

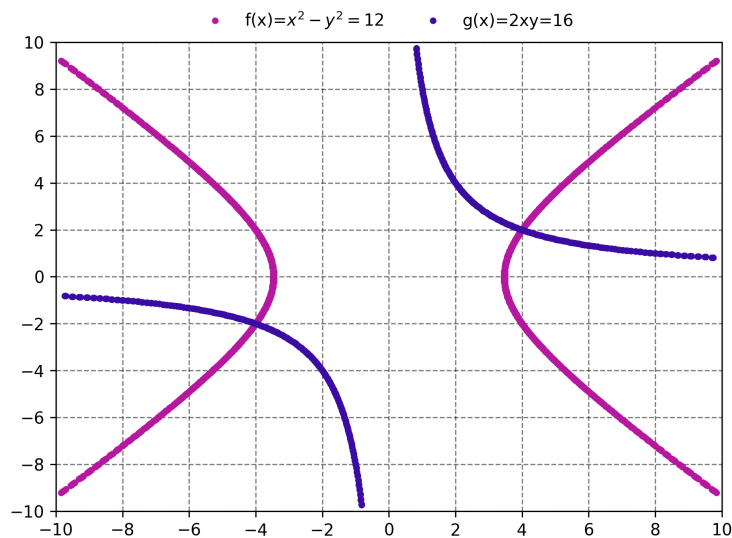
## Tarea 01 - Optimización

### Giovanni Gamaliel López Padilla

## Problema 01

Let  $f_1(x_1, x_2) = x_1^2 - x_2^2$ ,  $f_2(x_1, x_2) = 2x_1x_2$ . Represent the level sets associated with  $f_1(x_1, x_2) = 12$  and  $f_2(x_1, x_2) = 16$  on the same figure using Python. Indicate on the figure, the points  $x = [x_1, x_2]^T$  for which  $f(x) = [f_1(x_1, x_2), f_2(x_1, x_2)]^T = [12, 16]^T$ .

En la figura 1 se representan de los level sets de  $f_1(x_1, x_2) = x_1^2 - x_2^2 = 12$  y  $f_2(x_1, x_2) = 2x_1x_2 = 16$ .



**Figura 1:** Representación de los level sets para las funciones  $f_1$  y  $f_2$ .

## Problema 02

Consider the function  $f(x) = (a^T x)(b^T x)$ , where  $a$ ,  $b$  and  $x$  are  $n$ -dimensional vectors. Compute the gradient  $\nabla f(x)$  and the Hessian  $\nabla^2 f(x)$ .

Se tiene que:

$$f(x) = (a^T x)(b^T x)$$

como  $a$ ,  $b$  y  $x$  son vectores, entonces se puede hacer el cambio de  $a^T x = x^T a$ , por lo que la función  $f(x)$  es:

$$f(x) = (x^T a)(b^T x)$$

Calculando la derivada de  $f$  con respecto a  $x$ , se obtiene que:

$$\frac{df}{dx} = \frac{d(a^T x)(b^T x)}{dx}$$

aplicando la regla de la cadena de tal manera que  $f(x) = h^T(x)g(x)$  donde  $h(x) = x^T a$  y  $g(x) = b^T x$ . Entonces:

$$\begin{aligned}\frac{df}{dx} &= h^T(x) \frac{dg}{dx} + g^T(x) \frac{dh}{dx} \\ \frac{df}{dx} &= a^T x(b^T) + x^T b(a^T) \\ \frac{df}{dx} &= x^T a b^T + x^T b a^T\end{aligned}$$

como  $\nabla f = Df^T$ , entonces, el gradiente de f es:

$$\nabla f(x) = (ba^T + ab^T)x$$

Calculando  $\nabla^2 f(x)$  se obtiene que es igual a:

$$\nabla^2 f(x) = ba^T + ab^T$$

## Problema 03

Compute the gradient of

$$f(\theta) = \frac{1}{2} \sum_{i=1}^n [g(x_i) - g(Ax_i + b)]^2$$

with respect to  $\theta$ , where  $\theta = [a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2]^T$ ,  $x_i \in \mathbb{R}^2$  are defined as follows:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad b = [b_1, b_2]^T$$

and  $g : \mathbb{R}^2 \rightarrow \mathbb{R} \in C^1$ .

