Tarea 02 - Optimización

Giovanni Gamaliel López Padilla

Problema 01

Show that $x - sinx = o(x^2)$, as $x \to 0$

Realizando el límite cuando $x \to 0$ de $f(x) = \frac{x-sinx}{x^2}$, se obtiene lo siguente:

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{x - sinx}{x^2}$$

aplicando L'Hopital a este límite se encuentra que:

$$\begin{aligned}
\lim_{x\to 0} f(x) &= \lim_{x\to 0} \frac{x - sinx}{x^2} \\
&= \lim_{x\to 0} \frac{1 - cosx}{x} \\
&= \lim_{x\to 0} sinx \\
\lim_{x\to 0} f(x) &=
\end{aligned}$$

como obtenemos que el límite cuando $x \to 0$ de $f(x)$ tiende a 0, entonces

$$x - sinx \in o(x^2)$$

s

Problema 02

Suppose that $f(x) = o(g(x))$. Show that $f(x) = O(g(x))$. Tip: Show that for any given $\epsilon > 0$, there exists $\delta > 0$ such that if $0 < ||x|| < \delta$, then $|f(x)| < \epsilon|g(x)|$.

Se tiene que:

$$f(x) \in o(g(x)) \to \lim_{x\to 0} \frac{f(x)}{g(x)} = 0$$

eso con lleva a que para un $\epsilon > 0$ existe un $\delta > 0$ y tal que

$$\left|\frac{f(x)}{g(x)}\right| < \epsilon \qquad ||x|| < \delta$$

Por ende

$$\begin{aligned}\left|\frac{f(x)}{g(x)}\right| &< \left|\frac{f(x)}{g(x)} - L\right| + |L| \\ &< \epsilon^* + |L| \\ &\equiv \epsilon\end{aligned}$$

por lo tanto:

$$|f(x)| < \epsilon|g(x)| \qquad \forall||x|| < \delta$$

Problema 03

Show that if functions $f : R^n \to R$ and $g : R^n \to R$ satisfy $f(x) = -g(x) + o(g(x))$ and $g(x) > 0$ for all $x \neq 0$, then for all $x \neq 0$ sufficiently small, we have $f(x) < 0$.

Siguiendo la definición de la notació o pequeña, se tiene que:

$$o(g(x)) = \underset{x\to 0}{lim}\ \frac{h(x)}{g(x)}$$

entonces podemos definir que $h(x) = g(x)e(x)$, donde

$$\underset{x\to 0}{lim}\ e(x) = 0$$

Con esto, la función $f(x)$ puede ser escrita de la siguiente manera:

$$\begin{aligned}f(x) &= -g(x) + e(x)g(x) \\ f(x) &= g(x)(e(x) - 1)\end{aligned}$$

Calculando el límite cuando $f(x)$ tiende a cero, se obtiene lo siguiente:

$$\begin{aligned}\underset{x\to 0}{lim}\ f(x) &= \underset{x\to 0}{lim}\ g(x)(e(x) - 1) \\ &= \underset{x\to 0}{lim}\ g(x)\underset{x\to 0}{lim}\ (e(x) - 1) \\ &= \underset{x\to 0}{lim}\ g(x)\left(\underset{x\to 0}{lim}\ e(x) - \underset{x\to 0}{lim}\ -1\right) \\ &= \underset{x\to 0}{lim}\ g(x)\,(-1)\end{aligned}$$

como $g(x) > 0 \forall x \neq 0$, entonces, se puede obtener el límite cuando x tiende a 0 es positivo, por lo tanto:

$$\begin{aligned}&\underset{x\to 0}{lim}\ f(x) = \lim x0 - g(x) \\ &\underset{x\to 0}{lim}\ f(x) < 0\end{aligned}$$

Problema 04

Compute the stationary points of $f(x,y) = \frac{3x^4-4x^3-12x^2+18}{12(1+4y^2)}$ and determine their corresponding type (ie: minimum, maximum or saddle point)

Para obtener los valors críticos de $f(x,y)$ tenemos que calcular el gradiente del mismo. Calculando la derivada parcial con respecto a x de la función $f(x,y)$ se obtiene lo siguiente:

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{3x^4-4x^3-12x^2+18}{12(1+4y^2)}\right)\\ \frac{\partial f(x,y)}{\partial x} &= \frac{12x^3-12x^2-24x}{12(1+4y^2)}\end{aligned}$$

