Diferenciciación en $\mathbb{R}^n$: Derivadas de orden superior

R. Álvarez-Nodarse
Universidad de Sevilla

Supongamos que $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, A es un abierto de $\mathbb{R}^n$ tiene derivadas parciales $D_i f = \dfrac{\partial f(x)}{\partial x_i}$ en A, $i = 1, \ldots, n$.

Derivadas de orden superior

Supongamos que $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, A es un abierto de $\mathbb{R}^n$ tiene derivadas parciales $D_i f = \dfrac{\partial f(x)}{\partial x_i}$ en A, $i = 1, \ldots, n$. Supongamos que dichas derivadas parciales $D_i f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ admiten a su vez derivadas parciales $D_j(\cdot)$ en A. Dichas derivadas parciales se denominan derivadas parciales de segundo orden y se denotan por

$$D_j(D_i f)(x) = D_{j,i} f(x) = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}, \quad i, j = 1, 2, \ldots, n.$$

Derivadas de orden superior

Supongamos que $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, A es un abierto de $\mathbb{R}^n$ tiene derivadas parciales $D_i f = \dfrac{\partial f(x)}{\partial x_i}$ en A, $i = 1, \dots, n$. Supongamos que dichas derivadas parciales $D_i f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ admiten a su vez derivadas parciales $D_j(\cdot)$ en A. Dichas derivadas parciales se denominan derivadas parciales de segundo orden y se denotan por

$$D_j(D_i f)(x) = D_{j,i} f(x) = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}, \quad i, j = 1, 2, \dots, n.$$

Si las funciones $D_{j,i} f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ admiten derivadas parciales entonces podemos definir las derivadas parciales de orden 3

$$D_k(D_j(D_i f))(x) = D_{k,j,i} f(x) = \frac{\partial^3 f(x)}{\partial x_k \partial x_j \partial x_i}, \quad i, j, k = 1, 2, \dots, n.$$

Y así, sucesivamente.

Una pregunta natural es cuando las derivadas cruzadas son iguales, i.e.,

$$\frac{\partial^2 f(x)}{\partial x_j \partial x_i} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.$$

Veamos un ejemplo. Sea $f : \mathbb{R}^2 \mapsto \mathbb{R}$ definida por

$$f(x,y) = \begin{cases} x^2 \arctan \dfrac{y}{x} - y^2 \arctan \dfrac{x}{y}, & \text{si } xy \neq 0, \\ 0, & \text{si } xy = 0. \end{cases}$$

Theorem (Schwarz)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ con A abierto, y sea $x_0 \in A$. Si en A existen las derivadas parciales $\dfrac{\partial f(x)}{\partial x_i}$, $\dfrac{\partial f(x)}{\partial x_j}$ y $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ y la derivada $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ es continua en x_0, entonces en A existe la derivada $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$ y $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i} = \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$.

Derivadas de orden superior

Theorem (Schwarz)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ *con* A *abierto, y sea* $x_0 \in A$. *Si en* A *existen las derivadas parciales* $\dfrac{\partial f(x)}{\partial x_i}$, $\dfrac{\partial f(x)}{\partial x_j}$ *y* $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ *y la derivada* $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ *es continua en* x_0, *entonces en* A *existe la derivada* $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$ *y* $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i} = \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$.

Corolario (Bonnet)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ *con* A *abierto, y sea* $a \in A$ *tal que existen las derivadas parciales* $\dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ *y* $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$ *en un entorno de* $a \in A$ *y ambas son continuas en* a. *Entonces* $\dfrac{\partial^2 f(a)}{\partial x_j \partial x_i} = \dfrac{\partial^2 f(a)}{\partial x_i \partial x_j}$.

Derivadas de orden superior

Theorem (Heffter-Young)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, A abierto y sea $a \in A$. Supongamos que existen las derivadas parciales $\dfrac{\partial f(x)}{\partial x_i}$, y $\dfrac{\partial f(x)}{\partial x_j}$ en un entorno de a y son diferenciables en a. Entonces $\dfrac{\partial^2 f(a)}{\partial x_j \partial x_i} = \dfrac{\partial^2 f(a)}{\partial x_i \partial x_j}$.

Theorem (Heffter-Young)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, A abierto y sea $a \in A$. Supongamos que existen las derivadas parciales $\dfrac{\partial f(x)}{\partial x_i}$, y $\dfrac{\partial f(x)}{\partial x_j}$ en un entorno de a y son diferenciables en a. Entonces $\dfrac{\partial^2 f(a)}{\partial x_j \partial x_i} = \dfrac{\partial^2 f(a)}{\partial x_i \partial x_j}$.

Tanto el Teorema de Schwarz como el de Heffter-Young dan condiciones suficientes. Para comprobarlo escojamos la función

$$f(x,y) = \begin{cases} x^2 y \cos\dfrac{1}{x}, & \text{si } x \neq 0, \\ 0, & \text{si } x = 0. \end{cases}$$

Esta función cumple con las condiciones de ambos teoremas en todo $\mathbb{R}^2 \setminus \{(0,y), \forall y \in \mathbb{R}\}$. Veamos que ocurre $(0,0)$.

