FACULTADE DE MATEMÁTICAS

Trabajo Fin de Grado

Diferenciación en espacios de Banach

Elena Parrado Villasante

2020/2021

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

GRAO DE MATEMÁTICAS

Trabajo Fin de Grado

Diferenciación en espacios de Banach

Elena Parrado Villasante

Julio 2021

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Trabajo propuesto

Área de Conocimiento: Análisis Matemático
Título: Diferenciación en espacios de Banach
Breve descripción del contenido
En este trabajo se extenderán los conceptos estudiados en la materia de Diferenciación de funciones de varias variables reales. Así, mientras que la mencionada materia trata los conceptos de diferenciabilidad de funciones definidas en $\mathbb{R}^n$, consideraremos ahora espacios más generales (en particular, espacios de Banach). Para esto, introduciremos la derivada de Fréchet definida en tales espacios y estudiaremos la generalización de los resultados más importantes relativos a la diferenciabilidad de funciones, mostrando algunas de sus aplicaciones, como puede ser la resolución de problemas variacionales.

Índice general

Resumen

El objetivo de este trabajo será extender los conceptos relativos a la diferenciabilidad de funciones de varias variables reales a funciones definidas sobre espacios más generales. En primer lugar introduciremos el concepto de espacio de Banach y estudiaremos algunas propiedades que cumplen las funciones definidas en estos. A continuación, veremos como ampliar lo conocido en $\mathbb{R}^n$ a estos nuevos espacios, introduciendo los términos de derivadas de Fréchet y Gâteau. Por último demostraremos los Teoremas de la Función Inversa e Implícita.

Abstract

The aim of this work is to extend the concepts related to the differentiability of real multivariable functions to functions defined on more general spaces. First of all, we will introduce the concept of Banach space and we will study some properties of functions defined on there. Then, we will see how to expand what is known in $\mathbb{R}^n$ to these new spaces, introducing the Fréchet and Gâteau derivative terms. Finally, we will prove the Inverse and Implicit Function Theorems.

Introducción

Es en el siglo XVII cuando, al comenzar los estudios sobre la velocidad de los cuerpos en movimiento, nace lo que conocemos como cálculo diferencial, debido a que esta velocidad debe calcularse teniendo en cuenta la distancia que recorre el cuerpo de estudio en un tiempo infinitesimalmente pequeño. A pesar de que han pasado cuatro siglos desde que aparece por primera vez el cálculo diferencial, este se mantuvo prácticamente invariante a lo largo del tiempo.

En el Grado de Matemáticas aparece por primera vez el concepto de cálculo diferencial propiamente dicho en la asignatura de *Continuidad y Derivabilidad de Funciones de una Variable Real*, aplicado a funciones reales de una variable, para luego generalizarse a funciones de varias variables reales en la materia de *Diferenciación de Funciones de varias Variables Reales*. Lo que haremos a lo largo de todo este trabajo será ver cómo podemos extender los conceptos aprendidos en dichas materias a conjuntos de funciones definidas sobre espacios más generales.

Estos espacios a los que nos referimos y sobre los cuales queremos ver los resultados de manera general son los espacios de Banach. Se trata de espacios vectoriales normados y completos, es decir, toda sucesión de Cauchy es convergente en ellos. Cabe destacar que en los capítulos centrales, que son el 2 y el 3, no es necesario que las funciones con las que trabajamos estén definidas sobre espacios de Banach, bastará simplemente con que estén definidas sobre espacios vectoriales normados. Sin embargo, en el capítulo en el que tratamos los Teoremas de la Función Implícita e Inversa necesitaremos que nuestras funciones estén definidas en espacios de Banach. Estos dos teoremas son los grandes ejes del cálculo diferencial. Gracias al primero de ellos podemos establecer si una función tendrá inversa o no únicamente estudiando su diferencial. El segundo, por otro lado, nos ofrecerá la posibilidad de expresar una ecuación de varias variables en función de solo una de ellas. Ambos teoremas son equivalentes entre si en el sentido de que podemos demostrar uno a partir del otro.

Veremos ahora detalladamente lo que se recoge a lo largo de todo este trabajo. Empezaremos definiendo en el Capítulo 1 el concepto de espacio de Banach, para el cual necesitaremos conocer qué propiedades cumple una norma y qué entendemos cuando hablamos de espacio completo. Esto quedará clarificado con un par de ejemplos. Continuaremos con la definición de función continua y enunciaremos una serie de resultados sobre aplicaciones lineales como por ejemplo la caracterización de la continuidad para este tipo de aplicaciones. En la última sección de este capítulo introduciremos las funciones lipschitzianas y contractivas, además de las definiciones de homeomorfismos e isomorfismos y algunos resultados que necesitaremos en el desarrollo del último capítulo del trabajo.

En el Capítulo 2 veremos la analogía que existe entre la diferencial de funciones reales de variables reales y la de funciones definidas en espacios vectoriales normados arbitrarios. Hablaremos de las derivadas de Fréchet y de Gâteau, además de establecer una relación geométrica entre una función y su diferencial, introduciendo lo que se conoce como funciones tangentes. Al hablar de diferencial enunciaremos también algunas propiedades que nos ayudarán a la hora de su estudio, como la unicidad, su continuidad o la relación entre las derivadas de Fréchet y Gâteau, ya que la existencia de la primera implicará la existencia de la segunda. Continuaremos viendo qué ocurre con la diferencial de la composición, suma o producto de aplicaciones, así como la aplicación constante o lineal. Además, al definir funciones en espacios producto nacen las derivadas parciales de una función. Por otro lado, trataremos el Teorema del valor medio, para el que será necesario hablar de derivadas por la derecha e izquierda, así como de alguna caracterización relacionada con estas. Gracias a este teorema podremos establecer qué condiciones han de cumplir las derivadas parciales de una función para que esta sea diferenciable en un punto. Para terminar, como aplicación del Teorema del valor medio definiremos las funciones estrictamente tangentes y estrictamente diferenciales.

En el Capítulo 3 hablaremos en un primer momento sobre la diferencial segunda de una función y las propiedades que cumple esta y más tarde extenderemos estos conceptos a la diferencial de orden n. Una vez introducido esto gracias a la Fórmula de Taylor veremos como aproximar localmente funciones.

Para terminar, en el último capítulo enunciaremos y probaremos los Teoremas de la función Inversa e Implícita, además de hablar sobre difeomorfismos, que serán necesarios para la prueba del primero de dichos teoremas.

Capítulo 1

Espacios de Banach

A lo largo de este capítulo, denotaremos por $\mathbb{K}$ al cuerpo $\mathbb{R}$ o al cuerpo $\mathbb{C}$. Además, una vez introduzcamos el concepto de espacio de Banach trabajaremos generalmente en dos de estos espacios que denotaremos por E y F.

Nos centraremos en introducir las primeras nociones necesarias para definir un espacio de Banach, viendo un par de ejemplos que nos ayudarán a visualizarlos mejor. Además, en este capítulo veremos resultados que necesitaremos a lo largo de todo el trabajo relacionado con los espacios de Banach.

Para su desarrollo nos guiaremos por el Capítulo 1 de [1], el Capítulo 1 de [2], el Capítulo de 1 [6] y por el Capítulo 5 de [3].

Definición 1.1. Definimos una norma en el espacio vectorial E como una función real

$$\|\cdot\| : E \to \mathbb{R}$$

tal que para cualesquiera x, $y \in E$ y para todo $\lambda \in \mathbb{K}$ se cumplen las siguiente propiedades:

1. $\|x\| \geq 0$.
2. $\|x + y\| \leq \|x\| + \|y\|$ (desigualdad triangular).
3. $\|\lambda x\| = |\lambda| \cdot \|x\|$.
4. $\|x\| = 0$ si y solo si $x = 0$.

Por simplicidad, a no ser que se indique lo contrario y exceptuando el valor absoluto $|\cdot|$, suponemos que nuestros espacios están dotados con la norma $\|\cdot\|$, que no tiene por qué ser igual en todos. Esta notación será la que utilicemos únicamente a lo largo de este capítulo.

