Última actualización: 27 de mayo de 2022

Clase 34

Definición 1 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ una función. Diremos que f es derivable en $\overline{x}_0$ si existe una función lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que*

$$\lim_{\overline{x} \to \overline{x}_0} \frac{f(\overline{x}) - [L(\overline{x} - \overline{x}_0) + f(\overline{x}_0)]}{\|\overline{x} - \overline{x}_0\|} = 0.$$

Dado que la función L es única con esta propiedad (lo demostramos) la llamaremos ***la derivada de f en $\overline{x}_0$*** *y la denotaremos por* $\boldsymbol{Df(\overline{x}_0)}$.

Proposición 2 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$, $f : U \longrightarrow \mathbb{R}$ una función y $\overline{u} \in \mathbb{R}^n$ de norma 1. Si f es derivable en $\overline{x}_0$, entonces existe la derivada direccional de f en $\overline{x}_0$ en la dirección de $\overline{u}$ y*

$$D_{\overline{u}} f(\overline{x}_0) = Df(\overline{x}_0)(\overline{u}).$$

En esta clase, aprovechando que la derivada de una función de $\mathbb{R}^n$ en R en un punto es una función lineal y por tanto continua, introduciremos un vector de gran relevancia.

El Vector Gradiente

Comenzaremos esta sesión recordando que, para cada natural $n \geq 2$, el conjunto

$$S^{n-1} = \{\overline{u} \in \mathbb{R}^n \mid \|\overline{u}\| = 1\}$$

es un subconjunto cerrado y acotado (compacto) de $\mathbb{R}^n$, por lo que que cualquier función continua, en particular una lineal de $\mathbb{R}^n$ en $\mathbb{R}$, alcanza su valor mínimo y su valor máximo sobre S^{n-1}. Así, si $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ es una función derivable en $\overline{x}_0$, entonces deben existir $\overline{u}_1, \overline{u}_2 \in S^{n-1}$ tales que

$$Df(\overline{x}_0)(\overline{u}_1) \leq Df(\overline{x}_0)(\overline{u}) \leq Df(\overline{x}_0)(\overline{u}_2)$$

para todo $\overline{u} \in S^{n-1}$. Si además recordamos la relación entre la derivada y las derivadas direccionales, tenemos que

$$D_{\overline{u}_1} f(\overline{x}_0) \leq D_{\overline{u}} f(\overline{x}_0) \leq D_{\overline{u}_2} f(\overline{x}_0).$$

para todo $\overline{u} \in S^{n-1}$. Es decir, existen dos direcciones en las que la razón de cambio de la función f es mínima y otra en la que la razón de cambio es máxima. De hecho en la siguiente proposición mostraremos que $\overline{u}_1 = -\overline{u}_2$ y, si $Df(\overline{x}_0)$ no es la función lineal cero, entonces $\overline{u}_1$ y $\overline{u}_2$ son únicos.

Proposición 3 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ una función derivable en $\overline{x}_0$. Entonces existe $\overline{u}_0 \in S^{n-1}$ tal que*

$$D_{-\overline{u}_0} f(\overline{x}_0) \leq D_{\overline{u}} f(\overline{x}_0) \leq D_{\overline{u}_0} f(\overline{x}_0)$$

para todo $\overline{u} \in S^{n-1}$. Si además $Df(\overline{x}_0)$ no es la función constante cero, entonces $\overline{u}_0$ es único.

Demostración. Como $Df(\overline{x}_0)$ es una función continua y S^{n-1} es un conjunto cerrado y acotado, entonces existe $\overline{u}_0 \in S^{n-1}$ tal que

$$D_{\overline{u}} f(\overline{x}_0) \leq D_{\overline{u}_0} f(\overline{x}_0)$$

para todo $\overline{u} \in S^{n-1}$. Note que $-\overline{u}_0 \in S^{n-1}$, por lo que

$$\begin{aligned} D_{-\overline{u}_0} f(\overline{x}_0) &= Df(\overline{x}_0)(-\overline{u}_0) \\ &= -Df(\overline{x}_0)(\overline{u}_0) \\ &= -D_{\overline{u}_0} f(\overline{x}_0) \\ &\leq -D_{-\overline{u}} f(\overline{x}_0) \\ &= -Df(\overline{x}_0)(-\overline{u}) \\ &= D_{\overline{u}} f(\overline{x}_0). \end{aligned}$$