Definición

Diremos que $f \in C^{(k)}(A)$ *si* f *admite todas las derivadas parciales hasta orden* k *y estas son continuas en* A.

Definición

Diremos que $f \in C^{(k)}(A)$ *si* f *admite todas las derivadas parciales hasta orden* k *y estas son continuas en* A.

Supongamos que la función $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en A. Entonces podemos definir la función derivada de f en A, $Df : A \subset \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, donde $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ denota al espacio de todas las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Derivadas de orden superior

Definición

Diremos que $f \in C^{(k)}(A)$ si f admite todas las derivadas parciales hasta orden k y estas son continuas en A.

Supongamos que la función $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en A. Entonces podemos definir la función derivada de f en A, $Df : A \subset \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, donde $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ denota al espacio de todas las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Diremos que f es dos veces diferenciable en un punto $a \in A$ si la función Df anterior es diferenciable en a y denotaremos a la derivada segunda de f en a por $D^2 f(a)$. Nótese que de lo anterior se sigue que $D^2 f(a)$ es una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Definición

Diremos que $f \in C^{(k)}(A)$ *si* f *admite todas las derivadas parciales hasta orden* k *y estas son continuas en* A.

Supongamos que la función $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en A. Entonces podemos definir la función derivada de f en A, $Df : A \subset \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, donde $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ denota al espacio de todas las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Diremos que f es dos veces diferenciable en un punto $a \in A$ si la función Df anterior es diferenciable en a y denotaremos a la derivada segunda de f en a por $D^2 f(a)$. Nótese que de lo anterior se sigue que $D^2 f(a)$ es una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Este procedimiento se puede extender obteniéndose la derivada tercera $D^3 f(a)$ que es una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$, y así sucesivamente.

Veamos que ocurre en el caso especial (y de gran importancia en las aplicaciones) de la segunda derivada. Asumiremos que $f \in C^2(A)$ y $a \in A$.

Como ya hemos mencionado $D^2 f(a)$ es una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, i.e., $D^2 f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$, pero el espacio $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ es isométrico al espacio de las aplicaciones bilineales $\mathcal{L}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$.

Por tanto, dados $x, y \in \mathbb{R}^n$ tenemos $D^2 f(a)(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ y $\big(D^2 f(a)(x)\big)(y) \in \mathbb{R}^m$. Es decir, $D^2 f(a)$ se puede considerar como la aplicación bilineal $D^2 f(a)$ definida en $\mathbb{R}^n \times \mathbb{R}^n$ por $D^2 f(a)(x, y) = \big(D^2 f(a)(x)\big)(y)$.

En el caso particular $f : A \subset \mathbb{R}^n \to \mathbb{R}$, la derivada segunda en un punto se puede representar mediante una matriz cuadrada $n \times n$, que se denomina **matriz hessiana**.
En efecto, si f es dos veces diferenciable en a, hemos visto que $D^2 f(a)$ puede ser interpretada como una aplicación bilineal $B(x, y)$ de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$. Pero las aplicaciones bilineales $B(x, y)$ de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$ se identifican con las matrices cuadradas $n \times n$ mediante la expresión $B(x, y) = xBy^t$. Puesto que $D^2 f(a)$ se obtiene derivando la función Df entonces la matriz asociada a $D^2 f(a)$ tiene por entradas las segundas derivadas parciales de f:

$$D^2 f(a) = \begin{pmatrix} D_{11} f(a) & \cdots & D_{n1} f(a) \\ \vdots & \ddots & \vdots \\ D_{1n} f(a) & \cdots & D_{nn} f(a) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_n} \end{pmatrix} := H_f(a).$$

Es conveniente mencionar que al ser $f \in C^2(A)$, todas las derivadas cruzadas son iguales.

Por inducción es posible probar que si f es k veces diferenciable en a entonces la derivada k-ésima de f aplicada a un vector $h \in \mathbb{R}^n$ se expresa por

$$D^k f(a)(h) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f(a)}{\partial x_{i_1} \cdots \partial x_{i_k}} h_{i_1} \cdots h_{i_k} =$$

$$= \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^k f(a),$$

donde hemos usado la notación $D^k f(a)(h) := D^k f(a)(h, h, \ldots, h)$ (recuérdese que $D^k f(a)$ es una aplicación multilineal (k-lineal concretamente).