Calculando la derivada de f con respecto  $\theta$  se obtiene lo siguiente:

$$\begin{aligned}\frac{df}{d\theta} &= \frac{d}{d\theta} \left( \frac{1}{2} \sum_i [g(x_i) - g(Ax_i + b)]^2 \right) \\ &= \frac{1}{2} \sum_i \frac{d}{d\theta} [g(x_i) - g(Ax_i + b)]^2 \\ &= \frac{1}{2} \sum_i 2 [g(x_i) - g(Ax_i + b)] \frac{d}{d\theta} (g(x_i) - g(Ax_i + b)) \\ &= \frac{1}{2} \sum_i 2 [g(x_i) - g(Ax_i + b)] \left( \frac{dg(x_i)}{d\theta} - \frac{dg(Ax_i + b)}{d\theta} \right)\end{aligned}$$

La derivada de  $g(x_i)$  con respecto  $\theta$  es igual a cero, ya que la función no depende de los elemento de la matriz  $A$  o el vector  $b$ . Calculando la derivada de  $g(Ax_i + b)$  con respecto  $\theta$  se obtiene que:

$$\begin{aligned}\frac{dg(Ax_i + b)}{d\theta} &= \frac{dg(Ax_i + b)}{d(Ax_i + b)} \frac{\partial Ax_i + b}{\partial \theta} \\ \frac{dg(Ax_i + b)}{d\theta} &= \frac{dg(Ax_i + b)}{d(Ax_i + b)} \begin{pmatrix} x_{i_1} & x_{i_2} & 0 & 0 & 1 & 0 \\ 0 & 0 & x_{i_1} & x_{i_2} & 0 & 1 \end{pmatrix}\end{aligned}$$

por lo tanto, la derivada de  $f$  con respecto a  $\theta$  es:

$$\frac{df}{d\theta} = - \sum_i [g(x_i) - g(Ax_i + b)] \left( \frac{dg(Ax_i + b)}{d(Ax_i + b)} \right) \begin{pmatrix} x_{i_1} & x_{i_2} & 0 & 0 & 1 & 0 \\ 0 & 0 & x_{i_1} & x_{i_2} & 0 & 1 \end{pmatrix}$$

como  $\nabla F = Df^T$ , entonces, el gradiente de  $f$  con respecto  $\theta$  es:

$$\nabla f(\theta) = - \sum_i [g(x_i) - g(Ax_i + b)] \begin{pmatrix} x_{i_1} & 0 \\ x_{i_2} & 0 \\ 0 & x_{i_1} \\ 0 & x_{i_2} \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \left( \frac{dg(Ax_i + b)}{d(Ax_i + b)} \right)$$

## Problema 04

Let  $f(r, \theta)$  be  $\mathbb{R}^2 \rightarrow \mathbb{R}$  with  $r = \sqrt{x^2 + y^2}$  and  $\theta = \arctan(y/x)$ . Compute  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$ .

Se tiene que una derivada parcial puede escribir como la ecuación 1 debido a la regla de la cadena.

$$\frac{\partial f}{\partial q} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial q} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial q} \quad (1)$$

donde  $q$  es una variable generalizada que puede ser  $x$  o  $y$ .

Calculando  $\frac{\partial r}{\partial q}$  y  $\frac{\partial \theta}{\partial q}$ , se obtiene lo siguiente:

$$\begin{aligned}\frac{\partial r}{\partial x} &= \frac{x}{\sqrt{x^2 + y^2}} & \frac{\partial \theta}{\partial x} &= -\frac{y}{x^2 + y^2} \\ \frac{\partial r}{\partial y} &= \frac{y}{\sqrt{x^2 + y^2}} & \frac{\partial \theta}{\partial y} &= \frac{x}{x^2 + y^2}\end{aligned}$$

Por lo tanto,  $\frac{\partial f}{\partial x}$  y  $\frac{\partial f}{\partial y}$  es:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r} \frac{x}{\sqrt{x^2 + y^2}} - \frac{\partial f}{\partial \theta} \frac{y}{x^2 + y^2}$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r} \frac{y}{\sqrt{x^2 + y^2}} + \frac{\partial f}{\partial \theta} \frac{x}{x^2 + y^2}$$

## Problema 05

The directional derivative  $\frac{\partial f}{\partial v}(x_0, y_0, z_0)$  of a differentiable function  $f$  are  $\frac{3}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}$  in the directions of vectors  $[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T, [\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]^T$  and  $[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]^T$ . Compute  $\nabla f(x_0, y_0, z_0)$ .

Se tiene el siguiente sistema:

$$\begin{aligned} [0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}] \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} &= \nabla_{v_1} f \\ [\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}] \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} &= \nabla_{v_2} f \\ [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0] \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix} &= \nabla_{v_3} f \end{aligned}$$

donde  $\nabla_{v_1} f = \frac{3}{\sqrt{2}}, \nabla_{v_2} f = \frac{1}{\sqrt{2}}$  y  $\nabla_{v_3} f = -\frac{1}{\sqrt{2}}$

Entonces, se obtiene el siguiente sistema de ecuaciones:

$$\begin{aligned} \frac{1}{\sqrt{2}} \frac{\partial f}{\partial y} + \frac{1}{\sqrt{2}} \frac{\partial f}{\partial z} &= 3 \\ \frac{1}{\sqrt{2}} \frac{\partial f}{\partial x} + \frac{1}{\sqrt{2}} \frac{\partial f}{\partial z} &= 1 \\ \frac{1}{\sqrt{2}} \frac{\partial f}{\partial x} + \frac{1}{\sqrt{2}} \frac{\partial f}{\partial y} &= -1 \end{aligned}$$