Por lo tanto,

$$D_{-\overline{u}_0} f(\overline{x}_0) \leq D_{\overline{u}} f(\overline{x}_0) \leq D_{\overline{u}_0} f(\overline{x}_0), \tag{1}$$

para todo $\overline{u} \in S^{n-1}$. Ahora supongamos que $Df(\overline{x}_0)$ no es la función constante cero y que existe $\overline{v}_0 \in S^{n-1}$, $\overline{v}_0 \neq \overline{u}_0$ tal que $Df(\overline{x}_0)(\overline{v}_0) = Df(\overline{x}_0)(\overline{u}_0)$. Note que si $\overline{v}_0 = -\overline{u}_0$, entonces $\overline{v}_0 + \overline{u}_0 = \overline{0}$, de donde

$$\begin{aligned} 0 &= Df(\overline{x}_0)(\overline{0}) \\ &= Df(\overline{x}_0)(\overline{v}_0 + \overline{u}_0) \\ &= Df(\overline{x}_0)(\overline{v}_0) + Df(\overline{x}_0)(\overline{u}_0) \\ &= Df(\overline{x}_0)(\overline{u}_0) + Df(\overline{x}_0)(\overline{u}_0) \\ &= 2Df(\overline{x}_0)(\overline{u}_0). \end{aligned}$$

Por lo que $Df(\overline{x}_0)(\overline{u}_0) = 0$ y de aquí que $Df(\overline{x}_0)(-\overline{u}_0) = 0$. Luego, de (1), se tiene que $Df(\overline{x}_0)(\overline{u}) = 0$ para todo $\overline{u} \in S^{n-1}$, lo que implica que $Df(\overline{x}_0)$ es la constante cero, lo que es una contradicción. Por lo tanto, $\overline{v}_0 \neq -\overline{u}_0$. Podemos entonces considerar el vector

$$\overline{u} = \frac{\overline{v}_0 + \overline{u}_0}{\|\overline{v}_0 + \overline{u}_0\|} \in S^{n-1}.$$

Así,

$$Df(\overline{x}_0)(\overline{u}) = \frac{1}{\|\overline{v}_0 + \overline{u}_0\|}\big(Df(\overline{x}_0)(\overline{v}_0) + Df(\overline{x}_0)(\overline{u}_0)\big) = \frac{2}{\|\overline{v}_0 + \overline{u}_0\|} Df(\overline{x}_0)(\overline{u}_0).$$

Ahora, como $\overline{u}_0, \overline{v}_0 \in S^{n-1}$ y $\overline{v}_0 \neq -\overline{u}_0, \overline{u}_0$, se tiene que $\|\overline{v}_0 + \overline{u}_0\| < \|\overline{v}_0\| + \|\overline{u}_0\| = 2$. Por lo que

$$Df(\overline{x}_0)(\overline{u}_0) < \frac{2}{\|\overline{v}_0 + \overline{u}_0\|} Df(\overline{x}_0)(\overline{u}_0) = Df(\overline{x}_0)(\overline{u}).$$

Lo que es una contradicción al hecho de que

$$D_{\overline{u}} f(\overline{x}_0) \leq D_{\overline{u}_0} f(\overline{x}_0)$$

para todo $\overline{u} \in S^{n-1}$. ■

La propiedad que caracteriza a $\overline{u}_0$ y su unicidad nos permiten definir un vector muy especial.

Definición 4 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ una función derivable en $\overline{x}_0$. Definimos el gradiente de f en $\overline{x}_0$, denotado por $\nabla f(\overline{x}_0)$, como el siguiente vector*

$$\nabla f(\overline{x}_0) = D_{\overline{u}_0} f(\overline{x}_0)\overline{u}_0,$$

donde $\overline{u}_0 \in \mathbb{R}^n$ es el vector de la proposición anterior.