Una vez tenemos definido el concepto de norma, recogemos la siguiente definición.

Definición 1.2. Un espacio normado es un espacio vectorial dotado con una norma.

Observación 1. Sea E un espacio normado, entonces es también un espacio métrico con la métrica d, definida como

$$d(x, y) = \|y - x\| \text{ para todo } x, y \in E.$$

Definición 1.3. Dado E un espacio normado, definimos en él los siguientes términos.

1. Sea $B(a, r) = \{x \in E \mid d(a, x) < r\}$ la bola abierta de centro $a \in E$ y radio $r > 0$.

2. Sea $B[a, r] = \{x \in E \mid d(a, x) \leq r\}$ la bola cerrada de centro $a \in E$ y radio $r \geq 0$.

3. Sea $S(a, r) = \{x \in E \mid d(a, x) = r\}$ la esfera de centro $a \in E$ y radio $r \geq 0$.

Definiremos ahora el último concepto necesario para obtener la definición de espacio de Banach y poder enunciar los resultados relacionados con estos.

Definición 1.4. Un espacio métrico se dice completo si toda sucesión de Cauchy contenida en el espacio converge a un elemento de este.

Definición 1.5. Un espacio normado $(E, \|\cdot\|)$ es un espacio de Banach si, con la distancia inducida por la norma, es un espacio métrico completo.

Ilustraremos este tipo de espacios con unos ejemplos que nos servirán para clarificar la definición.

Ejemplo 1.6. Sea $\mathcal{C}([a, b])$ el espacio vectorial de las funciones continuas de $[a, b] \subset \mathbb{R}$ en $\mathbb{R}$. $(\mathcal{C}([a, b]), \|\cdot\|)$, donde

$$\|f\| = \sup_{x \in [a,b]} |f(x)|$$

tiene estructura de espacio de Banach.

Para probar que se trata de un espacio de Banach tendremos que ver que efectivamente $\|\cdot\|$ es una norma y que $\mathcal{C}([a, b])$ con esta norma es un espacio completo.

Empezaremos comprobando que $\|\cdot\|$ es una norma, viendo que cumple las propiedades de la Definición 1.1 para cualesquiera $f, g \in E$ y para todo $\lambda \in [a, b]$. En efecto se tiene que:

1. $\|f\| = \sup\limits_{x \in [a,b]} |f(x)| \geq 0$.

2. $\|f + g\| = \sup\limits_{x \in [a,b]} |f(x) + g(x)| \leq \sup\limits_{x \in [a,b]} |f(x)| + \sup\limits_{x \in [a,b]} |g(x)| = \|f\| + \|g\|$.

3. $\|\lambda f\| = \sup_{x\in[a,b]} \|\lambda f(x)\| = |\lambda| \cdot \sup_{x\in[a,b]} |f(x)| = |\lambda| \cdot \|f\|$.

4. $\|f\| = \sup_{x\in[a,b]} |f(x)| = 0$ si y solo si trabajamos con la función $f = 0$.

Hemos probado entonces que $\|\cdot\|$ es una norma y por lo tanto E es un espacio normado.

Veamos ahora que se trata de un espacio completo. Supongamos que $\{f_n\}$ es una sucesión de Cauchy en E, y veamos que converge a una función $f \in E$.

Hemos supuesto que $\{f_n\}$ es una sucesión de Cauchy en E, por lo que podemos deducir que $\{f_n(x)\}$ es de Cauchy en $\mathbb{R}$ para todo $x \in [a,b]$. Como además sabemos que $\mathbb{R}$ es completo, se puede definir la función $f : [a,b] \to \mathbb{R}$ que será el límite de la sucesión, $f(x) = \lim_{n\to\infty} f_n(x)$.

Por ser $\{f_n\}$ una sucesión de Cauchy, para todo $\varepsilon > 0$ existe un número natural N tal que para todos los números naturales $n, m > N$ se tiene que

$$\|f_n - f_m\| < \varepsilon.$$

Por lo tanto, para cada $m > N$ y para cada $x \in [a,b]$ se tiene que

$$|f(x) - f_m(x)| = \lim_{n\to\infty} |f_n(x) - f_m(x)| \le \lim_{n\to\infty} \|f_n - f_m\| \le \varepsilon. \tag{1.1}$$

Deducimos de (1.1) que la sucesión $\{f_n\}$ converge uniformemente a f en $[a,b]$. Sabemos que el límite uniforme de una función continua es continuo, por lo tanto es evidente que $f \in \mathcal{C}([a,b])$. Queda probado que $\{f_n\}$ converge en E, y por lo tanto E es completo.

De esta manera hemos visto que efectivamente $(\mathcal{C}([a,b]), \|\cdot\|)$ es un espacio de Banach.

Ejemplo 1.7. Si tenemos un espacio de Banach F, el espacio vectorial que representa al conjunto de funciones continuas y acotadas de $B(0,r)$ en F, $\mathcal{C}_b(B(0,r),F)$, es un espacio de Banach con la norma del ejemplo anterior.

Ejemplo 1.8. Consideremos el espacio vectorial $\mathcal{L}^1([0,1])$ de funciones definidas en el intervalo $[0,1] \subset \mathbb{R}$ que son integrables en el sentido de Lebesgue. Definimos la norma

$$\|f\| = \int_0^1 |f(t)|dt.$$

La aplicación así definida es fácil ver que cumple las 3 primeras propiedades de la Definición 1.1:

1. $\|f\| = \int_0^1 |f(t)|dt \ge 0$.

2. $\|f+g\| = \int_0^1 |f(t)+g(t)|dt \leq \int_0^1 |f(t)|dt + \int_0^1 |g(t)|dt = \|f\| + \|g\|.$

3. $\|\lambda f\| = \int_0^1 |\lambda f(t)|dt = |\lambda| \int_0^1 |f(t)|dt = |\lambda| \cdot \|f\|.$

Sin embargo no cumple la última de las condiciones, ya que $\|f\| = 0$ no implica necesariamente que f sea la función nula. Por lo tanto no podemos afirmar que se trate de un espacio de Banach, ya que $\|\cdot\|$ no es una norma.

Lo que haremos será cambiar el espacio en el que estamos trabajando. Para ello definimos en $\mathcal{L}^1([0,1])$ la relación de equivalencia siguiente

$$f_1 \sim f_2 \text{ si y solo si } f_1(x) = f_2(x) \text{ para casi todo } x \in [0,1].$$

Consideremos ahora el espacio cociente por esta relación de equivalencia

$$L^1([0,1]) = \frac{\mathcal{L}^1([0,1])}{\sim}.$$

Entonces $(L^1([0,1]), \|\cdot\|)$ sí que es un espacio de Banach, ya que $\|\cdot\|$ cumple las cuatro propiedades de norma en $L^1([0,1])$. Las tres primeras se prueban igual que en $\mathcal{L}^1([0,1])$, mientras que la cuarta se sigue de que

$$\|f\| = 0 \text{ si y solo si } f(x) = 0 \text{ para casi todo } x \in [0,1],$$

y por la relación de equivalencia que tenemos definida en $L^1([0,1])$ esto ocurre si y solo si $f = 0$ en $L^1([0,1])$

Por la teoría de Lebesgue sabemos que $L^1([0,1])$ es un espacio completo y por lo tanto concluimos que efectivamente $(L^1([0,1]), \|\cdot\|)$ es un espacio de Banach.

1.1. Aplicaciones lineales continuas

En esta sección seguimos asumiendo, por simplicidad, las hipótesis expuestas al comienzo del capítulo. Veremos algunas condiciones que nos ayudarán a establecer si una función entre espacios de Banach es continua y qué propiedades cumple la composición de este tipo de aplicaciones.

Empezaremos hablando en general, definiendo lo que entendemos por continuidad de una función.

