Tema 6

Diferenciabilidad

Iniciamos el estudio del cálculo diferencial para funciones de varias variables en un contexto muy general, definiendo la diferenciabilidad de funciones entre espacios normados arbitrarios. La noción conocida para funciones reales de variable real que se generaliza directamente, no es la de derivada, sino la de *diferencial*, equivalente a ella. Veremos que la diferenciabilidad es una propiedad local, que implica la continuidad. Entre las reglas de diferenciación de funciones entre espacios normados, destacaremos la *regla de la cadena*.

6.1. Motivación

Recordemos el concepto de derivada para funciones reales de variable real, y su significado analítico, intentando ver la forma de generalizarlo al caso de una función entre dos espacios normados arbitrarios, digamos X e Y. Si A es un subconjunto no vacío de $\mathbb{R}$, sabemos que una función $f : A \to \mathbb{R}$ es derivable en un punto $a \in A \cap A'$, cuando existe $\lambda \in \mathbb{R}$ tal que

$$\boxed{\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lambda} \tag{1}$$

en cuyo caso $\lambda = f'(a)$ es la *derivada* de f en a. De entrada, cuando f esté definida en un subconjunto del espacio normado X, el cociente de incrementos que aparece en (1) no tiene sentido, simplemente no podemos dividir por el vector $x - a$. Para resolver este problema con el denominador, reescribimos (1) en la forma

$$\boxed{\lim_{x \to a} \frac{f(x) - f(a) - \lambda(x - a)}{x - a} = 0}$$

lo que a su vez se puede expresar de dos formas equivalentes:

$$\boxed{\lim_{x \to a} \frac{|f(x) - f(a) - \lambda(x - a)|}{|x - a|} = 0} \quad \text{o bien,} \quad \boxed{\lim_{x \to a} \frac{f(x) - f(a) - \lambda(x - a)}{|x - a|} = 0} \tag{2}$$

Esto resuelve el problema que teníamos con el denominador de (1): para $a, x \in X$, bastará escribir $\|x - a\|$ en lugar de $|x - a|$.

Recordemos que las igualdades (2) son las permiten entender intuitivamente el significado analítico de la derivada: la función $P:\mathbb{R}\to\mathbb{R}$ dada por $P(x)=f(a)+\lambda(x-a)$ para todo $x\in\mathbb{R}$, que viene definida por un polinomio de primer orden, es una buena aproximación de f cerca del punto a, puesto que la diferencia $f(x)-P(x)$ tiende a cero en el punto a, "mucho más rápidamente" que la diferencia $x-a$.

Observemos que el numerador de cualquiera de los límites que aparecen en (2) nos plantea un nuevo problema. En el caso general, tendremos $f(x)-f(a)\in Y$ mientras que $\lambda(x-a)\in X$. La solución está en considerar, en vez de las nociones de derivabilidad y derivada de f en a, las nociones equivalentes de diferenciabilidad y diferencial, que enseguida recordaremos. Lo que subyace a esta equivalencia es la identificación total del espacio normado $\mathbb{R}$ con el espacio normado $L(\mathbb{R},\mathbb{R})$ de todas las aplicaciones lineales continuas de $\mathbb{R}$ en $\mathbb{R}$, así que aclaremos dicha identificación.

Para cada $\alpha\in\mathbb{R}$ consideramos la aplicación lineal de $\mathbb{R}$ en $\mathbb{R}$ que consiste en multiplicar por α, esto es, la aplicación $T_\alpha\in L(\mathbb{R},\mathbb{R})$ dada por $T_\alpha(x)=\alpha x$ para todo $x\in\mathbb{R}$. Es claro que, definiendo $\Phi(\alpha)=T_\alpha$ para todo $\alpha\in\mathbb{R}$, obtenemos una aplicación lineal $\Phi:\mathbb{R}\to L(\mathbb{R},\mathbb{R})$. Recordando la norma de $L(\mathbb{R},\mathbb{R})$, es evidente que $\|T_\alpha\|=|\alpha|$ para todo $\alpha\in\mathbb{R}$, así que Φ conserva la norma: $\|\Phi(\alpha)\|=|\alpha|$ para todo $\alpha\in\mathbb{R}$. En particular, Φ es inyectiva, pero también es sobreyectiva, pues para toda $T\in L(\mathbb{R},\mathbb{R})$, basta tomar $\alpha=T(1)$ para tener $T=T_\alpha=\Phi(\alpha)$. Por tanto, Φ nos permite identificar totalmente los espacios normados $\mathbb{R}$ y $L(\mathbb{R},\mathbb{R})$.

Volviendo a nuestra función f derivable en el punto a, cuando la anterior identificación se aplica a la derivada $\lambda=f'(a)$, se obtiene una aplicación $T_\lambda\in L(\mathbb{R},\mathbb{R})$, que será la diferencial de f en a, y diremos que f es diferenciable en el punto a. Es claro que las igualdades (2) se expresan equivalentemente en términos de la diferencial, sin más que escribir $T_\lambda(x-a)$ en lugar de $\lambda(x-a)$. Más formalmente, decimos que $f:A\to\mathbb{R}$ es *diferenciable* en el punto $a\in A\cap A'$ cuando existe $T\in L(\mathbb{R},\mathbb{R})$ verificando que

$$\boxed{\lim_{x\to a}\frac{|f(x)-f(a)-T(x-a)|}{|x-a|}=0}\,, \quad \text{o bien,} \quad \boxed{\lim_{x\to a}\frac{f(x)-f(a)-T(x-a)}{|x-a|}=0} \qquad (3)$$

en cuyo caso T es única, y la llamamos *diferencial* de f en a, que se denota por $Df(a)$. Está claro que si $\lambda\in\mathbb{R}$ verifica (1), o equivalentemente (2), entonces tomando $T=T_\lambda\in L(\mathbb{R},\mathbb{R})$, tenemos (3). Recíprocamente, si $T\in L(\mathbb{R},\mathbb{R})$ verifica (3), basta tomar $\lambda=T(1)\in\mathbb{R}$ para obtener (2), luego también (1). Así pues, *f es diferenciable en a si, y sólo si, es derivable en a*, en cuyo caso se tiene $f'(a)=Df(a)(1)$, o lo que es lo mismo, $Df(a)(x)=f'(a)x$ para todo $x\in\mathbb{R}$. Queda claro que, para funciones reales de variable real, la distinción entre derivada y diferencial es sólo una cuestión de matiz, hablamos de la derivada cuando pensamos en el número real $f'(a)$, mientras que hablamos de la diferencial cuando pensamos en la aplicación lineal $Df(a)$, pero la identificación total de $\mathbb{R}$ con $L(\mathbb{R},\mathbb{R})$ hace que la derivada y la diferencial sean conceptos equivalentes.

