

2: Introducción al Cálculo Multivariante

Presentamos inicialmente algunos conceptos como los de función multivariante, derivada parcial, gradiente, así como la composición de funciones de varias variables y la regla de la cadena para su derivación. Todos ellos serán imprescindibles en el desarrollo subsiguiente, en el que se introduce un algoritmo de aprendizaje supervisado muy utilizado para entrenar redes neuronales. Nos valdremos para ello del método del gradiente descendente, en cuya aplicación efectiva se emplea la estrategia de retropropagación, indicada para el cálculo del gradiente de funciones mediante la aplicación recursiva de la regla de la cadena.

Definición 2.1

Definición 1. Una función $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^k$ lleva $\mathbf{x} \in \mathbb{R}^n$ en $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathbb{R}^k$. Se dice que \mathbb{R}^n es el dominio de \mathbf{f} y para cada $\mathbf{x} \in \mathbb{R}^n$, el vector $\mathbf{y} = \mathbf{f}(\mathbf{x})$ es la imagen de \mathbf{x} por \mathbf{f} . El conjunto de todas las imágenes se denomina también imagen de \mathbf{f} .

- $\mathbf{x} = (x_1, \dots, x_n)^T$.
- $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))^T$ y cada $f_i(\mathbf{x})$ es una función coordenada de \mathbf{f} .
- Si $k = 1$, $f(\mathbf{x}) \in \mathbb{R}$.

Observación 2.1

Los vectores $\mathbf{x} \in \mathbb{R}^n$ son vectores columna. Por comodidad, usamos la notación $\mathbf{x} = (x_1, \dots, x_n)^T$ para indicar el traspuesto del vector fila (x_1, \dots, x_n) . En ocasiones preferiremos la notación $\mathbf{f}(x, y)$, que resulta más simple que $\mathbf{f}((x, y)^T)$.

Ejemplo 2.1

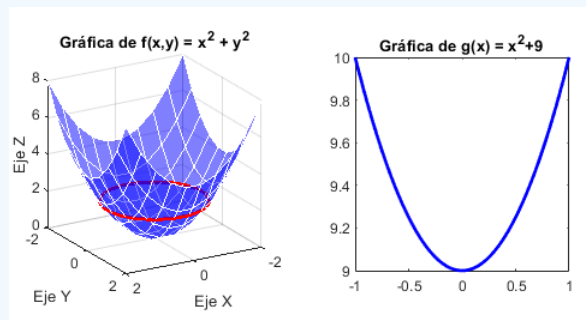
Funciones Lineales

Para números reales fijos a_1, a_2, a_3 , la función lineal f de dominio \mathbb{R}^3 dada por $f(x_1, x_2, x_3) = a_1x_1 + a_2x_2 + a_3x_3$ cumple que fijadas dos de las variables.

Por ejemplo $x_1 = -3$ y $x_2 = 5$, la función $g(x_3) = -3a_1 + 5a_2 + a_3x_3$ es una función lineal.

Ejemplo 2.2

Consideramos la función $f(x, y) = x^2 + y^2$, con dominio \mathbb{R}^2 e imagen en \mathbb{R} . El valor mínimo que toma la función es 0, en el punto $(0, 0)$. En todos los puntos (x, y) que estén sobre la circunferencia centrada en $(0, 0)$ y radio $r > 0$, la función f toma el valor r^2 . Si fijamos una de las variables, por ejemplo $y = 3$, obtenemos la función $g(x) = x^2 + 9$, cuya gráfica es una parábola. Podemos representar la función f , que es la superficie de \mathbb{R}^3 de la gráfica siguiente:

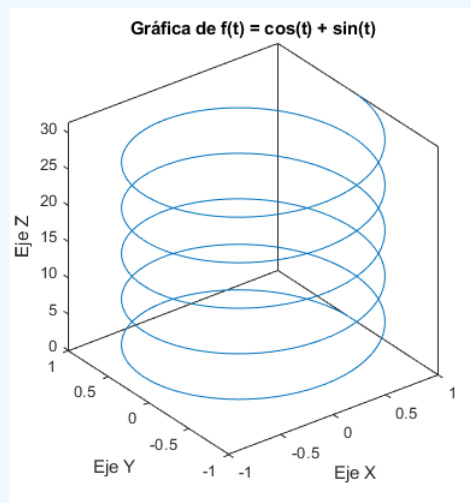


Observación 2.2

Cuando trabajemos con una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ y fijemos $n - 1$ variables, obtenemos una función con dominio e imagen en \mathbb{R} , que podemos analizar con las técnicas ya conocidas.