Calculando la derivada parcial con respecto a y de la función $f(x,y)$ se obtiene lo siguiente:

$$\begin{aligned}\frac{\partial f(x,y)}{\partial y} &= \frac{\partial}{\partial x}\left(\frac{3x^4-4x^3-12x^2+18}{12(1+4y^2)}\right)\\ \frac{\partial f(x,y)}{\partial y} &= \frac{-(8y)(3x^4-4x^3-12x^2+18)}{12(1+4y^2)^2}\end{aligned}$$

Como los puntos críticos son aquellos tal que $\nabla f(x^*,y^*) = 0$, se tiene que:

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= 0\\ \frac{12x^3-12x^2-24x}{12(1+4y^2)} &= 0\\ 12x^3-12x^2-24x &= 0\\ 12x(x-2)(x+1) &= 0\\ x_1 &= 0\\ x_2 &= 2\\ x_3 &= -1\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x,y)}{\partial y} &= 0\\ \frac{-(8y)(3x^4-4x^3-12x^2+18)}{12(1+4y^2)^2} &= 0\\ 8y &= 0\\ y &= 0\end{aligned}$$

Entonces, se obtiene los siguientes puntos críticos.

$$\begin{aligned}&P_1(0,0)\\ &P_2(2,0)\\ &P_3(-1,0)\end{aligned}$$

Calculando el Hessiano de $f(x, y)$ se obtiene lo siguiente:

$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{12x^3 - 12x^2 - 24x}{12(1+4y^2)}\right)$$
$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{36x^2 - 24x - 24}{12(1+4y^2)}$$

$$\frac{\partial^2 f(x,y)}{\partial y^2} = \frac{\partial}{\partial x}\left(\frac{-(8y)(3x^4 - 4x^3 - 12x^2 + 18)}{12(1+4y^2)^2}\right)$$
$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{(8)(3x^4 - 4x^3 - 12x^2 + 18)}{12}\left(\frac{4y^2 - 1}{(4y^2+1)^2}\right)$$

$$\begin{aligned}
\frac{\partial^2 f(x,y)}{\partial x \partial y} &= \frac{\partial^2 f(x,y)}{\partial y \partial x} \\
&= \frac{\partial}{\partial x}\left(\frac{\partial f(x,y)}{\partial y}\right) \\
&= \frac{\partial}{\partial x}\left(\frac{-(8y)(3x^4 - 4x^3 - 12x^2 + 18)}{12(1+4y^2)^2}\right) \\
\frac{\partial^2 f(x,y)}{\partial x \partial y} &= \frac{-(8y)(12x^3 - 12x^2 - 24)}{12(1+4y^2)^2}
\end{aligned}$$

Se evaluara $\nabla^2 f(x, y)$ para describir si se trata de un punto silla, mínimo o máximo de la función. En la tabla 1 se muestran los resultados obtenidos.

x	y	$\partial_{xx} f(x,y)$	$\partial_{yy} f(x,y)$	$\partial_x \partial_y f(x,y)$	$\lvert\nabla^2 f(x,y)\rvert$
0	0	-2	-12	0	24
2	0	6	9.3334	0	56
-1	0	3	-8.6667	0	-26

Tabla 1: Resultados de las segundas derivadas parciales de la función $f(x, y)$ evaluadas en sus puntos críticos.

Un punto es punto silla si $|\nabla^2 f(x, y)| < 0$, por lo tanto, el punto $(-1, 0)$ representa un punto silla.

Un punto es un máximo si $|\nabla^2 f(x, y)| > 0$ y $\partial_{xx} f(x, y) < 0$, por lo tanto el punto $(0, 0)$ es un punto máximo.

Un punto es un mínimo si $|\nabla^2 f(x, y)| > 0$ y $\partial_{xx} f(x, y) > 0$, por lo tanto el punto $(2, 0)$ es un punto mínimo.

Problema 05

Show that the function $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$ has only one stationary point, and that it is neither a maximum or minimum, but a saddle point. Plot the contour lines of f.

Calculando la primer derivada parcial con respecto a x_1 de la función $f(x)$ se obtiene lo siguiente:

$$\frac{\partial f(x)}{\partial x_1} = 8 + 2x_1$$

Calculando la primer derivada parcial con respecto a x_2 de la función $f(x)$ se obtiene lo siguiente:

$$\frac{\partial f(x)}{\partial x_2} = 12 - 4x_2$$

Por lo tanto, el punto crítico obtenido es $x_c = (-4, 3)$.

Calculando el hessiano de $f(x)$ se obtiene lo siguiente:

$$\nabla^2 f(x) = \begin{vmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & -4 \end{vmatrix}$$

$$|\nabla^2 f(x)| = -8$$

Como $\nabla^2 f(x) < 0$ para cualquier x, entonces x_c es un punto silla.

En la figura 1 se representan las curvas de nivel de la función $f(x)$.

