial guess

Best = $[0, -0.5]$ P

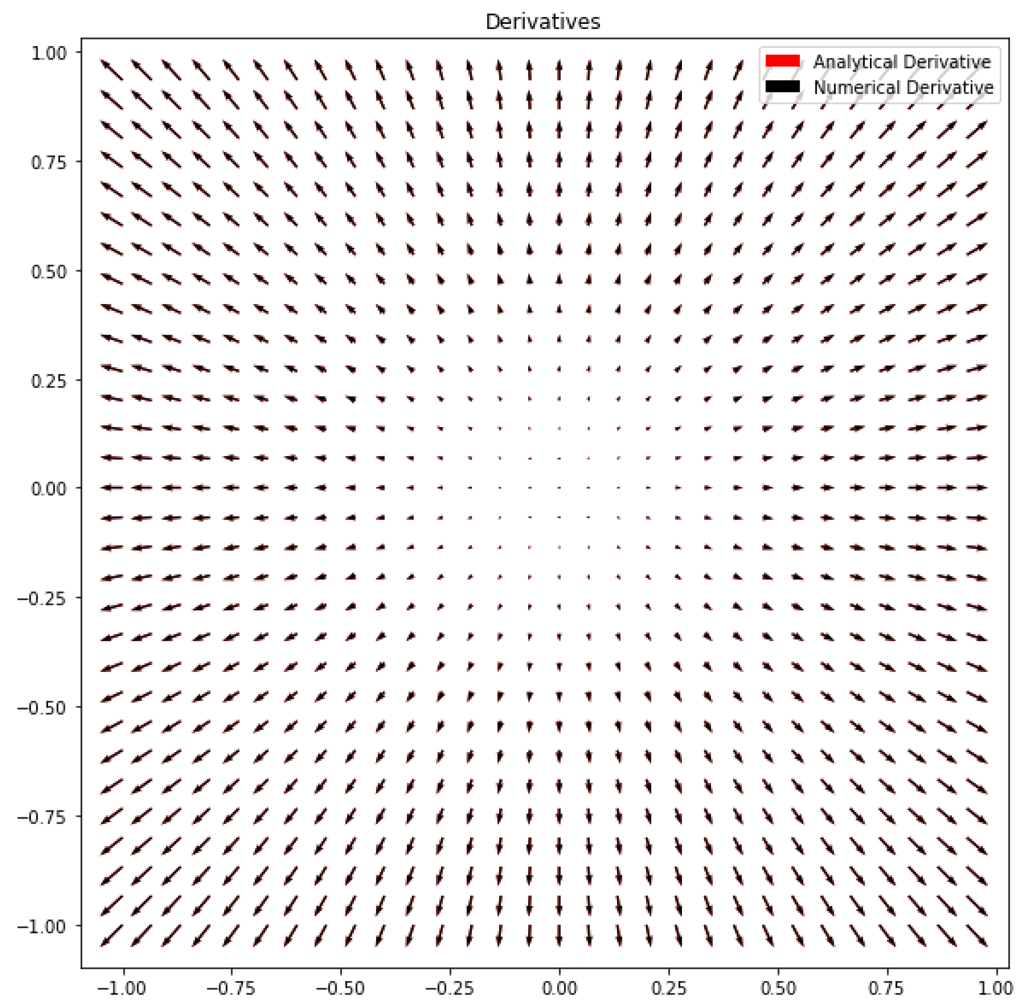
Worst = $[1, 0]$ Q



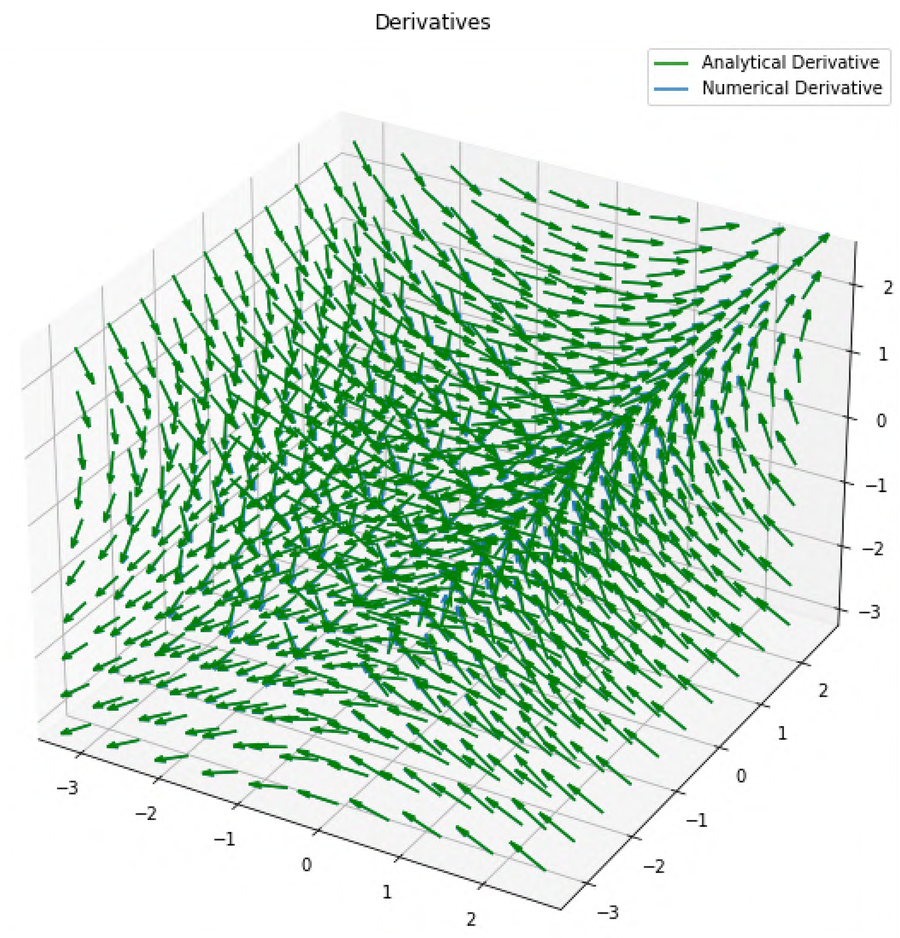
```
In [16]:
# import the necessary libraries

import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
```

