Tarea Dos

Optimización I

Esteban Reyes Saldaña

23 de febrero de 2021

Problema 1. La derivada direccional $\dfrac{\partial f}{\partial v}(x_0, y_0, z_0)$ de una función diferenciable f son $\dfrac{3}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}$ y $-\dfrac{1}{\sqrt{2}}$ en las direcciones de los vectores $\left[0, \dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right]^T$, $\left[\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right]^T$ y $\left[\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right]^T$. Calcule $\nabla f(x_0, y_0, z_0)$,

Solución. En clase se vio que dada una función diferenciable $f : \mathbb{R}^n \to \mathbb{R}$ entonces

$$D_a f(x, y, z) = \nabla f(x, y, z)^T a.$$

Sean

$$\begin{aligned}
u &= \left[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T \\
v &= \left[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right]^T \\
w &= \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right]^T.
\end{aligned}$$

y

$$\nabla f(x_0, y_0, z_0) = \left[\frac{\partial f}{\partial x}(x_0, y_0, z_0), \frac{\partial f}{\partial y}(x_0, y_0, z_0), \frac{\partial f}{\partial z}(x_0, y_0, z_0)\right]^T$$

Luego,

$$\begin{aligned}
\frac{3}{\sqrt{2}} = \frac{\partial f}{\partial u}(x_0,y_0,z_0) &= \nabla f(x_0,y_0,z_0)^T u\\
\frac{1}{\sqrt{2}} = \frac{\partial f}{\partial v}(x_0,y_0,z_0) &= \nabla f(x_0,y_0,z_0)^T v\\
-\frac{1}{\sqrt{2}} = \frac{\partial f}{\partial w}(x_0,y_0,z_0) &= \nabla f(x_0,y_0,z_0)^T w.
\end{aligned}$$

Realizando el producto punto entre el gradiente y los vectores u, v, w se obtiene el sistema

$$\begin{aligned}
\frac{1}{\sqrt{2}}\cdot\frac{\partial f}{\partial y}(x_0,y_0,z_0) + \frac{1}{\sqrt{2}}\cdot\frac{\partial f}{\partial z}(x_0,y_0,z_0) &= \frac{3}{\sqrt{2}}\\
\frac{1}{\sqrt{2}}\cdot\frac{\partial f}{\partial x}(x_0,y_0,z_0) + \frac{1}{\sqrt{2}}\cdot\frac{\partial f}{\partial z}(x_0,y_0,z_0) &= \frac{1}{\sqrt{2}}\\
\frac{1}{\sqrt{2}}\cdot\frac{\partial f}{\partial x}(x_0,y_0,z_0) + \frac{1}{\sqrt{2}}\cdot\frac{\partial f}{\partial y}(x_0,y_0,z_0) &= -\frac{1}{\sqrt{2}}
\end{aligned}$$

que es lineal respecto a las derivadas parciales en las direcciones canónicas. Así que se procede a resolver el sistema matricial asociado

$$\begin{aligned}
\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & | & \frac{3}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & | & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & | & -\frac{1}{\sqrt{2}} \end{pmatrix}
&\sim \begin{pmatrix} 0 & 1 & 1 & | & 3\\ 1 & 0 & 1 & | & 1\\ 1 & 1 & 0 & | & -1 \end{pmatrix}\\
&\sim \begin{pmatrix} 0 & 1 & 1 & | & 3\\ 1 & 0 & 1 & | & 1\\ 1 & 0 & -1 & | & -4 \end{pmatrix}\\
&\sim \begin{pmatrix} 0 & 1 & 1 & | & 3\\ 1 & 0 & 1 & | & 1\\ 2 & 0 & 0 & | & -3 \end{pmatrix}\\
&\sim \begin{pmatrix} 0 & 1 & 1 & | & 3\\ 1 & 0 & 1 & | & 1\\ 1 & 0 & 0 & | & -3/2 \end{pmatrix}\\
&\sim \begin{pmatrix} 0 & 1 & 0 & | & 1/2\\ 0 & 0 & 1 & | & 5/2\\ 1 & 0 & 0 & | & -3/2 \end{pmatrix}.
\end{aligned}$$

De lo anterior, se concluye que

$$\nabla f(x_0, y_0, z_0) = \left[-\frac{3}{2}, \frac{1}{2}, \frac{5}{2}\right]^T.$$

Problema 2. Demuestre que las curvas de nivel de la función $f(x, y) = x^2 + y^2$ son ortogonales a las curvas de nivel de $g(x, y) = \dfrac{y}{x}$ para todo (x, y).