Sin embargo, en el caso general, la distinción entre derivada y diferencial se vuelve crucial. Las igualdades (2) no tenían sentido, pero las de (3) sí lo tienen, con sólo sustituir $T\in L(\mathbb{R},\mathbb{R})$ por $T\in L(X,Y)$. Queda así motivada la noción general de funcion diferenciable.

¿ $T_\alpha(x) = \alpha x$ lineal? $\Leftrightarrow \forall a,b \in \mathbb{R}, \forall x,y \in X \quad T_\alpha(ax+by) = aT_\alpha(x) + bT_\alpha(y)$

$$\left.\begin{array}{l} T_\alpha(ax+by) = \alpha(ax+by) = \alpha a x + \alpha b y \\ a \cdot T_\alpha(x) + b \cdot T_\alpha(y) = a\alpha x + b\alpha y \end{array}\right\} \text{Coinciden } \checkmark$$

¿ $\Phi(\alpha) = T_\alpha$ lineal?

$$\Phi(a\alpha + b\beta) = T_{a\alpha+b\beta} = a \cdot T_\alpha + b \cdot T_\beta \quad \checkmark$$

Como $T_{a\alpha+b\beta}(x) = (a\alpha + b\beta)\, x = a(\alpha x) + b(\beta x) = a \cdot T_\alpha(x) + b \cdot T_\beta(x)$

$x = a + tv \quad t \in \mathbb{R}$ con $|t| < \frac{\delta}{\|v\|} \Rightarrow x - a = tv$

$$\|x - a\| = \|tv\| = |t| \cdot \|v\| < \frac{\delta}{\|v\|} \cdot \|v\| = \delta$$

$$\Downarrow$$

$$x \in \mathring{A} \subset A$$

6.2. Funciones diferenciables

En todo lo que sigue, X e Y serán espacios normados arbitrarios, si bien se descartan siempre los casos triviales $X = \{0\}$ e $Y = \{0\}$. Fijamos una función $f : A \to Y$, donde A es un subconjunto no vacío de X, y un punto $a \in A^\circ$.

Pues bien, se dice que f es **diferenciable** en el punto a cuando existe una aplicación lineal y continua $T \in L(X,Y)$ verificando las siguientes condiciones, que son equivalentes:

$$\boxed{\lim_{x\to a} \frac{\|f(x) - f(a) - T(x-a)\|}{\|x-a\|} = 0\,, \quad \text{o bien,} \quad \lim_{x\to a} \frac{f(x) - f(a) - T(x-a)}{\|x-a\|} = 0} \qquad (4)$$

En ambas igualdades tenemos el límite en el punto a de una función definida en $A \setminus \{a\}$, que tiene sentido, puesto que a es punto de acumulación de $A \setminus \{a\}$. La primera igualdad es conceptualmente más sencilla, pues dicha función toma valores en $\mathbb{R}$, mientras en la segunda tenemos el límite de una función con valores en Y. Conviene disponer de ambas, para usar una u otra según convenga.

También podemos trasladar el cálculo al origen, usando el conjunto $B = \{x - a : x \in A\}$, que verifica $0 \in B^\circ$. Mediante el cambio de variable $x = a + h \in A$ con $h \in B$, teniendo en cuenta que $x \to a$ cuando $h \to 0$ y que $x \neq a$ para $h \neq 0$, deducimos de (4) que

$$\boxed{\lim_{h\to 0} \frac{\|f(a+h) - f(a) - T(h)\|}{\|h\|} = 0} \quad \text{o bien,} \quad \boxed{\lim_{h\to 0} \frac{f(a+h) - f(a) - T(h)}{\|h\|} = 0} \qquad (5)$$

Recíprocamente, de (5) se deduce (4) usando el cambio de variable $h = x - a$. Por tanto, f es diferenciable en a si, y sólo si, existe $T \in L(X,Y)$ verificando (5).

Está claro que, salvo la hipótesis $a \in A^\circ$, más restrictiva que $a \in A \cap A'$, hemos generalizado la noción conocida para el caso $X = Y = \mathbb{R}$. Poco a poco, iremos analizando si los resultados que conocemos en dicho caso particular son o no válidos a plena generalidad. De momento, haremos cinco observaciones clave sobre la definición de diferenciabilidad en un punto.

6.2.1. Unicidad de la diferencial

- *Si f es diferenciable en a, la aplicación $T \in L(X,Y)$ que aparece en (4) es única.*

Por ser $a \in A^\circ$, existe $\delta > 0$ tal que $B(a,\delta) \subset A$. Para $v \in X \setminus \{0\}$, usamos en (4) el cambio de variable $x = a + tv \in A$ con $t \in \mathbb{R}$ y $|t| < \delta/\|v\|$, teniendo en cuenta que $x \to a$ cuando $t \to 0$ y que $x \neq a$ para $t \neq 0$. Obtenemos que

$h = tv$

$$0 = \lim_{t\to 0} \frac{\|f(a+tv) - f(a) - t\,T(v)\|}{|t|\;\|v\|} = \frac{1}{\|v\|} \lim_{t\to 0} \left\| \frac{f(a+tv) - f(a)}{t} - T(v) \right\|$$

donde hemos usado la linealidad de T. Así pues,

$$\boxed{T(v) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t}} \qquad (6)$$

Como esto es válido para todo $v \in X \setminus \{0\}$, queda claro que T es única: si dos aplicaciones lineales $T_1, T_2 : X \to Y$ verifican (4), también deberán verificar (6), de donde $T_1 = T_2$. ■

Nótese que, para comprobar la unicidad de T se ha usado que T es lineal, pero no que sea continua; la razón para exigir su continuidad se verá enseguida. También conviene resaltar que en el razonamiento anterior se usa que $a \in A^\circ$. Suponiendo sólo $a \in A \cap A'$, los límites que aparecen en (4) tendrían sentido, pero en general no podríamos asegurar la unicidad de la diferencial. De hecho, la condición $a \in A^\circ$ no es imprescindible para que T sea única, pero sí es la hipótesis más cómoda para asegurarse dicha unicidad.