Definición 1.9. Sean E y F dos espacios de Banach y $f : E \to F$. Se dice que f es continua en un punto $a \in E$ si para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para todo $x \in E$ con $\|x - a\| < \delta$ se cumple que $\|f(x) - f(a)\| < \varepsilon$.

Definición 1.10. Una función f diremos que es continua si lo es en todos los puntos de su dominio.

Una vez hemos introducido la definición de continuidad entre dos espacios de Banach, establecemos el siguiente teorema.

Teorema 1.11. *Sean E y F dos espacios de Banach y $f : E \to F$ una función lineal, las siguientes condiciones son equivalentes:*

1. *f es una función continua para todo $x \in E$.*
2. *f es continua en el origen.*
3. *Existe $M > 0$ tal que $\|f(x)\| \leq M$ para todo $x \in B[0, 1]$.*

Demostración. Veamos cada una de las implicaciones de manera individual.

Es claro que $1 \Rightarrow 2$ se cumple de manera trivial.

Veamos que $2 \Rightarrow 3$. Para esta demostración partimos de la suposición de que la función f es continua en 0. Por lo tanto, si tomamos una bola abierta $B(0, 1) \subset F$, su imagen inversa por f es un entorno abierto V de 0 en E, es decir, existe un $r > 0$ tal que $B(0, r) \subset V$. Se tiene entonces que

$$\|x\| \leq r \text{ implica que } \|f(x)\| \leq 1.$$

Haciendo el cambio de variable $y = rx$, como f es una función lineal, llegaríamos a que

$$\|f(y)\| = r \cdot \|f(x)\|.$$

Entonces si $x \in B[0, 1]$, $y = rx \in B[0, r]$, por lo que $\|f(y)\| \leq 1$. Consecuentemente se deduce que

$$\|x\| \leq 1 \text{ implica que } \|f(x)\| \leq \frac{1}{r}.$$

Queda entonces probado que $\|f(x)\|$ está acotado en la bola $B[0, 1]$.

Queda ver por último que $3 \Rightarrow 1$. Supongamos que existe un $M > 0$ tal que $\|f(x)\| \leq M$ para todo $x \in B[0, 1]$. De lo anterior se deduce que

$$\|f(x)\| \leq M\|x\| \text{ para todo } x \in E.$$

En efecto, si $x = 0$ es trivial y si $x \neq 0$, $\|x\| = r$, tomando $y = \frac{1}{r}x$ cumple que $\|y\| = 1$. Por lo que $\|f(y)\| \leq M$ y $\|f(x)\| = r\|f(y)\| \leq M\|x\|$. Bajo estas hipótesis queremos ver que la función f es continua en cualquier punto $a \in E$. Por ser f lineal se cumple que $f(x - a) = f(x) - f(a)$ y si acotamos $\|x - a\| \leq \frac{\varepsilon}{M}$ para todo $\varepsilon > 0$, se tiene que

$$\|f(x) - f(a)\| \leq M \cdot \frac{\varepsilon}{M} = \varepsilon.$$

Por lo tanto queda probada la continuidad de f en cualquier punto $a \in E$. $\square$

Introducimos a continuación el espacio vectorial de las aplicaciones lineales continuas de un espacio vectorial normado E en otro F, que denotaremos por $\mathcal{L}(E,F)$. En este espacio podemos definir la norma siguiente

$$\|f\| = \sup_{\|x\|\leq 1} \|f(x)\| = \sup_{\|x\|\neq 0} \frac{\|f(x)\|}{\|x\|}.$$

De la expresión previa se puede deducir la relación fundamental dada por

$$\|f(x)\| \leq \|f\|_{\mathcal{L}} \cdot \|x\|, \quad \text{para todo } x \in E.$$

Es fácil comprobar que $\|\cdot\|_{\mathcal{L}}$ es una norma sobre el espacio $\mathcal{L}(E,F)$, por lo que este es un espacio vectorial normado. Teniendo estas relaciones en cuenta, enunciaremos un lema de gran utilidad que nos relaciona la norma de la composición de dos aplicaciones lineales con la norma de cada aplicación.

Lema 1.12. *Sean T y S dos aplicaciones lineales, entonces*

$$\|T \circ S\|_{\mathcal{L}} \leq \|T\|_{\mathcal{L}} \cdot \|S\|_{\mathcal{L}}.$$

Demostración. Por hipótesis S y T son dos aplicaciones lineales, entonces $T \circ S$ es lineal. Se tiene que

$$\|T \circ S\|_{\mathcal{L}} = \sup_{x\neq 0} \frac{\|(T \circ S)(x)\|}{\|x\|} = \sup_{\|x\|=1} \|(T \circ S)(x)\|.$$

Por la definición de la norma $\|T\|_{\mathcal{L}}$, se cumple que

$$\|T\|_{\mathcal{L}} \geq \frac{\|T(x)\|}{\|x\|} \quad \text{para todo} \quad x \neq 0$$

y por lo tanto

$$\|T\|_{\mathcal{L}} \cdot \|x\| \geq \|T(x)\| \quad \text{para todo} \quad x \neq 0.$$

Teniendo en cuenta de nuevo la composición de las dos funciones

$$\|T \circ S\|_{\mathcal{L}} = \sup_{\|x\|=1} \|(T \circ S)(x)\| = \sup_{\|x\|=1} \|T(S(x))\| \leq \|T\|_{\mathcal{L}} \sup_{\|x\|=1} \|S(x)\| = \|T\|_{\mathcal{L}} \cdot \|S\|_{\mathcal{L}}.$$

□

Como hemos dicho antes, el espacio $\mathcal{L}(E,F)$ con la norma $\|\cdot\|_{\mathcal{L}}$ es un espacio vectorial normado. Veamos, con el siguiente teorema, bajo qué hipótesis podemos asegurar que se trata de un espacio de Banach.

Teorema 1.13. *Si F es un espacio de Banach entonces $\mathcal{L}(E,F)$ es un espacio de Banach.*

Demostración. Suponemos que tenemos una sucesión de Cauchy $\{f_n\}$ en el espacio $\mathcal{L}(E, F)$ y queremos ver que converge en este para poder asegurar que se trata de un espacio completo y, por lo tanto, de Banach. Para cada $r > 0$, consideramos las restricciones de $\{f_n\}$ a la bola $B[0, r]$, que serán las funciones $f_n^{(r)}$, que formarán una sucesión de Cauchy en el espacio vectorial $\mathcal{C}_b(B(0, r), F)$, y por el Ejemplo 1.7 este espacio es completo. Por lo tanto sabemos que la sucesión $\left\{f_n^{(r)}\right\}$ converge uniformemente a una función $f^{(r)}$ continua y acotada.

Si tomamos un $r' < r$ y restringimos la función $f^{(r)}$ a la bola $B[0, r']$, esta será igual a una función $f^{(r')}$. Por lo tanto la colección de funciones $f^{(r)}$ define una función f en todo el espacio E, tal que la restricción de esta a la bola $B[0, r]$ sea $f^{(r)}$. Para cada $x \in E$ y debido a la convergencia uniforme en la bola $B(0, r)$, se tiene que

$$f(x) = \lim_{n\to\infty} f_n(x).$$

Tomando dos puntos $x, y \in E$, como tenemos que las funciones f_n son lineales, llegamos a las siguientes igualdades, que prueban que la función f es lineal también.

$$f(x + y) = \lim_{n\to\infty} f_n(x + y) = \lim_{n\to\infty} f_n(x) + \lim_{n\to\infty} f_n(y) = f(x) + f(y)$$

y

$$f(\lambda x) = \lim_{n\to\infty} f_n(\lambda x) = \lim_{n\to\infty} \lambda f(x) = \lambda f(x).$$

Visto esto, podemos afirmar que para cada bola $B[0, r]$, $\|f(x)\|$ está acotada y por lo tanto, por el Teorema 1.11, f es lineal y continua.