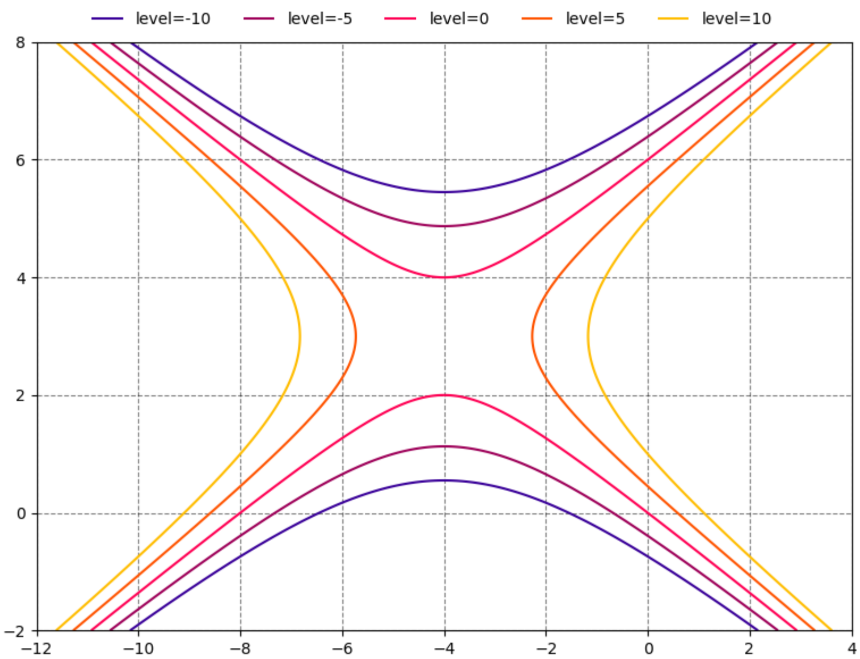

Figura 1: Curvas de niel de la función $f(x)$.

Problema 06

Compute the gradient $\nabla(x)$ and Hessian $\nabla^2 f(x)$ of the Rosenbrock function

$$f(x) = \sum_{i=1}^{N-1}[100(x_{i+1} - x_i^2) + (1 - x_i)^2]$$

where $x = [x_1, \ldots, x_N]^T \in R^N$ If n = 2 show that $x^* = [1, 1]^T$ is the only local minimizer of this function, and that the Hessian matrix at that point is positive definite. Plot the contour lines of f.

Para $n = 2$ se obtiene que $f(x)$ es:

$$f(x) = 100(x2 - x_1^2)^2 + (1 - x_1)^2$$

Calculando el gradiente se obtiene lo siguiente:

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

$$\frac{\partial f(x)}{\partial x_2} = 200(x_2 - x_1^2)$$

A partir de $\partial_{x_2} f(x) = 0$, se obtiene la realcion $x_2 = x_1^2$. Usando esta relación en $\partial_{x_1} f(x)$ se obtiene que $x_1 = 0, 1$, por ende $x_2 = 0, 1$. Dando así, los puntos críticos de $f(x)$ son $\{(0, 0), (1, 1)\}$.

Calculando el hessiano de $f(x)$ se obtiene que:

$$\frac{\partial^2 f(x)}{\partial x_1^2} = 800x_1^2 - 400(x_2 - x_1^2) + 2x - 2(1 - x_1)$$

$$\frac{\partial^2 f(x)}{\partial x_2^2} = 200$$

$$\frac{\partial^2 f(x)}{\partial x_2 x_1} = -400x_1$$

Por lo tanto, el hessiano es:

$$\nabla^2 f(x) = \begin{bmatrix} 800x_1^2 - 400(x_2 - x_1^2) + 2x - 2(1 - x_1) & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

Evaluando en $(0, 0)$ se obtiene que $|\nabla^2 f(0, 0)| = -400$, por lo que ese punto representa un punto silla. En cambio evaluando se obtiene que $|\nabla^2 f(1, 1)| = 400$ y $\partial_{x_1}^2 f(1, 1) = 802$. Por lo que este representa un punto mínimo.

En la figura 2 se representa las curvas de nivel de la función $f(x)$.