Si f es diferenciable en a, a la única $T \in L(X,Y)$ que verifica (4), la llamamos **diferencial de f en a**, y se denota por $Df(a)$. Por tanto, $Df(a) \in L(X,Y)$ se caracteriza por

$$\boxed{\lim_{x \to a} \frac{f(x) - f(a) - Df(a)(x-a)}{\|x-a\|} = 0} \tag{7}$$

y como consecuencia verifica también que

$$\boxed{Df(a)(v) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} \qquad \forall v \in X} \tag{8}$$

Conviene dejar claro que, en general, de (8) no se deduce (7).

6.2.2. Relación con la continuidad

- *Si f es diferenciable en a, entonces f es continua en a.*

Para todo $x \in A \setminus \{a\}$ podemos claramente escribir

$$\boxed{f(x) = f(a) + Df(a)(x-a) + \|x-a\| \, \frac{f(x) - f(a) - Df(a)(x-a)}{\|x-a\|}}$$

La igualdad (7) nos dice sobradamente que el último sumando tiene límite cero en el punto a. Por otra parte, como $Df(a)$ es continua, también tenemos $\lim_{x \to a} Df(a)(x-a) = 0$, y concluimos que $\lim_{x \to a} f(x) = f(a)$. ■

Resaltamos que la continuidad de $Df(a)$ ha sido esencial para probar que f es continua. Esto explica que al definir la diferenciabilidad hayamos exigido $T \in L(X,Y)$. En los casos que nos interesan, X tendrá dimensión finita, luego toda aplicación lineal de X en Y será continua.

6.2.3. Significado analítico

Podemos interpretar la diferenciabilidad exactamente igual que para funciones reales de variable real: que f sea diferenciable en a significa que, “cerca” del punto a, f admite una “buena aproximación” mediante una función sencilla: la función $g : X \to Y$ dada por

$$\boxed{g(x) = f(a) + Df(a)(x-a) = f(a) - Df(a)(a) + Df(a)(x) \qquad \forall x \in X}$$

Nótese que g es la composición de $Df(a) \in L(X,Y)$ con una traslación, luego es **afín y continua**. Basta pensar en el caso particular $X = Y = \mathbb{R}$, para darse cuenta de que las funciones afines y continuas de X en Y son la generalización natural de los polinomios de primer orden en $\mathbb{R}$. Tras las constantes, y las lineales continuas, las afines continuas son las funciones de X en Y más sencillas que podemos imaginar. Pues bien, en virtud de (7), tenemos:

$$\boxed{\lim_{x \to a} \frac{f(x) - g(x)}{\|x - a\|} = 0}$$

Esto significa que, "cerca" del punto a, la función g es una "buena aproximación" de f, pues la diferencia $f(x) - g(x)$ tiende a cero en el punto a "más rápidamente" que $\|x - a\|$. Cabe esperar que, cerca del punto a, la función f se comporte de forma similar a como lo haga g. En el fondo, todos los resultados del cálculo diferencial consistirán en aprovechar esta idea.

Alternativamente, pensemos que f describe la forma en que la variable $y \in Y$ "depende" de otra variable "independiente" $x \in A$, mediante la "relación funcional" $y = f(x)$, con lo que cada variación $\Delta x = x - a$, da lugar a una variación $\Delta y = f(x) - f(a)$. Cuando f es diferenciable en el punto a, la igualdad (7) significa que, si Δx es "pequeña", Δy es "aproximadamente" igual a $Df(a)\Delta x$. Esta interpretación explica el nombre de la diferencial: controla la relación entre diferencias de las variables, permitiendo pensar que Δy depende de Δx de manera lineal y continua. Para variables con valores vectoriales, este es el tipo de dependencia más sencillo que podemos imaginar. En casos particulares que iremos considerando, este *significado analítico* de la diferencial nos llevará a sus interpretaciones geométrica y física.

6.2.4. Carácter local

Está claro que la diferenciabilidad de una función en un punto es una propiedad local: se expresa mediante un límite, que obviamente sólo involucra los valores de la función en un entorno de dicho punto. Además, como sólo trabajamos en puntos interiores del conjunto en el que nuestra función está definida, este carácter local se expresa de forma muy sencilla:

- *Si $U \subset A$ y $a \in U^\circ$, entonces f es diferenciable en a si, y sólo si, $f\big|_U$ es diferenciable en a, en cuyo caso se tiene $Df(a) = D\big(f\big|_U\big)(a)$.*

En el resultado anterior siempre podemos tomar $U = A^\circ$, luego f es diferenciable en a si, y sólo si, lo es $f\big|_{A^\circ}$, que es una función definida en el conjunto abierto A°. Puesto que sólo estudiamos la diferenciabilidad de f en puntos de A° y, para dichos puntos, podemos trabajar con $f\big|_{A^\circ}$ en lugar de f, está claro que no perdemos generalidad trabajando solamente con funciones definidas en conjuntos abiertos, como haremos casi siempre a partir de ahora.

6.2.5. Independencia de las normas

- *La diferenciabilidad de f en a, así como su diferencial $Df(a)$, se conservan al sustituir las normas de X e Y por otras equivalentes a ellas.*

Denotando por $\|\cdot\|_1$ a las normas de partida en X e Y, y por $\|\cdot\|_2$ otras equivalentes a ellas, existen constantes $\lambda, \rho \in \mathbb{R}^+$ tales que $\lambda \|x\|_1 \leqslant \|x\|_2$ para todo $x \in X$ y $\|y\|_2 \leqslant \rho \|y\|_1$ para todo $y \in Y$. Si f es diferenciable en a con las normas $\|\cdot\|_1$, para $x \in A \setminus \{a\}$ tenemos

$$\boxed{0 \leqslant \frac{\|f(x)-f(a)-Df(a)(x-a)\|_2}{\|x-a\|_2} \leqslant \frac{\rho}{\lambda} \frac{\|f(x)-f(a)-Df(a)(x-a)\|_1}{\|x-a\|_1}}$$

y deducimos que $Df(a)$ sigue siendo la diferencial de f en a para las normas $\|\cdot\|_2$. ■

En el caso que nos interesa, los espacios normados X e Y tendrán dimensión finita, luego para estudiar la diferenciabilidad de una función f podemos considerar cualquier norma en dichos espacios.