✓ Ejemplo 2.3

La función $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $\mathbf{f}(t) = (\cos(t), \sin(t))$ viene definida por las funciones coordenadas $f_1(t) = \cos(t)$ y $f_2(t) = \sin(t)$. La gráfica de \mathbf{f} es una hélice contenida en \mathbb{R}^3 . La imagen de \mathbf{f} es el conjunto $C := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, ya que para todo $t \in \mathbb{R}$ se cumple que $\cos^2(t) + \sin^2(t) = 1$.



✓ Ejemplo 2.4

Nos encontraremos con la función sigmoide $\sigma(x) = \frac{1}{1 + e^{-x}}$, que es derivable y (¡Compruébalo!) su derivada es $\sigma'(x) = \sigma(x)(1 - \sigma(x))$

Revisamos para una función $f: \mathbb{R} \rightarrow \mathbb{R}$ (es decir: $n = k = 1$), la noción de derivada en un punto x .

✎ Definición 2.2

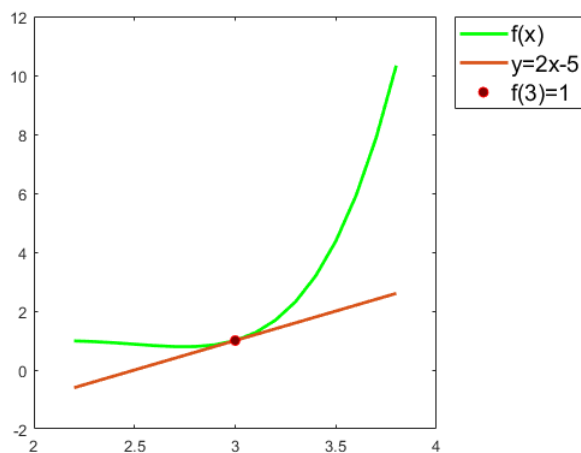
La derivada de f en x , que notamos $f'(x)$, viene dada (¡Cuando existe!) por:

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (2.1)$$

Recuerda que $f'(x)$ es la pendiente de la recta que más se asemeja a la gráfica de la función entre las que pasan por el punto $(x, f(x))$. Como la recta que une los puntos (a, b) y (c, d) tiene pendiente $\frac{d-b}{c-a}$, la recta que une $(x, f(x))$ con $(x+h, f(x+h))$ tiene como pendiente $\frac{f(x+h) - f(x)}{h}$, por lo que $f'(x)$ es el límite de las pendientes de estas rectas.

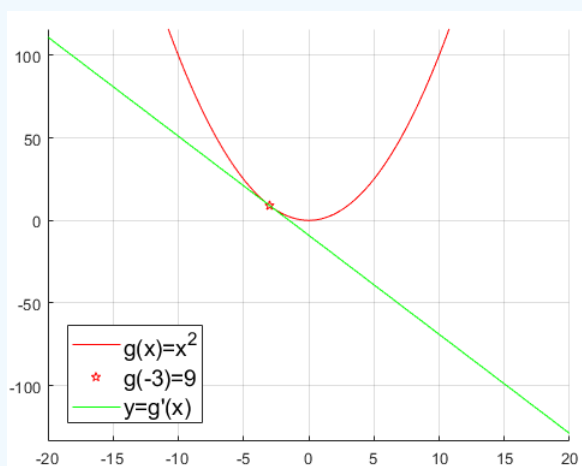
✓ Ejemplo 2.5

Si $f: \mathbb{R} \rightarrow \mathbb{R}$ cumple que $f(3) = 1$ y $f'(3) = 2$, el valor $f(3)$ no puede ser mínimo (ni máximo), ya que la gráfica cerca de 3 es similar a la recta $y = f(3) + f'(3)(x - 3) = 1 + 2(x - 3)$; es decir, $y = 2x - 5$.



✓ Ejemplo 2.6

Si pensamos en la función $g(x) = x^2$ y $x = -3$, como $g'(-3) = -6$, a la derecha de -3 la función tomará valores menores que $g(-3) = 9$. El valor mínimo de la función, que es $g(0) = 0$, se alcanza en un punto donde la derivada es nula.