Por supuesto que para calcular el vector gradiente es necesario conocer el vector $\overline{u}_0$, así que procederemos a calcularlo:
Sean $\beta = \{\overline{e}_1, ..., \overline{e}_n\}$ una base ortonormal de $\mathbb{R}^n$, $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ una función expresada en términos de las coordenadas $(x_1, ..., x_n)$ determinadas por la base β. Si f es derivable en $\overline{x}_0$ y $\overline{u} = (u_1, ..., u_n) \in S^{n-1}$, entonces podemos escribir

$$\begin{aligned} D_{\overline{u}} f(\overline{x}_0) &= Df(\overline{x}_0)(\overline{u}) \\ &= \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right) \cdot (u_1, u_2, ..., u_n) \\ &= \left\| \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right) \right\| \|(u_1, u_2, ..., u_n)\| \cos(\theta) \\ &= \left\| \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right) \right\| \cos(\theta). \end{aligned}$$

donde θ es el ángulo formado por los vectores $\left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right)$ y $(u_1, ..., u_n)$. Ahora si $Df(\overline{x}_0)$ no es la función constante cero, tenemos que el máximo valor de $Df(\overline{x}_0)(\overline{u})$, es decir, el máximo valor de $D_{\overline{u}} f(\overline{x}_0)$ se alcanza cuando $\cos(\theta) = 1$ y esto ocurre cuando $\theta = 0$. Por lo tanto el vector $\overline{u}_0$ debe ser paralelo al vector $\left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right)$ y como $\overline{u}_0$ también debe ser unitario, concluimos que

$$\overline{u}_0 = \frac{\left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right)}{\left\| \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right) \right\|}.$$

Una vez hallado $\overline{u}_0$, tenemos que

$$\begin{aligned} \nabla f(\overline{x}_0) &= D_{\overline{u}_0} f(\overline{x}_0)\overline{u}_0 \\ &= \left\| \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right) \right\| \frac{\left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right)}{\left\| \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right) \right\|} \\ &= \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right). \end{aligned}$$

Observación 5 *Sean $\beta = \{\overline{e}_1, ..., \overline{e}_n\}$ una base ortonormal de $\mathbb{R}^n$, $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ una función expresada en términos de las coordenadas $(x_1, ..., x_n)$ determinadas por la base β. Si f es derivable en $\overline{x}_0$, entonces*

$$\nabla f(\overline{x}_0) = \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right).$$

Una consecuencia de esta observación es que podemos usar (cuando f es derivable en $\overline{x}_0$) el gradiente de f en $\overline{x}_0$ para representar a $Df(\overline{x}_0)$, es decir,

$$Df(\overline{x}_0)(\overline{x}) = \nabla f(\overline{x}_0) \cdot \overline{x},$$

para cada $\overline{x} \in \mathbb{R}^n$.

Puede que se estén preguntando por qué definimos el gradiente como $\nabla f(\overline{x}_0) = D_{\overline{u}_0} f(\overline{x}_0)\overline{u}_0$, si pudimos definirlo desde el inicio como $\nabla f(\overline{x}_0) = \left(\frac{\partial f}{\partial x_1}(\overline{x}_0), ..., \frac{\partial f}{\partial x_n}(\overline{x}_0)\right)$, pero la intención fue resaltar que el gradiente de f en x_0 determina la dirección en la que f tiene la máxima razón de cambio. Intentaremos ilustrar esto con un ejemplo.

Ejemplo 6 *Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada en términos de las coordenadas de la base canónica por la siguiente regla de correspondencia*

$$f(x, y) = 8 - x^2 - y^2 + x - y$$

y $\overline{x}_0 = (2, 1)$. Demuestra que f es derivable en $\overline{x}_0$ y determina la dirección en que la derivada direccional de f en $\overline{x}_0$ alcanza su valor máximo.

Solución. Calcularemos el gradiente de f en $\overline{x}_0$ para luego construir la función lineal que propondremos como la derivada de f en $\overline{x}_0$. Note que

$$\frac{\partial f}{\partial x}(x, y) = -2x + 1 \quad \text{y} \quad \frac{\partial f}{\partial y}(x, y) = -2y - 1,$$

de donde

$$\begin{aligned} \nabla f(2, 1) &= \left(\frac{\partial f}{\partial x}(2, 1), \frac{\partial f}{\partial y}(2, 1)\right) \\ &= (-3, -3). \end{aligned}$$