Demostración. Observemos que dada una función diferenciable $f : \mathbb{R}^n \to \mathbb{R}$, se tiene que ∇f es siempre ortogonal a la gráfica de dicha función. Así, dada $f(x, y)$ diferenciable, $\nabla f(x, y)$ es ortogonal a la gráfica de la función, es decir, es ortogonal a toda curva de nivel. Luego, para probar la ortogonalidad entre curvas de nivel entre $f(x, y)$ y $g(x, y)$ se probará la ortogonalidad entre los gradientes.

$$\begin{aligned}
\nabla f(x, y) &= [2x, 2y]^T \\
\nabla g(x, y) &= \left[-\frac{y}{x^2}, \frac{1}{x}\right]^T
\end{aligned}$$

Entonces,

$$\begin{aligned}
< \nabla f(x, y), \nabla g(x, y) > &= -\frac{2xy}{x^2} + \frac{2y}{x} \\
&= -\frac{2y}{x} + \frac{2y}{x} \\
&= 0.
\end{aligned}$$

□

Problema 3. Calcule los puntos estacionarios de $f(x, y) = \dfrac{3x^4 - 4x^3 - 12x^2 + 18}{12(1 + 4y^2)}$ y determine su tipo correspondiente (máximo, mínimo o punto silla).

Solución. Buscamos los puntos tales que

$$\nabla f(x, y) = (0, 0).$$

Tenemos que

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

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

De donde $\nabla f(x,y) = 0$ sí y solo si

$$x(x^2-x-2) = x(x-2)(x+1) = 0 \quad (1)$$
$$(3x^4-4x^3-12x^2+18)y = 0. \quad (2)$$

De la ecuación (1) tenemos que

$$x = -1 \text{ o } x = 0 \text{ o } x = 2.$$

Sustituyendo estos valores en la ecuación (2) tenemos que
para $x = -1$,

$$\begin{aligned}(3(-1)^4-4(-1)^3-12(-1)^2+18)y &= 0\\ (3+4-12+18)y &= 0\\ 13y &= 0\\ y &= 0.\end{aligned}$$

Para $x = 0$,

$$\begin{aligned}(3(0)-4(0)-12(0)+18)y &= 0\\ 18y &= 0\\ y &= 0.\end{aligned}$$

Para $x = 2$,

$$\begin{aligned}(3(2)^4-4(2)^3-12(2)^2+18)y &= 0\\ -14y &= 0\\ y &= 0.\end{aligned}$$

Así que los **puntos estacionarios** son

$$(-1,0),(0,0) \text{ y } (2,0).$$

Para identificar el tipo de punto estacionario se observarán los valores propios de la matriz Hessiana asociada a dichos puntos. Obteniendo el Hessiano,

$$\begin{aligned}
\frac{\partial^2 f(x,y)}{\partial x^2} &= 3\frac{3x^2-2x-2}{1+4y^2}\\
\frac{\partial^2 f(x,y)}{\partial y \partial x} &= \frac{x(x^2-x-2)}{(1+4y^2)^2}(-8y)\\
&= -\frac{8xy(x^2-x-2)}{(1+4y^2)^2}\\
\frac{\partial^2 f(x,y)}{\partial x \partial y} &= -\frac{2}{3}\cdot\frac{(12x^3-12x^2-24x)y}{(1+4y^2)^2}\\
\frac{\partial^2 f(x,y)}{\partial y^2} &= -\frac{2}{3}(3x^4-4x^3-12x^2+18)\left(-\frac{2}{(1+4y^2)^3}(8y)+\frac{1}{(1+4y^2)^2}\right)\\
&= -\frac{2}{3}(3x^4-4x^3-12x^2+18)\left(\frac{-16y+1+4y^2}{(1+4y^2)^3}\right).
\end{aligned}$$