El tipo de invariancia recién obtenido se añade a otro, que es completamente evidente: la diferenciabilidad de una función no depende de las bases que podamos fijar en los espacios vectoriales X e Y, simplemente porque la hemos definido sin usar base alguna. Aunque sólo vamos a trabajar con la diferenciabilidad de funciones definidas en un abierto de $\mathbb{R}^N$ y con valores en $\mathbb{R}^M$, al haber definido la diferenciabilidad en abstracto, para espacios normados cualesquiera, tenemos a priori una noción invariante por cambios de base.

6.2.6. Diferenciabilidad global

Antes de presentar los primeros ejemplos de funciones diferenciables, pensemos en el tipo de función que nos va a interesar, que no será diferenciable en un solo punto, sino en todos los puntos del conjunto abierto en el que está definida.

Sean pues X e Y espacios normados, Ω un subconjunto abierto de X y $f : \Omega \to Y$ una función. Cuando f sea diferenciable en todo punto $x \in \Omega$, diremos simplemente que f es **diferenciable**, y $D(\Omega, Y)$ será el conjunto de todas las funciones diferenciables de Ω en Y, contenido en el conjunto $\mathcal{C}(\Omega, Y)$ de todas las funciones continuas de Ω en Y.

Para $f \in D(\Omega, Y)$ podemos considerar la función $Df : \Omega \to L(X,Y)$ que a cada punto $x \in \Omega$ hace corresponder la diferencial $Df(x) \in L(X,Y)$, y decimos que Df es la función diferencial de f, o simplemente la **diferencial** de f. Esta nomenclatura no debe causar confusión: por el contexto se sabe siempre si al hablar de la diferencial de f, nos referimos a la función Df o a la diferencial de f en un punto concreto $x \in \Omega$, que también es una función: $Df(x) \in L(X,Y)$. Ocurre aquí lo mismo que para funciones reales de variable real, no es lo mismo la derivada de una función en un punto que la función derivada.

Como $L(X,Y)$ es a su vez un espacio normado, tiene sentido plantearse la continuidad de la función Df. Decimos que $f \in D(\Omega, Y)$ es una **función de clase C^1** cuando Df es continua, y denotamos por $C^1(\Omega, Y)$ al conjunto de todas las funciones de clase C^1 de Ω en Y. También tendría sentido plantearse la diferenciabilidad de la función Df, pero esto es un asunto más delicado, que abordaremos en su momento, para el estudio de las diferenciales sucesivas de una función. Por ahora resaltamos la relación entre los conjuntos de funciones recién definidos:

$$C^1(\Omega, Y) \subset D(\Omega, Y) \subset \mathcal{C}(\Omega, Y)$$

6.3. Primeros ejemplos

Comentemos los ejemplos más obvios de funciones diferenciables entre espacios normados arbitrarios X e Y. Si $f : X \to Y$ es constante, cualquiera que sea $a \in X$, tenemos $f(x) = f(a)$ para todo $x \in X$, luego se cumple obviamente (4) sin más que tomar $T = 0$. Por tanto:

- *Si $f : X \to Y$ es constante, entonces f es diferenciable, con $Df(a) = 0$ para todo $a \in X$. Por tanto $f \in C^1(X,Y)$.*

Si $f : X \to Y$ es lineal y continua, fijado $a \in X$, tenemos entonces $f(x) - f(a) = f(x-a)$ para todo $x \in X$, luego se cumple (4) con $T = f$. Por tanto:

- *Si $f \in L(X,Y)$, entonces f es diferenciable con $Df(a) = f$ para todo $a \in X$. Vemos por tanto que Df es constante, luego $f \in C^1(X,Y)$.*

Está claro que la noción de diferenciabilidad no está pensada para trabajar con ejemplos tan obvios como los anteriores. Tan pronto como descendamos a casos particulares, es decir, concretemos los espacios normados X e Y, tendremos ejemplos menos obvios de funciones diferenciables.

Para funciones con valores en un espacio normado, las operaciones disponibles son la suma y el producto por escalares. Veamos pues una primera regla para el cálculo de diferenciales, que muestra la diferenciación como una operación *lineal*.

- *Si Ω es un abierto no vacío de X, $f,g : \Omega \to Y$ son diferenciables en $a \in \Omega$ y $\alpha, \beta \in \mathbb{R}$, entonces $\alpha f + \beta g$ es diferenciable en a con:*

$$D\big(\alpha f + \beta g\big)(a) = \alpha Df(a) + \beta Dg(a)$$

Si $f,g \in D(\Omega,Y)$, tendremos $\alpha f + \beta g \in D(\Omega,Y)$ con $D(\alpha f + \beta g) = \alpha Df + \beta Dg$. Por tanto, si $f,g \in C^1(\Omega,Y)$, se tiene $\alpha f + \beta g \in C^1(\Omega,Y)$. Así pues, $D(\Omega,Y)$ y $C^1(\Omega,Y)$ son subespacios vectoriales de $\mathcal{C}(\Omega,Y)$.

Basta pensar que, para todo $x \in \Omega \setminus \{a\}$, se tiene claramente

$$\frac{(\alpha f + \beta g)(x) - (\alpha f + \beta g)(a) - \big(\alpha Df(a) + \beta Dg(a)\big)(x-a)}{\|x-a\|}$$
$$= \alpha\, \frac{f(x) - f(a) - Df(a)(x-a)}{\|x-a\|} + \beta\, \frac{g(x) - g(a) - Dg(a)(x-a)}{\|x-a\|}$$

Por ser f y g diferenciables en a, los dos sumandos del segundo miembro tienen límite 0 en dicho punto, luego lo mismo le ocurre a la función que aparece en el primer miembro. Puesto que $\alpha Df(a) + \beta Dg(a) \in L(X,Y)$, tenemos el resultado deseado. ■

6.4. Regla de la cadena

Comprobamos ahora que la composición de aplicaciones preserva la diferenciabilidad. Esta es, sin duda, la regla más útil para el cálculo de diferenciales.