Para comprender cómo varía una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ que depende de varias variables empezamos fijando todas las variables salvo una. Consideramos la función resultante (que tiene dominio e imagen en \mathbb{R}). Para el ejemplo $f(x, y) = x^2 + y^2$, fijada la variable y obtenemos una función $g_1(x) = x^2 + y^2$, que es decreciente para $x < 0$ y creciente para $x > 0$, ver la figura del ejemplo Ejemplo 2 (se da lo análogo para $g_2(y) = x^2 + y^2$, con x fijo). Nos centramos, por ejemplo, en el punto del plano $(-3, 4)$, para el que $f(-3, 4) = 25$, queremos encontrar valores h_1 y h_2 para los que $f(-3 + h_1, 4 + h_2) < 25$. Podemos conseguirlo siempre $-3 < -3 + h_1 < 0$ y $0 < 4 + h_2 < 4$. Mientras la derivada $g'_1(-3) < 0$ nos alienta a elegir $h_1 > 0$, la derivada $g'_2(4) > 0$ nos impulsa a elegir $h_2 < 0$. La misma estrategia podrá ser utilizada para buscar mínimos de funciones escalares acudiendo a las nociones de derivada parcial y gradiente de la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$. La derivada parcial con respecto a una variable x_j reflejará la tasa de variación de f en la dirección de ese eje.

Definición 2.3

La derivada parcial de la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ con respecto a la variable x_j en $\mathbf{x} = (x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n)^T$, que denotamos $\frac{\partial}{\partial x_j} f(\mathbf{x})$, viene dada por

$$\lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{j-1}, x_j + h, x_{j+1}, \dots, x_n) - f(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n)}{h} \quad (2.2)$$

✓ Ejemplo 2.7

La función $f(x, y) = x^2 + y^2$ tiene como derivadas parciales $\frac{\partial}{\partial x} f(x, y) = 2x$ y $\frac{\partial}{\partial y} f(x, y) = 2y$. La función $f(x_1, x_2, x_3) = a_1 x_1 + a_2 x_2 + a_3 x_3$ tiene como derivada parcial con respecto a x_j el número a_j , para $1 \leq j \leq 3$. La función $f(x, y) = \sin(xy^3) + x^2 y$ tiene como derivadas parciales $\frac{\partial}{\partial x} f(x, y) = y^3 \cdot \cos(xy^3) + 2xy$ y $\frac{\partial}{\partial y} f(x, y) = 3xy^2 \cos(xy^3) + x^2$.

✎ Definición 2.4

Dada una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ con derivadas parciales $\frac{\partial}{\partial x_j} f(\mathbf{x})$, el gradiente de f se denota por $\nabla f(\mathbf{x})$ y tiene por coordenadas las funciones $\frac{\partial}{\partial x_j} f(\mathbf{x})$, con $1 \leq j \leq n$.

✓ Ejemplo 2.8

La función $f(x, y) = x^2 + y^2$ tiene como gradiente la función $\nabla f(x, y) = (2x, 2y)^T$. La función $f(x_1, x_2, x_3) = a_1 x_1 + a_2 x_2 + a_3 x_3$ tiene como gradiente el vector constante $\nabla f(\mathbf{x}) = (a_1, a_2, a_3)^T$.

Observación 2.3

Como veremos más adelante, el hecho de que la dirección del gradiente $\nabla f(\mathbf{x})$ sea la de máximo crecimiento de la función f en cada punto \mathbf{x} y la dirección dada por su opuesto, $-\nabla f(\mathbf{x})$, la de mayor decrecimiento, permite establecer una estrategia para minimizar la función escalar $f(\mathbf{x})$. Para la función $f(x, y) = x^2 + y^2$ en el punto $(-3, 4)$ hemos elegido valores (h_1, h_2) de signos opuestos al gradiente $(-6, 8)$ para obtener puntos $(-3 + h_1, 4 + h_2)$ en los que f toma valores menores que $f(-3, 4) = 25$.

Si trabajamos con una función $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^k$ de k coordenadas f_1, f_2, \dots, f_k , cuando existan las n derivadas parciales de cada una de ellas podemos considerar la función matricial de k filas y n columnas formada por los k gradientes ∇f_k . Se trata de la *matriz jacobiana* de la función \mathbf{f} .

This page titled [2: Introducción al Cálculo Multivariable](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Joaquín López Herraiz](#).