Tarea 05 - Optimización
Giovanni Gamaliel López Padilla

Problema 01

Calcule y clasifique los puntos críticos de la siguiente función

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

Muestra la función usando python

Se tiene que el gradiente de la función $f(x_1, x_2)$ es:

$$\nabla f(x_1 x_2) = \begin{bmatrix} 4x_1(x_1^2 + x_2^2 - 1) \\ 4x_2(x_1^2 + 2x_2 - 2) \end{bmatrix}$$

y su hessiano es el siguiente:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 12x_1 + 4x_2 - 4 & 8x_1 x_2 \\ 8x_1 x_2 & 4x_1 + 24x_2 - 8 \end{bmatrix}$$

Con el gradiente de $f(x_1, x_2)$ se pueden encontrar los valores críticos, los cuales son:

$$x_1 = 0 \qquad x_1^2 + x_2^2 = 1 \qquad \rightarrow \qquad x_2 = -1, 1$$

$$x_2 = 0 \qquad x_1^2 + 2x_2^2 = 2 \qquad \rightarrow \qquad x_1 = -1, 1$$

por ende, los puntos críticos son $x = \{(0,0), (0,1), (0,-1), (1,0), (-1,0)\}$

Evualuando el hessiano con los puntos críticos se obtiene la siguiente tabla:

Punto crítico	$(0,1)$	$(0,-1)$	$(1,0)$	$(-1,0)$	$(0,0)$
$\nabla^2 f(x)$	$\begin{bmatrix} 0 & 0 \\ 0 & 16 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 16 \end{bmatrix}$	$\begin{bmatrix} 8 & 0 \\ 0 & -4 \end{bmatrix}$	$\begin{bmatrix} 8 & 0 \\ 0 & -4 \end{bmatrix}$	$\begin{bmatrix} -4 & 0 \\ 0 & -8 \end{bmatrix}$

Tabla 1: Hessianos de cada punto crítico encontrado

Con estos resultados se pueden obtener que los puntos $(0,1)$ y $(0,-1)$ son mínimos. Por el hecho que la función esta acotada para número mayores e iguales a cero y la función evaluada en estos puntos es cero ($f(0,1) = f(0,-1) = 0$) se afirma que estos dos puntos representan mínimos globales.

Por otro lado los puntos $(1,0)$ y $(-1,0)$ representan puntos silla debido a que su hessiano no es positivo definido.

Por ultimo el punto $(0,0)$ representa un máximo de la función y como la misma no tiene una cota superior, entonces este representa un máximo local.

En la figura 1 se ve representada la función con los puntos crítcos marcados con estrellas