Teorema. *Sean X,Y,Z tres espacios normados, Ω y U abiertos no vacíos de X e Y respectivamente, y sean $f:\Omega\to U$ y $g:U\to Z$ dos funciones. Si f es diferenciable en un punto $a\in\Omega$ y g es diferenciable en $b=f(a)$, entonces $g\circ f$ es diferenciable en a, con*

$$D\big(g\circ f\big)(a) = Dg(b)\circ Df(a) = Dg\big(f(a)\big)\circ Df(a)$$

Por tanto, si $f\in D(\Omega,Y)$ y $g\in D(U,Z)$, entonces $g\circ f\in D(\Omega,Z)$.

Demostración. Para mayor claridad, la dividimos en tres etapas.

(a). Primero traducimos la diferenciabilidad de f y g, con una notación que haga más fáciles los cálculos. Por una parte, definimos una función $\Phi:\Omega\to\mathbb{R}_0^+$ de la siguiente forma:

$$\Phi(x) = \frac{\|f(x)-f(a)-Df(a)(x-a)\|}{\|x-a\|}\quad \forall x\in\Omega\setminus\{a\}\quad \text{y}\quad \Phi(a)=0$$

Por ser f diferenciable en a, tenemos que Φ es continua en a, y verifica:

$$\|f(x)-f(a)-Df(a)(x-a)\| = \Phi(x)\,\|x-a\| \qquad \forall x\in\Omega \tag{9}$$

Análogamente, definimos $\Psi:U\to\mathbb{R}_0^+$ por

$$\Psi(y) = \frac{\|g(y)-g(b)-Dg(b)(y-b)\|}{\|y-b\|}\quad \forall y\in U\setminus\{b\}\quad \text{y}\quad \Psi(b)=0$$

Entonces Ψ es continua en b y verifica:

$$\|g(y)-g(b)-Dg(b)(y-b)\| = \Psi(y)\,\|y-b\| \qquad \forall y\in U \tag{10}$$

Por último, sólo para abreviar la notación, definimos también $\Lambda:\Omega\to\mathbb{R}_0^+$ por

$$\Lambda(x)=\Big\|\big(g\circ f\big)(x)-\big(g\circ f\big)(a)-\big(Dg(b)\circ Df(a)\big)(x-a)\Big\| \qquad \forall x\in\Omega$$

Puesto que $Dg(b)\circ Df(a)\in L(X,Z)$, bastará comprobar que $\lim\limits_{x\to a}\dfrac{\Lambda(x)}{\|x-a\|}=0$.

(b). En una segunda fase, obtenemos una desigualdad clave, que relaciona Λ con Φ y Ψ.

Para $x\in\Omega$, escribimos $y=f(x)\in U$, y usamos en Z la desigualdad triangular:

$$0\leqslant\Lambda(x)\leqslant\Big\|g(y)-g(b)-Dg(b)(y-b)\Big\|+\Big\|Dg(b)\big[y-b-Df(a)(x-a)\big]\Big\| \tag{11}$$

Trabajamos ahora por separado con los dos sumandos que han aparecido.

El último es el más sencillo, usamos la definición de la norma en $L(Y,Z)$ junto con (9):

$$\begin{aligned}\Big\|Dg(b)\big[y-b-Df(a)(x-a)\big]\Big\| &\leqslant \|Dg(b)\|\,\|y-b-Df(a)(x-a)\|\\ &= \|Dg(b)\|\,\|f(x)-f(a)-Df(a)(x-a)\|\\ &= \|Dg(b)\|\,\Phi(x)\,\|x-a\|\end{aligned} \tag{12}$$

En vista de (10), el otro sumando es:

$$\| g(y) - g(b) - Dg(b)(y-b) \| = \Psi(y) \, \| y - b \| = \Psi\big(f(x)\big) \, \| f(x) - f(a) \|$$

La desigualdad triangular en Y, junto con (9) y la definición de la norma en $L(X,Y)$, nos dan

$$\begin{aligned} \| f(x) - f(a) \| &\leqslant \| f(x) - f(a) - Df(a)(x-a) \| + \| Df(a)(x-a) \| \\ &\leqslant \big[\Phi(x) + \| Df(a) \| \big] \, \| x - a \| \end{aligned}$$

de donde deducimos que

$$\| g(y) - g(b) - Dg(b)(y-b) \| \leqslant \Psi\big(f(x)\big) \big[\Phi(x) + \| Df(a) \| \big] \| x - a \| \tag{13}$$

Usando (12) y (13), deducimos de (11) la desigualdad que buscábamos:

$$0 \leqslant \Lambda(x) \leqslant \Psi\big(f(x)\big) \big[\Phi(x) + \| Df(a) \| \big] \| x - a \| + \| Dg(b) \| \, \Phi(x) \, \| x - a \| \tag{14}$$

(c). Usando la desigualdad (14), junto con las propiedades conocidas de Φ y Ψ, concluimos fácilmente la demostración. Para $x \in \Omega \setminus \{a\}$, de (14) deducimos que

$$0 \leqslant \frac{\Lambda(x)}{\| x - a \|} \leqslant \Psi\big(f(x)\big) \big[\Phi(x) + \| Df(a) \| \big] + \| Dg(b) \| \, \Phi(x) \tag{15}$$

luego bastará probar que el último miembro de esta desigualdad tiene límite 0 en el punto a.