Finalmente, $\|f - f_n\|$, por como está definida esta norma, tiende a 0, ya que

$$\|f - f_n\| = \sup_{\|x\|\leq 1} \|f(x) - f_n(x)\|$$

y sabemos que en la bola $B[0, 1]$ la sucesión $\{f_n\}$ converge uniformemente a f. Queda probado que $\mathcal{L}(E, F)$ es un espacio de Banach. $\square$

1.2. Isomorfismos en espacios de Banach

A lo largo de esta sección veremos algunas definiciones y resultados que nos serán de gran utilidad a la hora de desarrollar algunas demostraciones posteriores. Nos guiaremos, además de por las referencias citadas al comienzo del capítulo, por el Capítulo 4 de [1].

Para la compresión de los dos resultados que vamos a enunciar necesitaremos un par de definiciones.

Definición 1.14. Sean E y F espacios vectoriales normados y $U \subset E$. Diremos que $f : U \to F$ es una función Lipschitz continua o k-lipschitziana si existe una constante $k > 0$ tal que

$$\|f(x_1) - f(x_2)\| \leq k\|x_1 - x_2\| \quad \text{para todo } x_1, x_2 \in U.$$

Definición 1.15. Una función f Lipschitz continua con constante de Lipschitz $k < 1$, se dice contractiva o k-contractiva.

Enunciamos ahora el primero de los resultados, un lema que nos ayudará a la hora de demostrar el teorema posterior.

Lema 1.16. *Sean un espacio de Banach E, $a \in E$ y $r > 0$ tales que $B(a, r) \subset E$, y una aplicación continua $f : B(a, r) \to E$ tal que la aplicación $\varphi : B(a, r) \to E$ definida como*

$$\varphi(x) = x - f(x)$$

es k-contractiva. Si $f(a) = b$, entonces para todo $y \in B(b, (1 - k)r)$ existe un único x en $B(a, r)$ tal que $f(x) = y$.

Demostración. Empezaremos demostrando la unicidad. Tomamos $x, x' \in B(a, r)$ y tenemos que

$$f(x) - f(x') = (x - x') - (\varphi(x) - \varphi(x'))$$

de donde, tomando normas, llegamos a

$$\|f(x) - f(x')\| \geq \|x - x'\| - \|\varphi(x) - \varphi(x')\|.$$

Habíamos supuesto que φ es contractiva y por lo tanto llegamos a que

$$\|f(x) - f(x')\| \geq (1 - k) \cdot \|x - x'\|. \tag{1.2}$$

Entonces, como para estos dos puntos de la bola se tiene que $f(x) = f(x')$, volviendo a la expresión anterior llegamos a que $x = x'$.

Veamos ahora la existencia de este punto. Su construcción la llevaremos a cabo por el método de las aproximaciones sucesivas, tomando una sucesión de puntos

$$\begin{cases} x_0 = a \\ x_1 = y + \varphi(x_0) \\ \vdots \\ x_{n+1} = y + \varphi(x_n) \\ \vdots \end{cases} \tag{1.3}$$

Es necesario que nos aseguremos que en cada paso del proceso el iterante que estamos construyendo, x_n, pertenezca a la bola $B(a, r)$. Esto es lo que nos permite construir x_{n+1}, ya que φ está definida en $B(a, r)$. Vamos a probar por inducción en n que se cumple que

$$\|x_n - a\| \leq \frac{1 - k^n}{1 - k} \|y - b\|, \tag{1.4}$$

y como por hipótesis $\|y - b\| < (1 - k)r$, tendremos que $\|x_n - a\| < (1 - k^n)r < r$.

Para $n = 1$ se tiene que

$$x_1 - a = y + \varphi(a) - a = y - f(a) = y - b,$$

y por lo tanto (1.4) se cumple para $n = 1$.

Suponemos que (1.4) se cumple para $n > 1$ y lo vamos a probar para $n + 1$. Podemos escribir

$$x_{n+1} - x_n = y + \varphi(x_n) - y - \varphi(x_{n-1}),$$

de donde, por ser φ k-lipschitziana, se sigue que

$$\|x_{n+1} - x_n\| \leq k\|x_n - x_{n-1}\|.$$

Por recurrencia,

$$\|x_{n+1} - x_n\| \leq k^n \|x_1 - a\| = k^n \|y - b\|.$$

Esta desigualdad prueba que la serie que tiene por término general $x_{n+1} - x_n$ es convergente y por lo tanto la sucesión $\{x_n\}$ es de Cauchy. Teniéndola en cuenta y junto con (1.4) queda probado que

$$\begin{aligned} \|x_{n+1} - a\| &\leq \|x_n - a\| + \|x_{n+1} - x_n\| \\ &\leq \left(\frac{1 - k^n}{1 - k} + k^n \right) \cdot \|y - b\| = \frac{1 - k^{n+1}}{1 - k} \|y - b\|. \end{aligned}$$

Si tomamos x como el límite de $\{x_n\}$, pasándolo a la expresión de (1.4) se tiene que

$$\|x - a\| \leq \frac{1}{1 - k} \|y - b\| < r,$$

y por como viene dada la sucesión en (1.3)

$$x = y + \varphi(x),$$

de donde concluimos que $y = f(x)$. $\square$

Veremos un par de conceptos que aparecerán en los resultados que enunciaremos un poco más adelante.

Definición 1.17. Sean E y F dos espacios de Banach y V y W dos abiertos de estos, respectivamente. Se dice que $f : V \to W$ es un homeomorfismo si es una aplicación continua, biyectiva y además su inversa $g = f^{-1} : W \to V$ es continua.

Definición 1.18. Dada una aplicación $f : E \to F$, con E y F espacios normados, decimos que es un isomorfismo si:

1. f es lineal y continua.
2. Existe una aplicación $g : F \to E$ lineal y continua, tal que $g \circ f = I_E$ y $f \circ g = I_F$.

Cabe destacar que para la notación de los isomorfismos entre dos espacios, como puede ser de E en F, utilizaremos la notación $\mathrm{Isom}(E, F)$.

Podemos introducir el siguiente teorema.

Teorema 1.19. *Sean un espacio de Banach E, $a \in E$ y $r > 0$ tales que $B(a, r) \subset E$, y una aplicación continua $f : B(a, r) \to E$ tal que la aplicación $\varphi : B(a, r) \to E$ dada por*

$$\varphi(x) = x - f(x)$$

es k-contractiva. Entonces existe un abierto $V \subset B(a, r)$ que contiene a a, tal que f es un homeomorfismo de V en la bola abierta $B(b, (1-k)r)$, con $f(a) = b$. Además, la aplicación inversa

$$g = f^{-1} : B(b, (1-k)r) \to B(a, r)$$

es $\left[\frac{1}{1-k}\right]$-lipschitziana.

Demostración. Para cada $y \in B(b, (1-k)r)$, denotemos por $g(y)$ al único $x \in B(a, r)$ tal que $f(x) = y$. Definimos entonces una aplicación

$$g : B(b, (1-k)r) \to B(a, r).$$

La desigualdad que tenemos en (1.2) nos indica que, dados $y, y' \in B(b, (1-k)r)$ se tiene que

$$\|g(y) - g(y')\| \leq \frac{1}{1-k}\|y - y'\|.$$

De esta manera podemos afirmar que la función g es $\left[\frac{1}{1-k}\right]$-lipschitziana y, en particular, continua. Sea ahora $V \subset B(a, r)$ la imagen de g, por lo tanto

$$V = f^{-1}(B(b, (1-k)r)),$$

y como la imagen recíproca de un abierto por una función continua es abierto, queda probado que V es un abierto en la bola $B(a, r)$, que a su vez lo es en E.

Es claro que las aplicaciones f y g son biyectivas, continuas y una es la inversa de la otra y por lo tanto son homeomorfismos. □

Para terminar este capítulo veremos el siguiente teorema.

Teorema 1.20 (de Banach). *Si E y F son dos espacios de Banach, toda aplicación lineal, continua y biyectiva $f : E \to F$ es un isomorfismo.*

Capítulo 2

La diferencial en espacios normados

A lo largo de este capítulo veremos el concepto de derivada de Fréchet y de Gâteau en espacios normados, haciendo referencia y generalizando conceptos que ya habíamos visto para funciones reales de varias variables reales. Introduciremos también las propiedades de la diferencial, que nos servirán de ayuda para estudiar la diferenciabilidad de funciones partiendo de otras más elementales, así como resultados de gran utilidad como puede ser el Teorema del valor medio.