Entonces afirmamos que $L : \mathbb{R}^2 \longrightarrow \mathbb{R}$, dada por

$$\begin{aligned} L(x, y) &= \nabla f(2, 1) \cdot (x, y) \\ &= (-3, -3) \cdot (x, y) \\ &= -3x - 3y, \end{aligned}$$

es la derivada de f en $\overline{x}_0$.
Primero note que

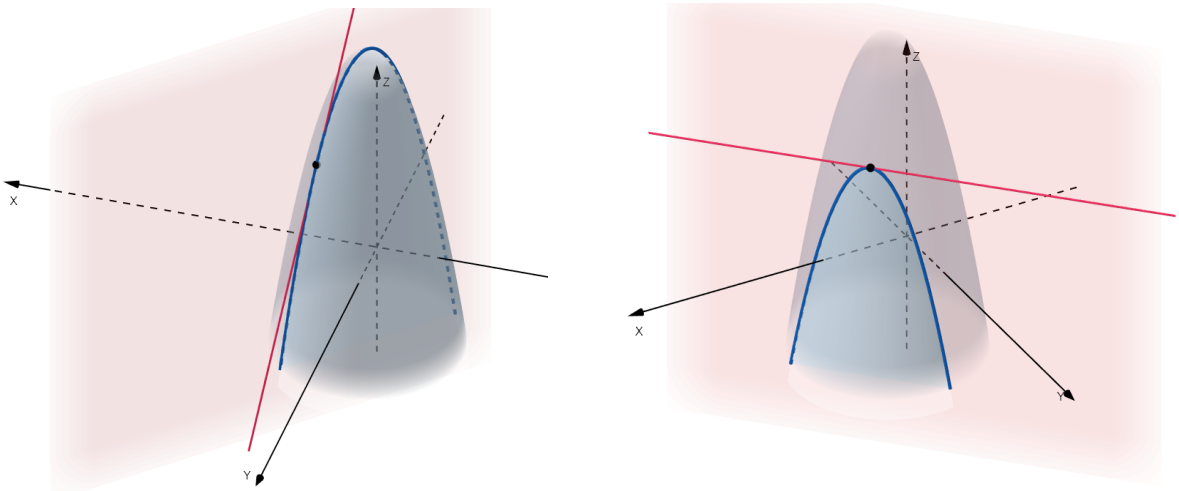
$$\begin{aligned} f(x, y) - [L(x - 2, y - 1) + f(2, 1)] &= 8 - x^2 - y^2 + x - y - [-3(x - 2) - 3(y - 1) + 4] \\ &= -x^2 + 4x - y^2 + 2y - 5 \\ &= -\left[(x - 2)^2 + (y - 1)^2\right] \\ &= -\|(x, y) - (2, 1)\|^2. \end{aligned}$$

Así,

$$\lim_{(x,y)\to(2,1)} \frac{f(x, y) - [L(x - 2, y - 1) + f(2, 1)]}{\|(x, y) - (2, 1)\|} = \lim_{(x,y)\to(2,1)} \frac{-\|(x, y) - (2, 1)\|^2}{\|(x, y) - (2, 1)\|} = 0$$

Por lo tanto, f es derivable en $(2,1)$ y la derivada $Df(2,1) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ está dada por

$$Df(2,1)(x,y) = -3x - 3y.$$



(a) En la dirección de $\nabla f(\overline{x}_0)$ se tiene la pendiente máxima.

(b) En la dirección de cualquier otro vector la pendiente es menor.

Figura 1

Finalmente, el vector

$$\begin{aligned} \overline{u}_0 &= \frac{\nabla f(2,1)}{\|\nabla f(2,1)\|} \\ &= \frac{(-3,-3)}{\|(-3,-3)\|} \\ &= \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right) \end{aligned}$$

determina la dirección en que la derivada direccional de f en $\overline{x}_0$ alcanza su valor máximo. ■

Intentemos ver el ejercicio anterior gráficamente, para ello consideremos la curva que resulta de la intersección de la gráfica de la función f y el plano perpendicular al plano XY cuya dirección está dada por el vector $\left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$ y que contiene al punto $(2,1,4) = (2,1,f(2,1))$. La recta tangente a dicha curva en el punto $(2,1,f(2,1))$ tiene la máxima pendiente, vea la Figura 1.