Como f es diferenciable, luego continua, en el punto a, y Ψ es continua en $b = f(a)$, vemos que $\Psi \circ f$ es continua en a. Además, Φ es continua en a, luego tenemos

$$\lim_{x \to a} \Psi\big(f(x)\big) = \Psi\big(f(a)\big) = \Psi(b) = 0 \qquad \text{y} \qquad \lim_{x \to a} \Phi(x) = \Phi(a) = 0$$

Así pues, concluimos claramente que

$$\lim_{x \to a} \Big[\Psi\big(f(x)\big) \big[\Phi(x) + \| Df(a) \| \big] + \| Dg(b) \| \, \Phi(x) \Big] = 0$$

y (15) nos dice que $\displaystyle \lim_{x \to a} \frac{\Lambda(x)}{\| x - a \|} = 0$, como queríamos demostrar. ■

Para poder estudiar la composición de dos funciones de clase C^1 necesitamos una sencilla observación sobre la norma de una composición de aplicaciones lineales y continuas:

- *Si X, Y, Z son espacios normados, $T \in L(X,Y)$ y $S \in L(Y,Z)$, entonces:*

$$\| S \circ T \| \leqslant \| S \| \, \| T \| \tag{16}$$

La comprobación de esta desigualdad es inmediata. Para $x \in X$ tenemos:

$$\| (S \circ T) x \| = \| S(Tx) \| \leqslant \| S \| \, \| Tx \| \leqslant \| S \| \, \| T \| \, \| x \|$$

donde hemos usado la definición de la norma en $L(Y,Z)$ y en $L(X,Y)$. Pero usándola también en $L(X,Z)$, la desigualdad obtenida nos dice que $\| S \circ T \| \leqslant \| S \| \, \| T \|$, como se quería. ■

Podemos ya completar la regla de la cadena en la forma que sigue.

- *Sean X,Y,Z espacios normados y Ω, U subconjuntos abiertos de X,Y respectivamente. Si $f \in C^1(\Omega, Y)$ verifica que $f(\Omega) \subset U$ y $g \in C^1(U,Z)$, entonces $g \circ f \in C^1(\Omega, Z)$.*

Para $x, a \in \Omega$, escribiendo $y = f(x) \in U$ y $b = f(a) \in U$ tenemos claramente

$$\begin{aligned} D(g \circ f)(x) - D(g \circ f)(a) &= Dg(y) \circ Df(x) - Dg(b) \circ Df(a) \\ &= Dg(y) \circ \big(Df(x) - Df(a)\big) + \big(Dg(y) - Dg(b)\big) \circ Df(a) \end{aligned}$$

La desigualdad triangular, junto con (16), nos dan:

$$\begin{aligned} \| D(g \circ f)(x) - D(g \circ f)(a) \| &\leqslant \| Dg(y) \| \, \| Df(x) - Df(a) \| + \| Dg(y) - Dg(b) \| \, \| Df(a) \| \\ &= \| Dg\big(f(x)\big) \| \, \| Df(x) - Df(a) \| + \| Dg\big(f(x)\big) - Dg\big(f(a) \| \, \| Df(a) \| \end{aligned}$$

Por ser f continua en a y Dg continua en $f(a)$, la función $Dg \circ f$ es continua en a. Por otra parte, Df también es continua en a, luego tenemos

$$\lim_{x \to a} Dg\big(f(x)\big) = Dg\big(f(a)\big) \quad \text{y} \quad \lim_{x \to a} Df(x) = Df(a)$$

Usando ahora que, en cualquier espacio normado, la norma es una función continua, de la desigualdad anterior deducimos claramente que

$$\lim_{x \to a} \big\| D(g \circ f)(x) - D(g \circ f)(a) \big\| = 0$$

luego $D(g \circ f)$ es continua en a, y esto es válido para todo $a \in \Omega$, como queríamos. ■

6.5. Funciones con valores en un producto

Como ocurrió con la continuidad, vamos a ver que la diferenciabilidad de una función que tome valores en un producto de espacios normados equivale a la de sus componentes. Esto permitirá reducir el estudio de la diferenciabilidad de campos vectoriales al caso de campos escalares. En lo que sigue $Y_1, Y_2, \ldots, Y_M$ son espacios normados y consideramos el espacio normado producto $Y = \prod_{j=1}^{M} Y_j$. Para cada $j \in \Delta_M$ recordemos la *j-ésima proyección coordenada* en Y, que es la aplicación sobreyectiva $\pi_j : Y \to Y_j$ dada por $\pi_j(y) = y(j)$ para todo $y \in Y$. Es claro que $\pi_j \in L(Y, Y_j)$ con $\|\pi_j\| = 1$, luego $\pi_j \in C^1(Y, Y_j)$ con $D\pi_j(y) = \pi_j$ para todo $y \in Y$.

En sentido contrario, usaremos la *inyección natural* de Y_j en Y, es decir, la aplicación inyectiva $I_j : Y_j \to Y$, definida como sigue: para cada $u \in Y_j$, $I_j(u)$ es el vector de Y cuyas componentes son todas nulas, salvo la j-ésima, que es u:

$$[I_j(u)](j) = u \quad \text{y} \quad [I_j(u)](i) = 0 \quad \forall i \in \Delta_M \setminus \{j\}$$

De nuevo es evidente que $I_j \in L(Y_j, Y)$ con $\|I_j\| = 1$, luego $I_j \in C^1(Y_j, Y)$ con $DI_j(u) = I_j$ para todo $u \in Y_j$. Nótese también que $\pi_j \circ I_j$ es la aplicación identidad en Y_j.

Si ahora Ω es un abierto no vacío de un espacio normado X, dada una función $f : \Omega \to Y$, escribimos $f = (f_1, f_2, \ldots, f_M)$ para indicar sus componentes. Esto significa que, para $x \in \Omega$ y $j \in \Delta_M$, la j-ésima componente del vector $f(x) \in Y$ es $f_j(x)$. Usando las proyecciones e inyecciones antes definidas, la relación de ida y vuelta entre f y sus componentes se expresa con gran comodidad:

$x = (x_1, \ldots, x_m)$ $\quad f(x) = (f_1(x_1), \ldots, f_m(x_m))$

$$f_j = \pi_j \circ f \quad \forall\, j \in \Delta_M \qquad \text{y} \qquad f = \sum_{j=1}^{M} I_j \circ f_j \tag{17}$$

$f_j(x) = \pi_j \circ f(x) = \pi_j(f_1(x_1), \ldots, f_m(x_m)) = (0, \ldots, \underset{j}{f_j(x_j)}, \ldots, 0)$

Fijado $j \in \Delta_M$, la regla de la cadena y la primera igualdad de (17), nos dicen que, si f es diferenciable en $a \in \Omega$, entonces f_j también lo es, con $Df_j(a) = \pi_j \circ Df(a)$.