Cabe destacar que durante el desarrollo de todo el capítulo sólo pediremos que nuestros espacios sean normados. Como los espacios de Banach son espacios normados, es claro que todos los resultados se pueden aplicar a estos.

Nos guiaremos por el Capítulo 1 de [4] y el Capítulo 2 de [7] para explicar todos estos conceptos.

2.1. El concepto de diferencial

Empezaremos esta sección recordando el concepto de diferencial para funciones reales de una o varias variables reales y extendiendo este concepto a espacios más generales como son los espacios normados. Veremos además los paralelismos existentes a la hora de definir los conceptos con los que trabajaremos en estos nuevos espacios.

Definición 2.1. Dada una función $f : \mathbb{R} \to \mathbb{R}$ diremos que es derivable en un punto $a \in \mathbb{R}$ si existe el siguiente límite

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

y es un número real. En este caso, diremos que este límite es la derivada de la función en el punto a y la denotaremos por $f'(a)$.

Diremos que una función f es derivable si lo es en todo punto de su dominio.

Este concepto de derivada nos permite expresar la idea geométrica de tangencia y nos indica que una función f es derivable en un punto a si y solo si su gráfica admite una recta tangente en ese punto y, además, $f'(a)$ es la pendiente de la recta tangente a f en a.

Una definición equivalente a la vista sobre la derivabilidad de una función, y que al generalizarla más adelante nos será de ayuda a la hora de demostrar algunos resultados, es la siguiente.

Definición 2.2. Diremos que la función f es derivable en $a \in \mathbb{R}$ si existe un número real $f'(a)$ tal que

$$|f(a+h) - f(a) - f'(a)h| = \phi(h),$$

donde ϕ es una función que ha de estar definida en un entorno de 0 y además

$$\lim_{h \to 0} \frac{|\phi(h)|}{|h|} = 0.$$

Podemos extender la noción de diferencial a funciones de la forma $f : \mathbb{R}^n \to \mathbb{R}$ y tendríamos la siguiente definición.

Definición 2.3. Diremos que $f : \mathbb{R}^n \to \mathbb{R}$ es diferenciable en un punto $a \in \mathbb{R}^n$ si, y solo si, existe una aplicación lineal que denotaremos por $Df(a)$ tal que

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

o, lo que es lo mismo,

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - Df(a)(h)|}{\|h\|} = 0.$$

La aplicación $Df(a)$ recibe el nombre de diferencial de f en a.

Otra definición equivalente de diferenciabilidad sería la siguiente.

Definición 2.4. Diremos que la función f es diferenciable en un punto $a \in \mathbb{R}^n$ si, y solo si, existe una aplicación lineal que denotaremos por $Df(a)$ y una función $\phi : A \subset \mathbb{R}^n \to \mathbb{R}$ tal que

1. ϕ está definida en un entorno del 0.
2. $\lim_{h \to 0} \dfrac{|\phi(h)|}{\|h\|} = 0.$
3. $f(a+h) = f(a) + Df(a)(h) + \phi(h).$

En el caso en el que la función f sea diferenciable en todo $a \in \mathbb{R}^n$, diremos que es diferenciable.

Podemos seguir generalizando este concepto a funciones $f : \mathbb{R}^n \to \mathbb{R}^m$, donde el operador $Df(a)$, que ya hemos dicho que se trata de una aplicación lineal, está definido por lo que denotaremos como la matriz jacobiana de f en el punto a.

Definición 2.5. Diremos que la función $f = (f_1, \ldots, f_m)$ es diferenciable en el punto a si, y solo si, todas las funciones f_i lo son en ese punto. En tal caso, la matriz que define la aplicación diferencial se llama matriz jacobiana y viene dada por

$$Jf(a) = \begin{pmatrix} D_1 f_1(a) & \cdots & D_n f_1(a) \\ \ldots & \cdots & \cdots \\ D_1 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix},$$

donde $D_i f_j(a)$ denota la derivada i-ésima de la componente j-ésima de f.

A partir de este punto estudiaremos este concepto de una manera más general para funciones definidas en un conjunto abierto de un espacio vectorial E que consta de una norma y que toma valores en otro espacio vectorial normado F. Las normas que definiremos en cada espacio son $|\cdot|$ y $\|\cdot\|$, respectivamente, que pueden coincidir o ser diferentes. Por tanto de ahora en adelante, salvo que se indique lo contrario, E y F denotarán dos espacios vectoriales normados.

Definición 2.6. Sean $U \subset E$ un subconjunto abierto y $f : U \to F$ una función. Diremos que f es Fréchet diferenciable en $a \in U$ si existe una transformación lineal acotada $Df(a)$, conocida como derivada o diferencial de Fréchet de f en el punto a, tal que

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - Df(a)(h)\|}{|h|} = 0.$$

Una definición equivalente a esta sería la siguiente.

Definición 2.7. Diremos que $f : U \subset E \to F$ es Fréchet diferenciable en $a \in U$ si existen una transformación lineal acotada $Df(a)$ y una función $\phi : A \subset E \to F$ tal que

1. ϕ está definida en un entorno del 0.
2. $\lim\limits_{h \to 0} \dfrac{\phi(h)}{|h|} = 0.$
3. $f(a+h) - f(a) - Df(a)(h) = \phi(h).$

Podemos ver el concepto de diferenciabilidad de una función entre espacios vectoriales normados de una manera geométrica, igual que veíamos en funciones reales de una variable real. Para desarrollar este nuevo punto de vista utilizaremos los Capítulos 2 y 3 de [1] y el Capítulo 8 de [3].

Empezaremos introduciendo una definición necesaria a la hora de exponer la nueva caracterización para la diferenciabilidad de funciones.

Definición 2.8. Dados $U \subset E$ un abierto y dos funciones, $f_1 : U \to F$ y $f_2 : U \to F$, diremos que f_1 y f_2 son tangentes en un punto $a \in U$ si

$$m(r) = \sup_{|x-a|\leq r} \|f_1(x) - f_2(x)\|,$$

definida para todo $r > 0$, cumple que

$$\lim_{r\to 0^+} \frac{m(r)}{r} = 0.$$

Observación 2. La relación que cumplen dos funciones al ser tangentes entre sí en un punto de su dominio es una relación de equivalencia. En efecto, dadas tres aplicaciones, $f, g, h : E \to F$, si f y g son tangentes en $a \in E$ y g y h son tangentes en a, entonces f y h son tangentes en a. Esto se deduce de la siguiente desigualdad

$$\|f(x) - h(x)\| \leq \|f(x) - g(x)\| + \|g(x) - h(x)\|.$$

Escribimos ahora la definición de diferenciabilidad teniendo en cuenta este concepto previo.

Definición 2.9. Dados $U \subset E$ un abierto y una función continua $f : U \subset E \to F$, se dice que f es diferenciable en un punto $a \in U$ si existe una aplicación lineal $g : E \to F$, tal que

$$x \mapsto f(a) + g(x - a)$$

es tangente a f en el punto a. La aplicación g es lo que se conoce como diferencial de f en a.

Una definición sencilla con la que trabajaremos más adelante a la hora de introducir el Teorema del valor medio y que utilizaremos ahora en un ejemplo es la siguiente.

Definición 2.10. Definimos el segmento cerrado que une los puntos $a, b \in E$, donde E es un espacio normado, como el conjunto siguiente

$$L[a, b] = \{x \in E \mid x = (1 - t)a + tb,\ 0 \leq t \leq 1\}.$$

El segmento abierto $L(a, b)$ se define de manera similar usando $0 < t < 1$.

Veremos a continuación un par de ejemplos en los que calcularemos la diferencial de Fréchet de algunas funciones.