Ejemplo: Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $f \in\in C^k(A)$. Sea la función $l : \mathbb{R} \mapsto \mathbb{R}^n$, $l(x) = x + th$, $t \in \mathbb{R}$, $h \in \mathbb{R}^n$. Definamos la función compuesta $\phi : \mathbb{R} \mapsto \mathbb{R}$, $\phi(t) = (f \circ l)(t)$. Como f y l son diferenciables con todas sus derivadas hasta orden k continuas entonces $\phi \in C^k([0,1])$. Calculemos las derivadas sucesivas de ϕ,

$$\phi'(t) = Df(x+ht) \cdot h = \sum_{i=1}^{n} \frac{\partial f(x+ht)}{\partial x_i} h_i$$

$$\phi''(t) = \frac{d}{dt}\left[\sum_{i=1}^{n} \frac{\partial f(x+ht)}{\partial x_i} h_i\right] = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[\sum_{i=1}^{n} \frac{\partial f(x+ht)}{\partial x_i} h_i\right] h_j =$$

$$\sum_{i,j=1}^{n} \frac{\partial^2 f(x+ht)}{\partial x_j \partial x_i} h_i h_j$$

$$\vdots$$

$$\phi^{(k)}(t) = \sum_{i_1,\cdots,i_k=1}^{n} \frac{\partial^k f(x+ht)}{\partial x_{i_1}\cdots\partial x_{i_k}} h_{i_1}\cdots h_{i_k} =$$

$$\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^k f(x+ht).$$

El cálculo de la derivada k-ésima arbitraria es muy engorroso, es por ello conveniente usar un paquete de cálculo simbólico (por ejemplo MAXIMA CAS).

$$\vdots$$

$$\phi^{(k)}(t) = \sum_{i_1,\cdots,i_k=1}^{n} \frac{\partial^k f(x+ht)}{\partial x_{i_1}\cdots\partial x_{i_k}} h_{i_1}\cdots h_{i_k} =$$

$$\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^k f(x+ht).$$

El cálculo de la derivada k-ésima arbitraria es muy engorroso, es por ello conveniente usar un paquete de cálculo simbólico (por ejemplo MAXIMA CAS).

Definición

Dado un $a \in A$ y $h \in \mathbb{R}^n$ definiremos al intervalo (cerrado) $[a, a+h]$ como el conjunto (intervalo) $[a_1, a_1+h_1] \times [a_2, a_2+h_2] \times \cdots \times [a_n, a_n+h_n]$.

Theorem (de Taylor con resto de Lagrange)

Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, $f \in C^k(A)$; $a \in A$; $[a, a+h] \subset A$ *para cierto* $h \neq 0$:

$$f(a+h) = f(a) + \sum_{l=1}^{k-1} \frac{1}{l!} D^l f(a)(h) + r_k(a,h),$$

$$r_k(a,h) = \frac{1}{k!} D^k f(a+\xi h)(h), \quad \xi \in (0,1).$$

Theorem (de Taylor con resto de Lagrange)

Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, $f \in C^k(A)$; $a \in A$; $[a, a+h] \subset A$ para cierto $h \neq 0$:

$$f(a+h) = f(a) + \sum_{l=1}^{k-1} \frac{1}{l!} D^l f(a)(h) + r_k(a,h),$$

$$r_k(a,h) = \frac{1}{k!} D^k f(a+\xi h)(h), \quad \xi \in (0,1).$$

Corolario (Teorema local de Taylor)

En las mismas condiciones del Teorema anterior:

$$f(a+h) = f(a) + \sum_{l=1}^{k} \frac{1}{l!} D^l f(a)(h) + o(\|h\|^k).$$

El corolario anterior nos indica otra manera de entender la diferenciabilidad en $\mathbb{R}^n$. Por sencillez, lo mostraremos en el caso de una función dos veces diferenciable. Si f tiene derivadas parciales de orden dos y estas son continuas entonces

$$f(a+h) - f(a) - Df(a)(h) - \frac{1}{2}D^2 f(a)(h) = o(\|h\|^2), \qquad (*)$$

donde $D^f(a)(h)$ es la forma bilineal

$$D^2 f(a)(h) = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \frac{\partial^2 f(a)}{\partial x_{i_1}\partial x_{i_2}} h_{i_1} h_{i_2} = h^T H_f(a) h.$$

Así pues f es dos veces diferenciable si existen la apliación lineal $Df(a)$ y la bilineal $D^2 f(a)$ tales que (*) sea cierta.

Lo anterior es fácilmente generalizable para cualquier $k \geq 3$.

Lo anterior nos indica que podemos restringirnos por simplicidad al caso cuando las funciones $f \in C^k(A)$.

Lo anterior es fácilmente generalizable para cualquier $k \geq 3$.

Lo anterior nos indica que podemos restringirnos por simplicidad al caso cuando las funciones $f \in C^k(A)$.

Así pues diremos que $f : A \in \mathbb{R}^n \mapsto \mathbb{R}$, es k veces diferenciable en a si f es $C^k(A)$ siendo A un abierto tal que $a \in A$, de forma que, por el teorema de Taylor tenemos asegurado que f es k veces diferenciable en A en el sentido antes explicado.

Lo anterior es fácilmente generalizable para cualquier $k \geq 3$.

Lo anterior nos indica que podemos restringirnos por simplicidad al caso cuando las funciones $f \in C^k(A)$.

Así pues diremos que $f : A \in \mathbb{R}^n \mapsto \mathbb{R}$, es k veces diferenciable en a si f es $C^k(A)$ siendo A un abierto tal que $a \in A$, de forma que, por el teorema de Taylor tenemos asegurado que f es k veces diferenciable en A en el sentido antes explicado.

Veamos algunos ejemplos.