Entonces

$$H_f(x,y) = \begin{pmatrix} \frac{3x^2-2x-2}{1+4y^2} & -\frac{8xy(x^2-x-2)}{(1+4y^2)^2} \\ -\frac{2y(12x^3-12x^2-24x)}{3(1+4y^2)^2} & -\frac{2(3x^4-4x^3-12x^2+18)(4y^2-16y+1)}{3(1+4y^2)^3}. \end{pmatrix}$$

Luego,

$$\begin{aligned}
H_f(-1,0) &= \begin{pmatrix} \frac{3+2-2}{1} & 0 \\ 0 & -\frac{2(3+4-12+18)(1)}{1}. \end{pmatrix}\\
&= \begin{pmatrix} 3 & 0 \\ 0 & -\frac{26}{3}. \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
H_f(0,0) &= \begin{pmatrix} -2 & 0 \\ 0 & -\frac{2(18)}{3}. \end{pmatrix}\\
&= \begin{pmatrix} -2 & 0 \\ 0 & -12. \end{pmatrix}
\end{aligned}$$

$$
\begin{aligned}
H_f(2,0) &= \begin{pmatrix} 3(4)-2(2)-2 & 0 \\ 0 & -\dfrac{2(3(16)-4(8)-12(4)+18)}{3}. \end{pmatrix} \\
&= \begin{pmatrix} 6 & 0 \\ 0 & -\dfrac{56}{3}. \end{pmatrix}
\end{aligned}
$$

Dado que las matrices Hessianas asociadas son diagonales, entonces sus eigenvalores son los valores en la diagonal. Dichos valores son reales, entonces, por un teorema visto en clase, concluímos que
$(-1,0)$ es un **punto silla** (dado que tiene eigenvalores tanto positivos como negativos).
$(0,0)$ es un **mínimo** (dado que todos sus eigenvalores son negativos).
$(2,0)$ es un **punto silla** (dado que tiene eigenvalores tanto positivos como negativos).

Problema 4. Calcule el gradiente $\nabla f(x)$ y el Hessiano $\nabla^2 f(x)$ de la función de Rosenbrock

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

donde $x = [x_1, \ldots, x_N]^T \in \mathbb{R}^N$.
Solución. Primero notemos que para $2 \leq j \leq N-1$ se puede reescribir $f(x)$ como

$$
\begin{aligned}
f(x) &= \sum_{i=1}^{j-2} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2] + [100(x_j - x_{j-1}^2)^2 + (1 - x_{j-1})^2] \\
& \quad [100(x_{j+1} - x_j^2)^2 + (1 - x_j)^2] + \sum_{i=j+1}^{N-1} [100(x_{i-1} - x_i^2)^2 - (1 - x_i)^2].
\end{aligned}
$$

Notemos que el primer y útlimo término son cosntantes con respecto a x_j, entonces se anularán al calcular la correspondiente derivada parcial.
Para D_1 se tiene que

$$
\begin{aligned}
\frac{\partial}{\partial x_1} f(x) &= 100(2)(x_2 - x_1^2)(-2x_1) - 2(1 - x_1) \\
&= -400(x_2 - x_1^2)x_1 - 2(1 - x_1).
\end{aligned}
$$

Para $D_j \in \{2, \ldots, N-1\}$,

$$\begin{aligned}\frac{\partial}{\partial x_j} f(x) &= [200(x_j - x_{j-1}^2)] + [200(x_{j+1} - x_j^2)(-2x_j)] - [2(1-x_j)] \\ &= 200(x_j - x_{j-1}^2) - 400(x_{j+1} - x_j^2)x_j - 2(1-x_j)\end{aligned}$$

Para D_N,

$$\frac{\partial}{\partial x_N} f(x) = 200(x_N - x_{N-1}^2).$$

De lo anterior tenemos que

$$\nabla f(x) = \begin{bmatrix} -400(x_2 - x_1^2)x_1 - 2(1-x_1) \\ \vdots \\ 200(x_j - x_{j-1}^2) - 400(x_{j+1} - x_j^2)x_j - 2(1-x_j) \\ \vdots \\ 200(x_N - x_{N-1}^2). \end{bmatrix} \tag{3}$$

Nota. Para calcular el Hessiano,se debe considerar como casos particulares, aquellos que se apliquen a $\dfrac{\partial f(x)(x)}{\partial x_1}$ y $\dfrac{\partial f(x)(x)}{\partial x_N}$ (como se observó en el gradiente).