Ejemplo 2.11. Sean $a, b \in E$ tales que $L[a,b] \subset E$. Dada $\alpha : [0,1] \subset \mathbb{R} \to L[a,b]$ una función definida como $\alpha(t) = a + t(b-a)$, veamos que $D\alpha(t) = \beta$, con $\beta(h) = (b-a) \cdot h$ para todo $t \in [0,1]$.

$$\begin{aligned}
&\lim_{h\to 0} \frac{\|\alpha(t+h) - \alpha(t) - D\alpha(t)(h)\|}{|h|} \\
&= \lim_{h\to 0} \frac{\|a + t(b-a) + h(b-a) - a - t(b-a) - h(b-a)\|}{|h|} = 0.
\end{aligned}$$

Queda probado así que $D\alpha(t) = \beta$ para todo $t \in [0,1]$.

Ejemplo 2.12. Sea $(E, \|\cdot\|)$, donde $E = \mathcal{C}([0,1])$ y $\|\cdot\| = \sup_{x\in[0,1]} |f(x)|$, un espacio de Banach. Queremos estudiar la diferenciabilidad de la función

$$\begin{aligned}
F : E &\to E \\
f &\rightsquigarrow f^2.
\end{aligned}$$

Dada $f_0 \in E$, veamos que la aplicación diferencial de F en f_0 viene dada por

$$\begin{aligned}
DF(f_0) : E &\to E \\
h &\rightsquigarrow DF(f_0)(h) = 2 \cdot f_0 \cdot h.
\end{aligned}$$

En efecto, aplicando la Definición 2.6

$$\begin{aligned}
\lim_{h\to 0} \frac{\|F(f_0+h) - F(f_0) - DF(f_0)(h)\|}{\|h\|} &= \lim_{h\to 0} \frac{\|f_0^2 + 2f_0h + h^2 - f_0^2 - 2f_0h\|}{\|h\|} \\
&= \lim_{h\to 0} \frac{\|h^2\|}{\|h\|} = 0.
\end{aligned}$$

Es fácil ver que en el ejemplo anterior la norma que definimos en el espacio no es realmente importante a la hora de ver si la función es diferenciable, ya que si ahora tomamos otra norma diferente, veremos que obtenemos el mismo resultado.

Ejemplo 2.13. Tengamos de nuevo $(E, \|\cdot\|)$, donde $E = \mathcal{C}([0,1])$, pero en este caso $\|\cdot\| = \int_0^1 |f(x)|dx$, un espacio de Banach. Queremos estudiar la diferenciabilidad de la función del ejemplo anterior, $F(f) = f^2$.

Dada $f_0 \in E$, veamos que la función diferencial viene dada por $DF(f_0)(h) = 2 \cdot f_0 \cdot h$. De nuevo, aplicando la Definición 2.6 tenemos que

$$\begin{aligned}
\lim_{h\to 0} \frac{\|F(f_0+h) - F(f_0) - DF(f_0)(h)\|}{\|h\|} &= \lim_{h\to 0} \frac{\|f_0^2 + 2f_0h + h^2 - f_0^2 - 2f_0h\|}{\|h\|} \\
&= \lim_{h\to 0} \frac{\|h^2\|}{\|h\|} = 0.
\end{aligned}$$

Tenemos las siguientes consecuencias de las definiciones dadas.

Proposición 2.14. *Dados $U \subset E$ un subconjunto abierto y $f : U \to F$, se cumple que:*

I. *Si f es diferenciable en $a \in U$, entonces la función es continua en ese punto.*

II. *Si f es diferenciable en $a \in U$, entonces la aplicación $Df(a) \in \mathcal{L}(E, F)$ es única.*

Demostración.

I. Sabemos por hipótesis que la función es diferenciable, por lo que por la Definición 2.7 existen una transformación lineal acotada $Df(a)$ y una función $\phi : A \subset E \to F$ tales que

   a) ϕ está definida en un entorno del 0.
   b) $\lim_{h \to 0} \dfrac{\phi(h)}{|h|} = 0$.
   c) $f(a+h) - f(a) = Df(a)h + \phi(h)$.

   Como el operador $Df(a) \in \mathcal{L}(E, F)$ tomamos la norma que habíamos definido en este espacio y por su definición sabemos que

$$\|Df(a)(h)\| \leq \|Df(a)\|_{\mathcal{L}} \cdot |h| \text{ para todo } h \in E.$$

   Aplicándolas a la expresión de *c*) se tiene que

$$\|f(a+h) - f(a)\| = \|Df(a)(h) + \phi(h)\| \leq \|Df(a)\|_{\mathcal{L}} \cdot |h| + \|\phi(h)\|,$$

   de donde, operando y sacando factor común $|h|$ obtenemos que

$$\|f(a+h) - f(a)\| \leq \left(\|Df(a)\|_{\mathcal{L}} + \frac{\|\phi(h)\|}{|h|} \right) |h|.$$

   Y es evidente que

$$\lim_{h \to 0} \left(\|Df(a)\|_{\mathcal{L}} + \frac{\|\phi(h)\|}{|h|} \right) |h| = 0$$

   ya que, por la definición de diferenciabilidad, el segundo sumando se va a ir a 0.

   Por lo tanto podemos concluir que $\lim_{h \to 0} f(a+h) = f(a)$, es decir, f es continua en el punto a.

II. Por hipótesis tenemos que f es diferenciable en a y por lo tanto satisface las condiciones de la Definición 2.7. Suponemos que tenemos $(Df)_1, (Df)_2 \in \mathcal{L}(E, F)$ y las respectivas aplicaciones ϕ_1 y ϕ_2 que deben estar definidas en un entorno del 0 y cumplir que $\lim_{h \to 0} \frac{\phi_i(h)}{|h|} = 0$, para $i = 1, 2$. Veamos que coinciden para poder demostrar su unicidad.

Tenemos que

$$f(a+h) - f(a) = (Df)_1(a)(h) + \phi_1(h),$$
$$f(a+h) - f(a) = (Df)_2(a)(h) + \phi_2(h),$$

de donde se deduce que

$$((Df)_1(a) - (Df)_2(a))(h) = \phi_1(h) - \phi_2(h)$$

para los elementos h de un cierto entorno de 0. Además, por la linealidad de $(Df)_1(a)$ y $(Df)_2(a)$, si $y \neq 0$ es un elemento arbitrario de E, se puede escribir

$$\begin{aligned} \|((Df)_1(a) - (Df)_2(a))(y)\| &= \frac{\|((Df)_1(a) - (Df)_2(a))(ty)\|}{|t|} \\ &= |y| \frac{\|\phi_1(ty) - \phi_2(ty)\|}{|t||y|}. \end{aligned}$$

Como esto se cumple para todo t, lo hará también en el caso de que tomemos el límite cuando t tiende a 0. Por lo que se tendría

$$\|((Df)_1(a) - (Df)_2(a))(y)\| \leq |y| \left(\lim_{t \to 0} \frac{\|\phi_1(ty)\|}{|t||y|} + \lim_{t \to 0} \frac{\|\phi_2(ty)\|}{|t||y|} \right) = 0,$$

que sabemos que es cierto por como habíamos definido las funciones ϕ_i para $i = 1, 2$. Por lo tanto tenemos que $(Df)_1 = (Df)_2$.

□

De igual manera que hablamos sobre el concepto de diferencial para funciones reales de varias variables reales para introducir el término de derivada de Fréchet y ver así su relación, vamos a introducir ahora el término de derivada direccional y definiremos a continuación la derivada de Gâteau, que se corresponde con una generalización de esta.

Si partimos de una función $f : \mathbb{R}^2 \to \mathbb{R}$ se define su derivada direccional en $a \in \mathbb{R}^2$ según el vector unitario $v = (v_1, v_2) \in \mathbb{R}^2$, como

$$D_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}.$$

Podemos ahora introducir el concepto para espacios vectoriales normados arbitrarios E y F.