Code to compute and plot Numerical and analytical Derivative for the given functions

```
In [55]:
#Part A: compute Numnerical derivative and display quiver plot that numerical and analytical derivatives lead to
# same result

x = np.arange( -1, 1, 2/30)
y = np.arange( -1, 1, 2/30)
x,y = np.meshgrid(x, y )

analytical_dx = 2*x
analytical_dy = 2*y

h = 0.01
numerical_dx = (((x + h)**2 + (y)**2) - (x**2 + y**2))/h
numerical_dy = (((x)**2 + (y + h)**2) - (x**2 + y**2))/h

fig, ax = plt.subplots(figsize =(10, 10))

ax.quiver(x, y, analytical_dx, analytical_dy, color='r')
ax.quiver(x, y, numerical_dx, numerical_dy)
ax.set_title('Derivatives')
ax.legend(['Analytical Derivative','Numerical Derivative'] , prop={"size":10})
ax.set_aspect('equal') ,

plt.show()
```




```
In [49]:
#Part B: compute Numnerical derivative and display quiver plot that numerical and analytical derivatives lead to
# same result

x = np.arange( -np.pi, np.pi, 2*np.pi/10)
y = np.arange( -np.pi, np.pi, 2*np.pi/10)
z = np.arange( -np.pi, np.pi, 2*np.pi/10)


x,y,z = np.meshgrid(x, y, z)

# Compute analytical derivatives

analytical_derivative_x = np.cos(x) + z  # Derivations in the written answer script
analytical_derivative_y = 0
analytical_derivative_z = x

# compute numerical derivatives

h = np.pi/10
numerical_derivative_x = ((np.sin(x + h) + (x + h)*z) - np.sin(x) - (x)*z)/h
numerical_derivative_y = 0
numerical_derivative_z = ((np.sin(x) + (x)*(z + h) ) - np.sin( x) - (x)*(z))/h


# plot the computed results

ax = plt.figure().add_subplot(projection='3d')

ax.quiver(x, y, z, analytical_derivative_x, analytical_derivative_y,
          analytical_derivative_z,
          color='g', length=0.5,  normalize=True)

ax.quiver(x, y, z,
          numerical_derivative_x, numerical_derivative_y, numerical_derivative_z,
          length=0.5, normalize=True)


ax.set_title('Derivatives')
ax.legend(['Analytical Derivative','Numerical Derivative'])
plt.rcParams['figure.figsize'] = [10, 10]

plt.show()
```



Code to Compute Hessian Numerically

```
In [50]:
# Part A compute Hessian of x^2 + Y^3 at point (1,2)

def function_a( x,y):
    value = x**2 + y**3

    return value


x, y, h = 1, 2, 0.001

dd_x =  (function_a(x + 2*h, y) - 2*function_a(x + h, y) + function_a(x, y))/(h**2)
dd_y =  (function_a(x, y + 2*h) - 2*function_a(x, y + h) + function_a(x, y))/(h**2)
dd_xy = (function_a(x + h, y + h) - function_a(x, y + h) - function_a(x + h, y) + function_a(x, y))/(h**2)

print("Hessian of function at point (1,2) is")

H = np.array( [ [ dd_x , dd_xy  ] ,
              [dd_xy ,dd_y ]  ])
print(H)
```

```
Hessian of function at point (1,2) is
[[ 2.      0.    ]
 [ 0.     12.006]]
```

```
In [51]:
# Part B compute Hessian of function at point (1,1)


def function_b(x,y):
        value = 1 / (1 + np.exp(3*x + 2*y))
        return value

x, y, h = 1, 1, 0.001

dd_x =  (function_b(x + 2*h, y) - 2*function_b(x + h, y) + function_b(x, y))/(h**2)
dd_y =  (function_b(x, y + 2*h) - 2*function_b(x, y + h) + function_b(x, y))/(h**2)
dd_xy = (function_b(x + h, y + h) - function_b(x, y + h) - function_b(x + h, y) + function_b(x, y))/(h**2)

print("Hessian of function at point (1,1) is")
H = np.array( [ [ dd_x , dd_xy  ] ,
              [dd_xy ,dd_y ]  ])
print(H)
```

```
Hessian of function at point (1,1) is
[[0.05885956 0.0392588 ]
 [0.0392588  0.02618526]]
```