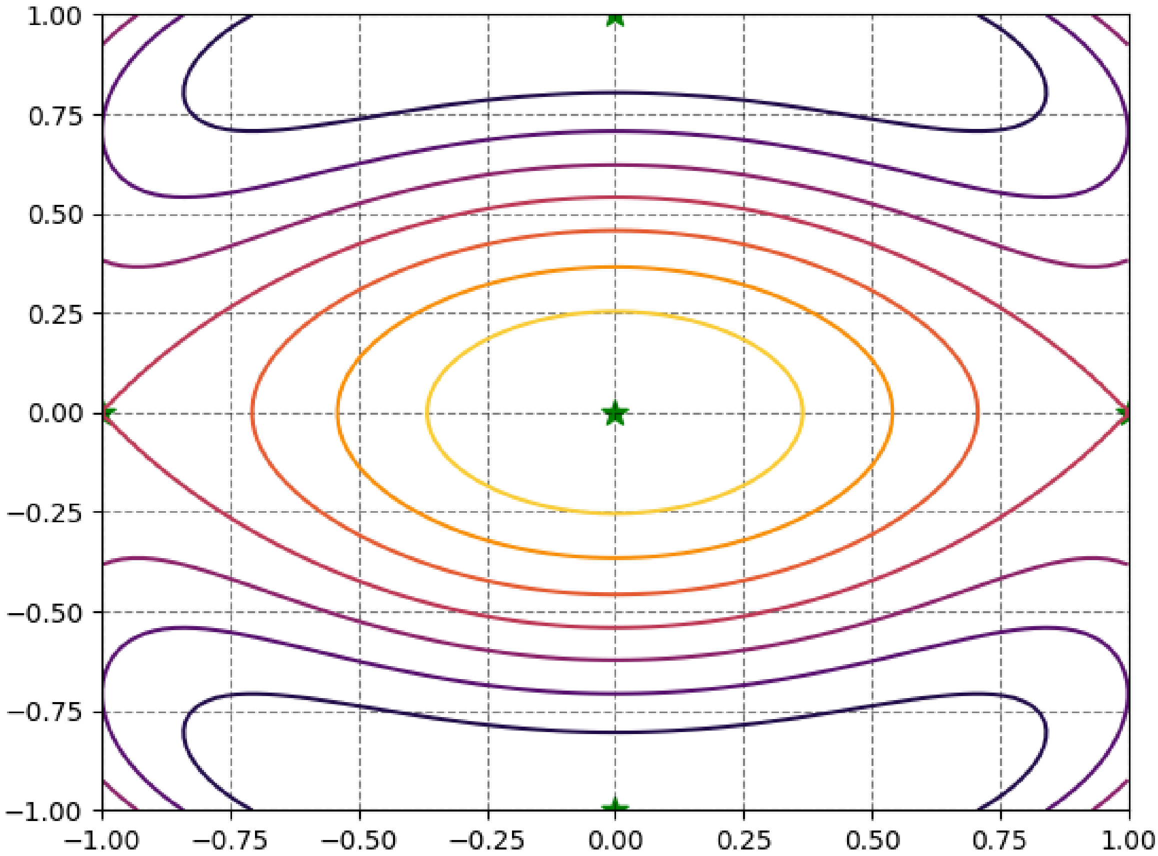

Figura 1

Problema 02

Sea A una matriz positiva definida. Muestra que

$$A_{ij} < \frac{A_{ii} + A_{jj}}{2}$$

Sea A' una matriz tal que sus elementos son submatrices de la matriz A tal que:

$$A' = \begin{bmatrix} A_{ii} & A_{ij} \\ A_{ji} & A_{jj} \end{bmatrix}$$

calculando el determinante de A' se tiene lo siguiente:

$$detA' = A_{ii}A_{jj} - A_{ij}^2$$

lo cual es una descomposición de Schur de la matriz A. Como A es una matriz positiva definida entonces

$$A_{ii}A_{jj} - A_{ij}^2 > 0$$

por ende

$$A_{ij} < \sqrt{A_{ii}A_{jj}}$$

por la desigualdad de $\sqrt{A_{ii}A_{jj}} \leq \frac{A_{ii}+A_{jj}}{2}$ entonces:

$$A_{ij} < \frac{A_{ii}+A_{jj}}{2}$$

Problema 03

Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa. Muestra que para todo x, y se cumple

$$f(y) \geq f(x) + \alpha(f(x) - f(z))$$

donde $\alpha > 0$ y $z = x + \frac{1}{\alpha}(x - y)$.
Sen dos parámetros β_1, β_2 tal que $\beta_1 > 0$, $\beta_2 \leq 0$ y $\beta_1 + \beta_2 = 1$. Con esto pueden definir de la siguiente manera:

$$\beta_1 = 1 + \frac{1}{\alpha} \qquad \beta_2 = -\frac{1}{\alpha}$$

Se tiene la siguiente operación:

$$\begin{aligned}
\frac{1}{\beta_1}(\beta_1 x + \beta_2 y) + \left(1 - \frac{1}{\beta_1}\right) y &= x + \frac{\beta_2}{\beta_1} y + y - \frac{1}{\beta_1} y \\
&= x + y + \frac{\beta_2 - 1}{\beta_1} y \\
&= x + y - \frac{\beta_1}{\beta_1} y \\
&= x + y - y \\
&= x
\end{aligned}$$