El cual puede ser llevado a ser escrito en el siguiente sistema matricial:

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0, z_0) \\ \frac{\partial f}{\partial y}(x_0, y_0, z_0) \\ \frac{\partial f}{\partial z}(x_0, y_0, z_0) \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$$

Resolviendo el sistema matricial, se obtiene que el gradiente de  $f$  evaluado en  $X_0$  es:

$$\nabla f(x_0, y_0, z_0) = \begin{bmatrix} -\frac{3}{2} & \frac{1}{2} & \frac{5}{2} \end{bmatrix}^T$$

## Problema 06

Show that the level curves of the function  $f(x, y) = x^2 + y^2$  are orthogonal to the level curves of  $g(x, y) = y/x$  for all  $(x, y)$ .

Calculando  $\nabla f$  Y  $\nabla g$ , se obtiene que:

$$\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \qquad \nabla g = \begin{bmatrix} -\frac{y}{x^2} \\ \frac{1}{x} \end{bmatrix}$$

Se tiene que si el producto punto entre dos vectores es igual a 0, entonces los dos vectores son ortogonales. Entonces calculando  $\nabla f^T \nabla g$ , se obtiene lo siguiente:

$$\begin{aligned} \nabla f^T \nabla g &= \begin{bmatrix} 2x & 2y \end{bmatrix} \begin{bmatrix} -\frac{y}{x^2} \\ \frac{1}{x} \end{bmatrix} \\ &= -\frac{2y}{x} + \frac{2y}{x} \\ &= 0 \end{aligned}$$

Como  $\nabla f^T \nabla g = 0$ , entonces las curvas de nivel de  $f$  y  $g$  son ortogonales para cualquier  $(x, y)$ .

## Problema 07

Let  $f, g, h$  be differentiable functions, with  $f : \mathbb{R}^n \rightarrow \mathbb{R}^3, g : \mathbb{R}^n \rightarrow \mathbb{R}^3$  and  $h : \mathbb{R}^n \rightarrow \mathbb{R}$

$$h(x) = f(x)^T g(x)$$

show that

$$Dh(x) = f(x)^T Dg(x) + g(x)^T Df(x)$$

Se tiene que:

$$h(x) = \begin{bmatrix} f_1 & f_2 & f_3 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix}$$

donde  $f_i, g_i$  son funciones que van de  $\mathbb{R}^n \rightarrow \mathbb{R}$ . Calculando el producto de matrices se obtiene que la función  $h$  es:

$$h(x) = f_1 g_1 + f_2 g_2 + f_3 g_3$$

Calculando  $Dh$  se obtiene que:

$$\begin{aligned}
Dh(x) &= D(f_1g_1 + f_2g_2 + f_3g_3) \\
&= D(f_1g_1) + D(f_2g_2) + D(f_3g_3) \\
&= f_1Dg_1 + g_1Df_1 + f_2Dg_2 + g_2Df_2 + f_3Dg_3 + g_3Df_3 \\
&= g_1Df_1 + g_2Df_2 + g_3Df_3 + f_1Dg_1 + f_2Dg_2 + f_3Dg_3 \\
&= \begin{bmatrix} g_1 & g_2 & g_3 \end{bmatrix} \begin{bmatrix} Df_1 \\ Df_2 \\ Df_3 \end{bmatrix} + \begin{bmatrix} f_1 & f_2 & f_3 \end{bmatrix} \begin{bmatrix} Dg_1 \\ Dg_2 \\ Dg_3 \end{bmatrix} \\
Dh(x) &= g^T Df + f^T Dg
\end{aligned}$$

## Problema 08

Consider the induced matrix norm

$$\|A\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

where  $\|\cdot\|_p$  is the  $l_p$  norm, ie

$$\|x\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

show that

$$\|AB\|_p \leq \|A\|_p \|B\|_p$$

Se tiene que:

$$\|AB\|_p = \max_{x \neq 0} \frac{\|ABx\|_p}{\|x\|_p}$$

Como A,B se encuentran en un espacio de  $\mathbb{R}^{n \times n}$  y x en  $\mathbb{R}^n$ . Entonces se cumplen las desigualdades de Cauchy-Schwarz. Agrupando al término Bx se obtiene lo siguiente:

$$\begin{aligned}
\max_{x \neq 0} \frac{\|ABx\|_p}{\|x\|_p} &= \max_{x \neq 0} \frac{\|A(Bx)\|_p}{\|x\|_p} \\
&\leq \max_{x \neq 0} \frac{\|A\|_p \|Bx\|_p}{\|x\|_p} \\
&\leq \max_{x \neq 0} \frac{\|A\|_p \|B\|_p \|x\|_p}{\|x\|_p} \\
&\leq \max_{x \neq 0} \|A\|_p \|B\|_p \\
&\leq \|A\|_p \|B\|_p
\end{aligned}$$

por lo tanto

$$\|AB\|_p \leq \|A\|_p \|B\|_p$$