Definición 2.15. Se define la derivada de Gâteau de $f : E \to F$ en el punto $a \in E$ a lo largo de un vector dado $h \in E$ como el siguiente límite

$$df(a;h) = \lim_{t \to 0} \frac{f(a+th) - f(a)}{t}. \tag{2.1}$$

Diremos que f es Gâteau diferenciable en a si existe la derivada de Gâteau para todo $h \in E$.

Se tiene la siguiente relación entre la diferenciabilidad de Fréchet y la de Gâteau.

Proposición 2.16. *Si f es Fréchet diferenciable en un punto a, entonces es Gâteau diferenciable en el punto y la derivada de Gâteau viene dada por $Df(a)(h)$.*

Demostración. Sabemos por hipótesis que la función f es Fréchet diferenciable y sean $h \in E$ y $t \in \mathbb{R}$. Por la Definición 2.7 existen una transformación lineal acotada $Df(a)$ y una función $\phi : A \subset E \to F$ tal que

1. ϕ está definida en un entorno del 0.
2. $\lim_{th \to 0} \frac{\phi(th)}{|th|} = 0.$
3. $f(a + th) - f(a) = Df(a)(th) + \phi(th).$

Operando y tomando límites nos quedaría

$$\lim_{t \to 0} \frac{f(a + th) - f(a)}{t} = \lim_{t \to 0} \left(Df(a)(h) + h\frac{\phi(th)}{th} \right) = Df(a)(h).$$

Por lo tanto queda probado que si la función es Fréchet diferenciable en a, entonces existe la derivada de Gâteau en ese punto y que $df(a; h) = Df(a)(h)$. □

El recíproco de la proposición anterior no es cierto, ya que una función puede tener derivada de Gâteau en un punto a lo largo de cualquier dirección y no ser Fréchet diferenciable en ese punto. Lo veremos a través del siguiente ejemplo.

Ejemplo 2.17. En este ejemplo nuestros espacios pasan a ser $E = \mathbb{R}^2$ y $F = \mathbb{R}$, y por lo tanto se trata de una función $f : \mathbb{R}^2 \to \mathbb{R}$ definida como sigue:

$$f(x, y) = \begin{cases} \frac{2xy^2}{x^2+y^4} & \text{si} \quad (x, y) \neq (0, 0), \\ 0 & \text{si} \quad (x, y) = (0, 0). \end{cases}$$

Probaremos que no es diferenciable en el punto $(0, 0)$ viendo que no es continua en él. Sabemos que la función f será continua en $(0, 0)$ si se cumple que

$$f(0, 0) = \lim_{(x,y) \to (0,0)} f(x, y).$$

Por lo tanto bastaría buscar un subconjunto a través del cual el límite sea distinto de $f(0, 0) = 0$, ya que en caso de que este existiese tendría que coincidir en todos los subconjuntos. Tomando $x = y^2$ tendríamos lo siguiente

$$\lim_{\substack{(x,y) \to (0,0) \\ x = y^2}} \frac{2xy^2}{x^2 + y^4} = \lim_{y \to 0} \frac{2y^4}{y^4 + y^4} = 1 \neq f(0, 0).$$

Por lo tanto como f no es continua en $(0,0)$, tampoco será Fréchet diferenciable, ya que como habíamos visto en la Proposición 2.14, que sea continua es condición necesaria para la diferenciabilidad. Pero, a pesar de esto, sí existe la derivada de Gâteau en el punto a lo largo de cualquier dirección h. Utilizaremos la expresión (2.1)

$$\lim_{t\to 0} \frac{f(th_1, th_2) - f(0,0)}{t} = \lim_{t\to 0} \frac{2t^3 h_1 h_2^2}{t\left((th_1)^2 + (th_2)^4\right)} = \lim_{t\to 0} \frac{2h_1 h_2^2}{h_1^2 + th_2^2},$$

donde tendremos que distinguir dos casos:

1. El caso en el que $h_1 = 0$ y entonces $\lim\limits_{t\to 0} \dfrac{0}{th_2^2} = 0$.
2. El caso en el que $h_1 \neq 0$ y entonces $\lim\limits_{t\to 0} \dfrac{2h_1 h_2^2}{h_1^2 + th_2^2} = \dfrac{2h_2^2}{h_1}$.

Concluimos que a pesar de que la función no sea Fréchet diferenciable, existe la derivada de Gâteau en todas las direcciones.

2.2. Propiedades de la diferencial

En esta sección veremos ciertas propiedades que cumple la diferencial, ya sea a la hora de componer y sumar funciones, al multiplicarlas por escalares o si el espacio de salida o de llegada se pueden escribir como espacios producto de espacios normados.

2.2.1. Reglas del cálculo

En esta parte del capítulo enunciaremos una serie de resultados que nos ayudarán a la hora de estudiar la diferenciabilidad de aplicaciones a partir de otras en las que esta propiedad se pueda estudiar de forma elemental.

Proposición 2.18 (Regla de la cadena). *Sean E, F, y G espacios vectoriales normados, y sean U y V dos conjuntos abiertos de E y F respectivamente. Sean $f : U \to F$ y $g : V \to G$ tales que para un punto $a \in U$ dado, tenemos que $f(a) = b \in V$. Si f es diferenciable en a y g lo es en b, entonces $h = g \circ f$ es diferenciable en a y además*

$$Dh(a) = Dg(f(a)) \circ Df(a).$$

Demostración. A lo largo de esta demostración seguiremos utilizando las normas, $|\cdot|$ y $\|\cdot\|$, que habíamos denotado para los espacios E y F, respectivamente. Además, trabajemos con una nueva norma en el espacio G denotada por $\|\cdot\|_G$.

Por hipótesis sabemos que f es diferenciable en a y que g lo es en b. Partiendo entonces de la Definición 2.7, sabemos que existen $Df(a)$ y $Dg(b)$ aplicaciones lineales y dos funciones $\phi_f : A \subset E \to F$ y $\phi_g : B \subset F \to G$ tales que

1. ϕ_f y ϕ_g están definidas en un entorno del 0.
2. $\lim_{x\to a} \dfrac{\phi_f(x-a)}{|x-a|} = 0$ y $\lim_{y\to b} \dfrac{\phi_g(y-b)}{\|y-b\|} = 0$.
3. Se tienen las siguientes igualdades:

$$\begin{aligned} f(x)-f(a) &= Df(a)(x-a)+\phi_f(x-a), \\ g(y)-g(b) &= Dg(b)(y-b)+\phi_g(y-b). \end{aligned}$$

Por como habíamos definido la aplicación $h = g \circ f$ llegamos a que

$$\begin{aligned} h(x)-h(a) = g(f(x))-g(f(a)) &= Dg(f(a))(f(x)-f(a))+\phi_g(f(x)-f(a)) \\ &= Dg(f(a))(Df(a)(x-a))+Dg(f(a))(\phi_f(x-a))+\phi_g(f(x)-f(a)). \end{aligned} \tag{2.2}$$

Si demostramos que

$$\lim_{x\to a} \frac{\|h(x)-h(a)-Dg(f(a))(Df(a)(x-a))\|_G}{|x-a|} = 0 \tag{2.3}$$

quedaría demostrado que nuestra función h es diferenciable en a y además se tiene que $Dh(a) = Dg(f(a)) \circ Df(a)$.