Recíprocamente, supongamos que f_j es diferenciable en a para todo $j \in \Delta_M$. Usando la regla de la cadena, la de diferenciación de una suma y la segunda igualdad de (17), deducimos que f es diferenciable en a con $Df(a) = \sum_{j=1}^{M} I_j \circ Df_j(a)$. En resumen, tal como habíamos anunciado, la diferenciabilidad de f en cada punto equivale a la de todas sus componentes, y tenemos la relación explícita entre las correspondientes diferenciales:

- *Sea Ω un abierto no vacío de un espacio normado X y sea $f = (f_1, f_2, \ldots f_M) : \Omega \to Y$. Entonces f es diferenciable en un punto $a \in \Omega$ si, y sólo si, f_j es diferenciable en a para todo $j \in \Delta_M$, en cuyo caso se tiene:*

$$Df_j(a) = \pi_j \circ Df(a) \quad \forall\, j \in \Delta_M \qquad y \qquad Df(a) = \sum_{j=1}^{M} I_j \circ Df_j(a) \tag{18}$$

Por tanto, $f \in D(\Omega, Y)$ si, y sólo si, $f_j \in D(\Omega, Y_j)$ para todo $j \in \Delta_M$. Análogamente se tiene $f \in C^1(\Omega, Y)$ si, y sólo si, $f_j \in C^1(\Omega, Y_j)$ para todo $j \in \Delta_M$.

Sólo queda comprobar la última afirmación. Para $a, x \in \Omega$ y $j \in \Delta_M$ tenemos:

$$\| Df_j(x) - Df_j(a) \| = \| \pi_j \circ \big(Df(x) - Df(a) \big) \| \leqslant \| Df(x) - Df(a) \|$$

Por tanto Df_j es continua siempre que Df lo sea. Recíprocamente, tenemos

$Df(x) - Df(a) = \sum_{j=1}^{M} I_j \circ Df_j(x) - \sum_{j=1}^{M} I_j \circ Df_j(a) = \sum_{j=1}^{M} (I_j \circ Df_j(x) - I_j \circ Df_j(a))$

$$\| Df(x) - Df(a) \| \leqslant \sum_{j=1}^{M} \big\| I_j \circ \big(Df_j(x) - Df_j(a) \big) \big\| \leqslant \sum_{j=1}^{M} \| Df_j(x) - Df_j(a) \| \quad \forall\, x, a \in \Omega$$

y deducimos que Df será continua siempre que lo sea Df_j para todo $j \in \Delta_M$. ■

Resaltamos el significado de las igualdades (18), análogas a las de (17), con $Df(a)$ en lugar de f: para cada $j \in \Delta_M$, la j-ésima componente de la diferencial $Df(a)$ coincide la diferencial $Df_j(a)$ de la j-ésima componente de f. Equivalentemente, podemos escribir:

$$Df(a) = \big(Df_1(a), Df_2(a), \ldots, Df_M(a) \big)$$

6.6. Productos y cocientes de funciones diferenciables

Para dos nuevas reglas de diferenciación manejamos funciones con valores en $\mathbb{R}$. Dado un abierto no vacío Ω de un espacio normado X, como ya hicimos con $\mathcal{C}(\Omega)$ y $\mathcal{C}(\Omega,\mathbb{R})$, abreviamos escribiendo $D(\Omega)$ y $C^1(\Omega)$ en lugar de $D(\Omega,\mathbb{R})$ y $C^1(\Omega,\mathbb{R})$, respectivamente. Pretendemos probar que $D(\Omega)$ y $C^1(\Omega)$ son subanillos de $\mathcal{C}(\Omega)$. Para ello nos basamos en el siguiente resultado:

- *La función $P:\mathbb{R}^2\to\mathbb{R}$, definida por $P(x,y)=xy$ para todo $(x,y)\in\mathbb{R}^2$, es de clase C^1 en $\mathbb{R}^2$ y para todo $(x,y)\in\mathbb{R}^2$ se tiene:*

$$DP(x,y)(h,k)=yh+xk \qquad \forall\,(h,k)\in\mathbb{R}^2$$

Fijado $(x,y)\in\mathbb{R}^2$, definimos $T\in L(\mathbb{R}^2,\mathbb{R})$ por $T(h,k)=yh+xk$ para todo $(h,k)\in\mathbb{R}^2$. Usando en $\mathbb{R}^2$ la norma euclídea, para todo $(h,k)\in\mathbb{R}^2$ tenemos claramente

$$\big|P(x+h,y+k)-P(x,y)-T(h,k)\big|=\big|(x+h)(y+k)-xy-yh-xk\big|=\big|hk\big|\leqslant\|(h,k)\|^2/2$$

y deducimos que

$$\lim_{(h,k)\to(0,0)}\frac{|P(x+h,y+k)-P(x,y)-T(h,k)|}{\|(h,k)\|}=0$$

es decir, P es diferenciable en el punto (x,y) con $DP(x,y)=T$, como se quería.

Para ver que DP es continua, dados $(x,y),(a,b),(h,k)\in\mathbb{R}^2$, se tiene

$yh+xk-(bh+ak)=(y-b)h+(x-a)k$

$$\Big|\big(DP(x,y)-DP(a,b)\big)(h,k)\Big|=|(y-b)h+(x-a)k|\leqslant\|(y-b,x-a)\|\,\|(h,k)\|$$

donde hemos usado la desigualdad de Cauchy-Schwartz. Por definición de la norma en $L(\mathbb{R}^2,\mathbb{R})$ deducimos que

$$\|DP(x,y)-DP(a,b)\|\leqslant\|(y-b,x-a)\|=\|(x,y)-(a,b)\| \qquad \forall\,(x,y),(a,b)\in\mathbb{R}^2$$

Esto prueba que DP no sólo es continua, sino que es una aplicación no expansiva. ■

Podemos ya probar la regla para la diferenciación de un producto de funciones:

- *Sea Ω un abierto de un espacio normado X y $f,g:\Omega\to\mathbb{R}$ funciones diferenciables en un punto $a\in\Omega$. Entonces fg es diferenciable en a con*

$$D(fg)(a)=g(a)Df(a)+f(a)Dg(a) \tag{19}$$

Si $f,g\in D(\Omega)$ se tendrá $fg\in D(\Omega)$ con $D(fg)=gDf+fDg$, luego $fg\in C^1(\Omega)$ siempre que $f,g\in C^1(\Omega)$. Así pues, $D(\Omega)$ y $C^1(\Omega)$ son subanillos de $\mathcal{C}(\Omega)$.