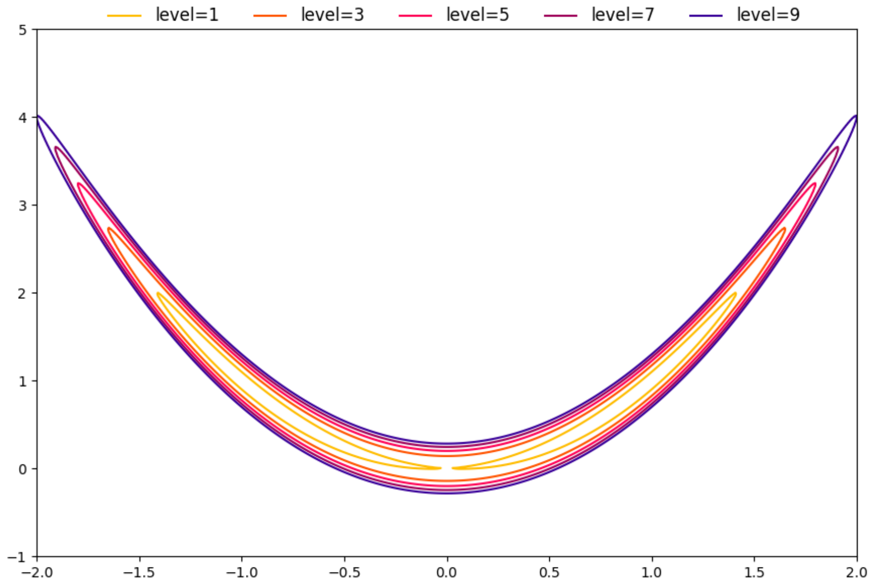

Figura 2: Curvas de niel de la función $f(x)$.

Problema 07

Show, without using the optimality conditions, that $f(x) > f(x^*)$ for all $x \neq x^*$ if

$$f(x) = \frac{1}{2}x^TQx - b^Tx$$

$Q = Q^T \succ 0$ **and** $Qx^* = b$**.**

Se tiene que $x \to x + x^* - x^* = x^* + (x - x^*)$, entonces:

$$\begin{aligned}
f(x^* + (x-x^*)) &= \frac{1}{2}(x^* + (x-x^*))^TQ(x^* + (x-x^*) - b^T(x-x^*) \\
&= \frac{1}{2}\left(x^{*T}Qx^* + (x-x^*)^TQx^* + x^{*T}Q(x-x^*) + (x-x^*)^TQ(x-x^*)\right) - b^Tx^* - b^T(x-x^*)^T \\
&= \frac{1}{2}x^{*T}Qx^* - b^Tx^* + (x-x^*)^TQx^* + \frac{1}{2}(x-x^*)^TQ(x-x^*) - b^T(x-x^*)^T \\
&= f(x^*) + (x-x^*)^TQx^* + \frac{1}{2}(x-x^*)^TQ(x-x^*) - b^T(x-x^*)^T \\
&= f(x^*) + (x-x^*)^T(Qx^* - b) + \frac{1}{2}(x-x^*)^TQ(x-x^*) \\
f(x) &= f(x^*) + \frac{1}{2}(x-x^*)^TQ(x-x^*)
\end{aligned}$$

Como Q es una matriz simetrica con eigenvalores positivos, entonces también es definida positiva. $(x - x^*)$ es un vector, entonces $(x-x^*)^TQ(x-x^*) > 0$, por lo tanto:

$$\begin{aligned}
f(x) &= f(x^*) + \frac{1}{2}(x-x^*)^TQ(x-x^*) \\
f(x) &> f(x^*)
\end{aligned}$$

Problema 08

Let $A \in R^{n \times n}$ be a real symmetric matrix. Show that

$$max\{x^T Ax : ||x|| = 1\} = \lambda_{max}(A)$$

Realizando una descomposición SVD a la matriz A, se obtiene que es igual a:

$$A = U^T DU$$

donde U es una matriz que contiene a los eigenvectores de A y D es una matriz diagonal que contiene en su diagonal los eigenvalores de A. Entonces:

$$\begin{aligned} x^T Ax &= x^T U^T AUx \\ &= (Ux)^T AUx \end{aligned}$$

La operación Ux representa un cambio de base, dando asi que $Ux \to y$, por ende:

$$\begin{aligned} x^T Ax &= y^T Dy \\ &= \sum_i \lambda_i y_i^2 \\ &= \sum_i \lambda_i (U_i x_i)^2 \end{aligned}$$

ordenando los eigenvalores de tal manera que $\lambda_1 > \lambda_2 > \ldots > \lambda_n$, entonces

$$\begin{aligned} x^T Ax &= \sum_i \lambda_i (U_i x_i)^2 \\ &\leq \lambda_1 \sum_i (U_i x_i)^2 \\ &\leq \lambda_1 ||x||^2 \\ &\leq \lambda_1 \\ max(x^T Ax) &= \lambda_1 \\ max(x^T Ax) &= \lambda_{max} \end{aligned}$$