Entonces, se tiene que:

$$f(x) = f\left(\frac{1}{\beta_1}(\beta_1 x + \beta_2 y) + \left(1 - \frac{1}{\beta_1}\right) y\right)$$

como f es convexa y $0 > \frac{1}{\beta_1} > 1$, Entonces

$$
\begin{aligned}
f(x) &= f\left(\frac{1}{\beta_1}(\beta_1 x + \beta_2 y) + \left(1 - \frac{1}{\beta_1}\right) y\right) \\
&\leq \frac{1}{\beta_1} f(\beta_1 x + \beta_2 y) + \left(1 - \frac{1}{\beta_1}\right) f(y) \\
&\leq \frac{1}{\beta_1} f(\beta_1 x + \beta_2 y) - \frac{\beta_2}{\beta_1} f(y) \\
\beta_1 f(x) &\leq f(\beta_1 x + \beta_2 y) - \beta_2 f(y) \\
\left(\frac{\alpha+1}{\alpha}\right) f(x) &\leq f\left(\frac{\alpha+1}{\alpha} x - \frac{1}{\alpha} y\right) + \frac{1}{\alpha} f(y) \\
(\alpha+1) f(x) &\leq \alpha f\left(x + \frac{1}{\alpha} x - \frac{1}{\alpha} y\right) + f(y) \\
f(y) &\geq (\alpha+1) f(x) - \alpha f\left(x + \frac{1}{\alpha} x - \frac{1}{\alpha} y\right) \\
f(y) &\geq f(x) + \alpha f(x) - \alpha f\left(x + \frac{1}{\alpha}(x - y)\right) \\
f(y) &\geq f(x) + \alpha f(x) - \alpha f\left(x + \frac{1}{\alpha}(x - y)\right) \\
f(y) &\geq f(x) + \alpha f(x) - \alpha f(z) \\
f(y) &\geq f(x) + \alpha \left(f(x) - f(z)\right)
\end{aligned}
$$

Problema 04

Sea

$$
f(x) = \frac{1}{2} x^T \begin{bmatrix} \frac{3}{2} & 2 \\ 0 & \frac{3}{2} \end{bmatrix} x + x^T \begin{bmatrix} 3 \\ -1 \end{bmatrix} - 22
$$

Parte 1

Si se usa al algoritmo de gradiente descendente con tamaño de paso fijo para mininizar la funcion anterior, diga el rango de valores que puede tomar el tamaño de paso para que el algoritmo converja al minimizador.

Al ser Q una matriz no simétrica, entonces llevaremos a la función f a una función cuadraica con Q simétrica como se encuentra en la siguiente ecuación:

$$
f(x) = \frac{1}{2} x^T \begin{bmatrix} \frac{3}{2} & 1 \\ 1 & \frac{3}{2} \end{bmatrix} x + x^T \begin{bmatrix} 3 \\ -1 \end{bmatrix} - 22
$$

donde los eigenvalores de la matriz son $\lambda = \{1/2, 5/2\}$. Se tiene que para el descenso del gradiente, el rango del paso fijo con convergencia de una ecuación cuadratica esta dado por la ecuación:

$$0 < \alpha < \frac{2}{A}$$

donde A es el máximo eigenvalor de la matriz Q. En este caso $A = 5/2$. Por lo tanto el paso fijo esta acotado por

$$0 < \alpha < \frac{4}{5}$$

Parte 2

Calcula el tamaño de paso exacto α_0 si el punto incial es $x_0 = [0,0]^T$?

Se tiene que se puede calcular el k-esimo tamaño de paso de la siguiente manera:

$$\alpha_k = \frac{g_k^T g_k}{g_k^T Q g_k}$$

En nuestro caso, el gradiente es:

$$g_k = Qx - b$$

para $x_0 = [0,0]^T$, se tiene que g_k es

$$g_k = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

por ende

$$g_k^T g_k = 10 \qquad g_k^T Q g_k = 9$$

por lo tanto el tamaño de paso para x_0 es :

$$\alpha_k = \frac{10}{9}$$