Para $D_{x_k x_1}$,

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_1} = \begin{cases} 1200x_1^2 - 400x_2 + 2 & si \quad k=1 \\ -400x_1 & si \quad k=2 \\ 0 & e.o.c \end{cases}$$

Para $D_{x_k x_j}$ con $j \in \{2, \ldots, N-1\}$,

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_j} = \begin{cases} 1200x_j^2 - 400x_2 + 2 & si \quad k=j \\ -400x_j & si \quad k=j+1 \\ -400x_{j-1} & si \quad k=j-1 \\ 0 & e.o.c \end{cases}$$

Para $D_{x_k x_N}$,

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_1} = \begin{cases} -400x_{N-1} & si \quad k=N-1 \\ 200 & si \quad k=N \\ 0 & e.o.c \end{cases}$$

Así que el Hessiano tiene la forma

$$H(x) = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 & 0 & 0 & \dots & 0 & 0 \\ -400x_1 & 1200x_2^2 - 400x_3 + 202 & -400x_2 & 0 \dots & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -400x_{N-1} & 200 \end{pmatrix}. \quad (4)$$

Problema 5. Demuestre, sin usar condiciones de optimalidad, que $f(x) > f(x^*)$ para todo $x \neq x^*$ si

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

$Q = Q^T > 0$ y $Qx^* = b$.

Demostración. Primero observemos que dados $x, y \in \mathbb{R}^n$ tenemos que, como $Q = Q^T$,

$$x^TQy = (Qy)^Tx = y^TQ^Tx = y^TQx.$$

Por lo que

$$x^TQy = y^TQx. \quad (5)$$

Ahora,

$$\begin{aligned} f(x) &= f(x^* + (x - x^*)) \\ &= \frac{1}{2}(x^* + (x - x^*))^TQ(x^* + (x - x^*)) - b^T(x^* + (x - x^*)) \\ &= \frac{1}{2}\left((x^*)^TQ + (x - x^*)^TQ\right)(x^* + (x - x^*)) - b^T(x^* + (x - x^*)) \\ &= \frac{1}{2}(x^*)^TQx^* + \frac{1}{2}(x^*)^TQ(x - x^*) + \frac{1}{2}(x - x^*)^TQx^* \\ &\quad + \frac{1}{2}(x - x^*)^TQ(x - x^*) - b^T(x^* + (x - x^*)) \\ &= \frac{1}{2}(x^*)^TQx^* + \underbrace{\frac{1}{2}(x - x^*)^TQx^*}_{(5)} + \frac{1}{2}(x - x^*)^TQx^* \\ &\quad + \frac{1}{2}(x - x^*)^TQ(x - x^*) - b^T(x^* + (x - x^*)) \end{aligned}$$

Ahora,

$$
\begin{aligned}
&= \frac{1}{2}(x^*)^T Q x^* + (x-x^*)^T Q x^* + \underbrace{\frac{1}{2}(x-x^*)^T Q (x-x^*)}_{\textbf{Definida Positiva}} \\
&\quad -b^T(x^* + (x-x^*)) \\
&> \frac{1}{2}(x^*)^T Q x^* + (x-x^*)^T Q x^* - b^T(x^* + (x-x^*)) \\
&= \frac{1}{2}(x^*)^T Q x^* - b^T x^* + (x-x^*)^T Q x^* - b^T(x-x^*) \\
&= f(x^*) + (x-x^*)^T Q x^* - (x-x^*)^T b \\
&= f(x^*) + (x-x^*)^T (Q x^* - b) \\
&= f(x^*) + (x-x^*)^T (\underbrace{b}_{\text{hipótesis}} - b) \\
&= f(x^*)
\end{aligned}
$$

Así que $f(x) > f(x^*)$. $\square$