→ Por el recuadro azul de la página anterior.

La función $h=(f,g):\Omega\to\mathbb{R}^2$ es diferenciable en a por serlo sus componentes f y g, pero es claro que $fg=P\circ h$ donde $P:\mathbb{R}^2\to\mathbb{R}$ es la función estudiada en el resultado anterior.

Puesto que P es diferenciable en todo punto de $\mathbb{R}^2$, la regla de la cadena nos dice que $f\,g$ es diferenciable en a. Para calcular su diferencial basta observar que, para todo $x \in X$ tenemos claramente:

$$\begin{aligned} D(f\,g)(a)(x) &= \big[DP\big(h(a)\big) \circ Dh(a)\big](x) = DP\big(f(a), g(a)\big)\big(Df(a)(x)\,, Dg(a)(x)\big) \\ &= g(a)\,Df(a)(x) + f(a)\,Dg(a)(x) \end{aligned} \qquad \blacksquare$$

Por definición de DP

Nótese que la igualdad (19) es la generalización natural de la bien conocida regla para la derivada de un producto de dos funciones reales de variable real.

Podemos ya dar nuevos ejemplos de funciones de clase C^1 en un abierto no vacío $\Omega \subset \mathbb{R}^N$. Recordemos que el conjunto $\mathcal{P}(\Omega)$ de las funciones polinómicas en Ω es el subanillo de $\mathcal{C}(\Omega)$ engendrado por las funciones constantes, junto con las restricciones a Ω de las proyecciones coordenadas. Como todas estas funciones son de clase C^1 en Ω, el subanillo que engendran está contenido en $C^1(\Omega)$ que, según hemos visto, es un subanillo de $\mathcal{C}(\Omega)$. Así pues:

- *Si Ω es un abierto no vacío de $\mathbb{R}^N$, se tiene $\mathcal{P}(\Omega) \subset C^1(\Omega)$, es decir, toda función polinómica en Ω es de clase C^1.*

Pasamos a estudiar ahora la diferenciabilidad de un cociente de funciones diferenciables:

- *Sea Ω un abierto no vacío de un espacio normado X y sean $f, g : \Omega \to \mathbb{R}$ diferenciables en un punto $a \in \Omega$, con $g(x) \neq 0$ para todo $x \in \Omega$. Entonces la función cociente f/g es diferenciable en a con*

$$D(f/g)(a) = \frac{1}{g(a)^2}\big(g(a)Df(a) - f(a)\,Dg(a)\big)$$

Por tanto, si $f, g \in D(\Omega)$ se tiene $f/g \in D(\Omega)$, y $f/g \in C^1(\Omega)$ siempre que $f, g \in C^1(\Omega)$.

Tenemos $1/g = \varphi \circ g$ donde $\varphi : \mathbb{R}^* \to \mathbb{R}$ viene dada por $\varphi(x) = 1/x$ para todo $x \in \mathbb{R}^*$, y es claro que $\varphi \in C^1(\mathbb{R}^*)$. Como g es diferenciable en a, la regla de la cadena nos dice que $1/g$ también lo es, y la regla para la diferenciabilidad de un producto nos da que $f/g = f\,(1/g)$ es diferenciable en a. Si $f, g \in C^1(\Omega)$, se tendrá también $1/g \in C^1(\Omega)$, luego $f/g \in C^1(\Omega)$.

Por último, para calcular la diferencial de f/g en a, la regla de la cadena nos da

$$D(1/g)(a)(x) = D\varphi\big(g(a)\big)\big(Dg(a)(x)\big) = \frac{-1}{g(a)^2}Dg(a)(x) \qquad \forall x \in X$$

y ahora basta usar la diferencial de un producto:

$$D(f/g)(a) = \frac{1}{g(a)}Df(a) + f(a)\,D(1/g)(a) = \frac{1}{g(a)^2}\big(g(a)Df(a) - f(a)\,Dg(a)\big) \qquad \blacksquare$$

Recordemos el conjunto $\mathcal{R}(\Omega)$ de las funciones racionales en un abierto no vacío $\Omega \subset \mathbb{R}^N$, es decir $f \in \mathcal{R}(\Omega)$ cuando $f = P/Q$ donde $P, Q \in \mathcal{P}(\Omega)$ y $Q(x) \neq 0$ para todo $x \in \Omega$. Del resultado anterior deducimos claramente:

- *Si Ω es un subconjunto abierto de $\mathbb{R}^N$, se tiene $\mathcal{R}(\Omega) \subset C^1(\Omega)$, es decir, toda función racional en Ω es de clase C^1.*

6.7. Ejercicios

1. Probar que la norma de un espacio normado $X \neq \{0\}$ nunca es diferenciable en 0.

2. Probar que la norma euclídea es diferenciable en todo punto de $\mathbb{R}^N \setminus \{0\}$ y calcular su diferencial. ¿En qué puntos de $\mathbb{R}^2$ es diferenciable la norma de la suma?

3. Probar que, tanto la diferenciabilidad de una función, como su diferencial, se conservan por traslaciones. Más concretamente, sean X,Y espacios normados, Ω un abierto no vacío de X y $f : \Omega \to Y$ diferenciable en un punto $a \in \Omega$. Fijados $x_0 \in X$ e $y_0 \in Y$, se define $\widehat{\Omega} = \{x \in X : x + x_0 \in \Omega\}$ y la función $\widehat{f} : \widehat{\Omega} \to Y$ dada por

$$\widehat{f}(x) = f(x+x_0) + y_0 \qquad \forall x \in \widehat{\Omega}$$

   Probar que $\widehat{f}$ es diferenciable en $a - x_0$ con $D\widehat{f}(a - x_0) = Df(a)$.

4. Sea Ω un abierto no vacío de un espacio normado X y $f : \Omega \to \mathbb{R}$ diferenciable en un punto $a \in \Omega$. Sea J un intervalo abierto tal que $f(\Omega) \subset J$ y $g : J \to \mathbb{R}$ una función derivable en el punto $b = f(a)$. Probar que $g \circ f$ es diferenciable en a. ¿Cual es la relación entre $D(g \circ f)(a)$ y $Df(a)$?

5. Dar un ejemplo de una función $f \in D(\mathbb{R}^N)$ tal que $f \notin C^1(\mathbb{R}^N)$.