Assignment 1

Exercise 1

$f : R^2 \to R.$

$$f(x_1, x_2) = 200\,(x_2 - x_1^2) + (1 - x_1)^2.$$

1. gradient $\nabla f : R^2 \to R^2$ and Hessian $H_f : R^2 \to R^{2\times 2}$,

we defined:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \qquad H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} -800\,x_1 (x_2 - x_1^2) + 2x_1 - 2 \\ 400\,(x_2 - x_1^2) \end{bmatrix}$$

$$H_f = \begin{bmatrix} 2400\,x_1^2 - 800x_2 + 2 & -800x_1 \\ -800\,x_1 & 400 \end{bmatrix}$$

2. The Taylor expansion of 'f' up to the second order around the point $(x_1, x_2) = (0,0)$

$$f(x_1, x_2) \approx f(0,0) + \nabla f(0,0) \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{1}{2} [x_1 \;\; x_2]\, H_f(0,0) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x_1, x_2) \approx 1 + (-2x_1 + 0) + \frac{1}{2} [x_1, x_2] \begin{bmatrix} 2 & 0 \\ 0 & 400 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 1 - 2x + \frac{1}{2} [2x_1, 400x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 1 - 2x + \frac{1}{2} (2x_1^2 + 400x_2^2)$$

$$= 1 - 2x + x_1^2 + 200\,x_2^2$$

Exercise 2

$$\min_{x \in R^n} J(x) = \frac{1}{2} x^T A x - b^T x.$$

1. Gradient of $J(x) = \nabla J(x) = \frac{\partial J}{\partial x} = Ax - b$

Hessian of $J(x) = H_J = \frac{\partial^2 J}{\partial x^2} = A.$

2. First Order necessary condition: used to identify critical points where gradient $= 0$. So

$\nabla J(x^*) = 0$

as our

$\nabla J(x) = Ax - b,$

Our first order necessary condition would be:

$$\nabla J(x^*) = Ax^* - b = 0.$$

3. Secondary order necessity condition states that critical point 'x^*', if $H(x^*)$ is positive semi-definite, then x^* is a local minimum.

$H(x^*) \geq 0$ as $H_J = A$ so $A \geq 0$.

'A' has to be positive semi-definite.

In case of secondary order sufficient condition 'x^*' has to be 'strict' local minimum.

$H(x^*) > 0$, as $H_J = A$, $A > 0$.

Hence, A has to be 'positive-definite'.

Exercise 3.

$f(x,y) = x^2 + \mu y^2$.

1. In quadratic form i.e. $\frac{1}{2} x^T A x - b^T x$, $A \in R^{2\times2}$, $b \in R^2$

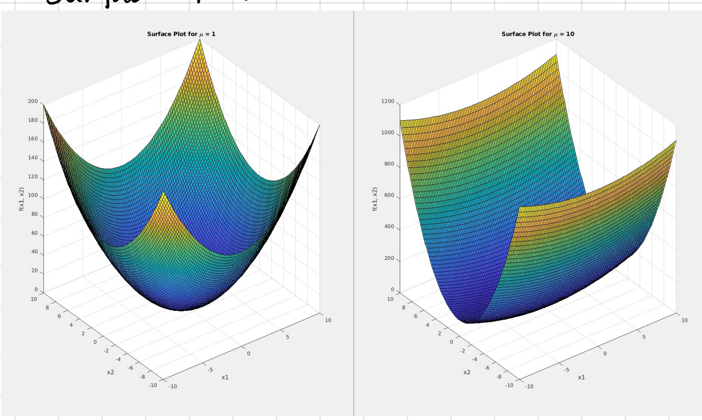
$$f(x,y) = \frac{1}{2} [x \quad y] \begin{bmatrix} 2 & 0 \\ 0 & 2\mu \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - [0 \quad 0] \begin{bmatrix} x \\ y \end{bmatrix}.$$

$$= \frac{1}{2} [2x \quad 2\mu y] \begin{bmatrix} x \\ y \end{bmatrix} - (0)$$

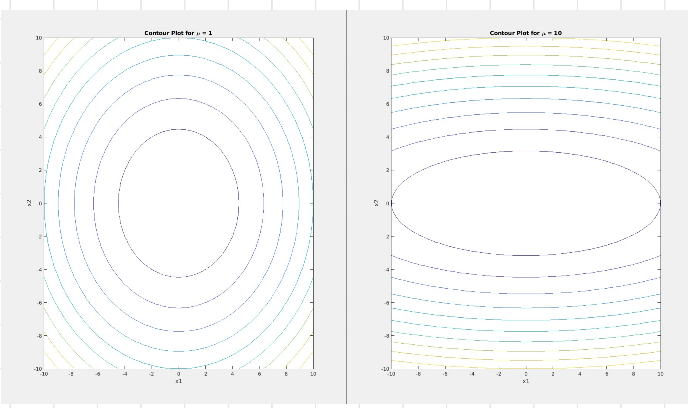
$$= \frac{1}{2} [2x^2 + 2\mu y^2)$$

$$= x^2 + \mu y^2.$$

3.2: Surface Plot



3.2 : Contour Plot.



from viewing isolines in the contour plots;
For $\mu = 1$, The isolines are more circular, which means the function varies more uniformly in all directions.

For $\mu = 10$, The isolines become more stretched along the y-axis, which means function varies more along y-direction due to higher weight on y^2 than x^2.

3.3

$$\alpha^* = \frac{\nabla f(x,y)^T \cdot \nabla f(x,y)}{\nabla f(x,y)^T \cdot H_f \cdot \nabla f(x,y)}$$

$$\nabla f(x,y) = \begin{bmatrix} \frac{df}{dx} \\ \frac{df}{dy} \end{bmatrix} = \begin{bmatrix} 2x \\ 2\mu y \end{bmatrix}$$

$$\alpha^* = \frac{[2x, 2\mu y] \begin{bmatrix} 2x \\ 2\mu y \end{bmatrix}}{[2x, 2\mu y] \begin{bmatrix} 2 & 0 \\ 0 & 2\mu \end{bmatrix} \begin{bmatrix} 2x \\ 2\mu y \end{bmatrix}}$$

$$= \frac{4x^2 + 4\mu^2 y^2}{[4x, 4\mu^2 y] \begin{bmatrix} 2x \\ 2\mu y \end{bmatrix}}$$

$$= \frac{4x^2 + 4\mu^2 y^2}{8x^2 + 8\mu^3 y^2} = \frac{x^2 + \mu^2 y^2}{2x^2 + 2\mu^3 y^2}$$

$$= \frac{1}{2} \left(\frac{x^2 + \mu^2 y^2}{x^2 + \mu^3 y^2} \right)$$