Desarrollando y teniendo en cuenta la ecuación (2.2) es fácil ver que

$$\begin{aligned} &\frac{\|h(x)-h(a)-Dg(f(a))(Df(a)(x-a))\|_G}{|x-a|} \\ &= \frac{\|Dg(f(a))(\phi_f(x-a))+\phi_g(f(x)-f(a))\|_G}{|x-a|} \\ &\leq \frac{\|Dg(f(a))(\phi_f(x-a))\|_G+\|\phi_g(f(x)-f(a))\|_G}{|x-a|} \\ &\leq \frac{\|Dg(f(a))\|_{\mathcal{L}}\cdot\|\phi_f(x-a)\|}{|x-a|}+\frac{\|\phi_g(f(x)-f(a))\|_G}{|x-a|}. \end{aligned}$$

Aplicando la segunda condición de la Definición 2.7, que habíamos enunciado antes, se tiene que

$$\begin{aligned} \lim_{x\to a} \frac{\|Dg(f(a))\|_{\mathcal{L}}\cdot\|\phi_f(x-a)\|}{|x-a|} &= 0, \\ \lim_{x\to a} \frac{\|\phi_g(f(x)-f(a))\|_G}{|x-a|} &= 0. \end{aligned}$$

Veamos que efectivamente esta segunda igualdad se cumple, y para ello podemos escribirlo como

$$\lim_{x\to a} \frac{\|\phi_g(f(x)-f(a))\|_G}{|x-a|} = \lim_{x\to a} \frac{\|\phi_g(f(x)-f(a))\|_G}{\|f(x)-f(a)\|}\cdot\frac{\|f(x)-f(a)\|}{|x-a|}.$$

Separamos ambos factores para poder trabajar con ellos de manera independiente. Para la primera expresión tendremos en cuenta el cambio de variable $f(x) = y$,

$$\lim_{x\to a} \frac{\|\phi_g(f(x) - f(a))\|_G}{\|f(x) - f(a)\|} = \lim_{y\to b} \frac{\|\phi_g(y - b)\|_G}{\|y - b\|} = 0.$$

Ahora, con la segunda expresión

$$\lim_{x\to a} \frac{\|f(x) - f(a)\|}{|x - a|} = \lim_{x\to a} \frac{\|Df(a)(x - a) + \phi_f(x - a)\|}{|x - a|}$$
$$\leq \lim_{x\to a} \frac{\|Df(a)\|_{\mathcal{L}} \cdot |x - a|}{|x - a|} + \frac{\|\phi_f(x - a)\|}{|x - a|} = \|Df(a)\|_{\mathcal{L}}.$$

Por lo tanto podemos afirmar que

$$\lim_{x\to a} \frac{\|\phi_g(f(x) - f(a))\|_G}{|x - a|} = \|Df(a)\|_{\mathcal{L}} \cdot 0 = 0.$$

De esta manera queda probada la igualdad de la ecuación (2.3) y por lo tanto concluímos la demostración. □

Proposición 2.19. *Dadas dos aplicaciones f, $g : E \to F$ diferenciables en un punto $a \in E$, se tiene que*

1. *$D(f + g)(a) = Df(a) + Dg(a)$.*
2. *$D(\lambda f)(a) = \lambda Df(a)$, donde $\lambda \in \mathbb{R}$.*
3. *$D(fg)(a) = f(a)Dg(a) + g(a)Df(a)$.*
4. *Si además se cumple que $f(a) \neq 0$, se tiene que*

$$D\left(\frac{g}{f}\right)(a) = \frac{f(a)Dg(a) - g(a)Df(a)}{f^2(a)}.$$

Proposición 2.20. *Sea $f : E \to F$ una aplicación constante, entonces es diferenciable en E y $Df(a) = 0$ para todo $a \in E$.*

Demostración. Por hipótesis se tiene que la función f es constante, y por lo tanto toma el mismo valor en todos los puntos de su dominio. Veamos que $Df(a) = 0$ para todo $a \in E$. En efecto,

$$\lim_{h\to 0} \frac{\|f(a + h) - f(a) - 0\|}{|h|} = \lim_{h\to 0} \frac{\|f(a + h) - f(a)\|}{|h|} = 0$$

por ser $f(a + h) = f(a)$ para todo h. □

Veamos una última proposición en la que pediremos que la función con la que vamos a trabajar sea una aplicación lineal.

Proposición 2.21. *Si $f : E \to F$ es una aplicación lineal, entonces es diferenciable en E y $Df(a) = f$ para todo $a \in E$.*

Demostración. Aplicando la Definición 2.6 se tiene que

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - f(h)\|}{|h|} = \lim_{h\to 0} \frac{\|f(a) + f(h) - f(a) - f(h)\|}{|h|} = 0.$$

□

2.2.2. Espacios producto

Podemos considerar situaciones especiales donde la aplicación f tiene como dominio o codominio espacios producto normados.

Empezaremos considerando el caso en el que podemos escribir $F = F_1 \times \cdots \times F_m$, como el producto de espacios vectoriales normados. Para $1 \leq i \leq m$, definimos la aplicación proyección

$$\begin{aligned} p_i : F &\longrightarrow F_i \\ x = (x_1, \ldots, x_m) &\longmapsto p_i(x) = x_i \end{aligned}$$

y sea $u_i : F_i \to F$ la aplicación inclusión definida por

$$u_i(x_i) = (0, \ldots, 0, x_i, 0, \ldots, 0),$$

donde x_i se encuentra en la posición i-ésima. Entonces notemos que se tienen las siguientes relaciones

$$\begin{cases} p_i \circ u_i = I_{F_i}, \\ \sum_{i=1}^m u_i \circ p_i = I_F, \end{cases}$$

donde I_{F_i} e I_F denotan la aplicación identidad en los espacios F_i y F, respectivamente.

Una propiedad que cumple la aplicación proyección vendría enunciada en el siguiente resultado, que será de gran utilidad en demostraciones posteriores.

Por simplicidad asumiremos a partir de aquí que $\|\cdot\|$ representa tanto la norma en F como en F_i.

Proposición 2.22. *Sean $F = F_1 \times \cdots \times F_m$ un espacio producto de espacios vectoriales normados y $p_i : F \to F_i$ la aplicación proyección, se tiene que $Dp_i(a) = p_i$, para todo $a \in F$.*

Demostración. Se tiene que:

$$\lim_{h\to 0} \frac{\|p_i(a+h) - p_i(a) - p_i(h)\|}{\|h\|} = \lim_{h\to 0} \frac{\|a_i + h_i - a_i - h_i\|}{\|h\|} = 0.$$

Queda probado entonces $Dp_i(a) = p_i$ para todo $a \in F$. □

Hemos visto una propiedad de la aplicación proyección y enunciaremos ahora una muy similar para la aplicación inclusión.

Lema 2.23. *Sean $F = F_1 \times \cdots \times F_m$ un espacio producto de espacios vectoriales normados y $u_i : F_i \to F$ la aplicación inclusión, se tiene que $Du_i(a) = u_i$, para todo $a \in F_i$.*

Demostración. Se tiene que:

$$\begin{aligned}
&\lim_{h\to 0} \frac{\|u_i(a+h) - u_i(a) - u_i(h)\|}{\|h\|} \\
&= \lim_{h\to 0} \frac{\|(0,\ldots,0,a+h,0,\ldots,0) - (0,\ldots,0,a,0,\ldots,0) - (0,\ldots,0,h,0,\ldots,0)\|}{\|h\|} = 0.
\end{aligned}$$

Queda probado entonces $Du_i(a) = u_i$ para todo $a \in F$. □

Proposición 2.24. *Dada una función $f : U \subset E \to F$, con $F = F_1 \times \cdots \times F_m$, diremos que es diferenciable en un punto $a \in U$ si, y solo si, $f_i = p_i \circ f : U \to F_i$ es diferenciable en a para cada i, $1 \le i \le m$. En este caso*

$$Df(a) = \sum_{i=1}^{m} u_i \circ Df_i(a). \tag{2.4}$$

Demostración. "⇒" Por hipótesis tenemos que la función f es diferenciable en a y, además, p_i es lineal y continua y por lo tanto diferenciable en todo punto y, en particular, en $f(a)$. Como la composición de diferenciables es diferenciable, queda probado que f_i lo es en a. Si ahora aplicamos la regla de la cadena y la Proposición 2.22

$$Df_i(a) = Dp_i \circ Df(a) = p_i \circ Df(a).$$

"⇐" Recíprocamente, si f_i es diferenciable para cada uno de los i, tenemos que

$$f = \sum_{i=1}^{m} u_i \circ f_i$$

y de nuevo, como u_i es una aplicación lineal, se tiene que f es diferenciable por ser suma y composición de aplicaciones diferenciables y además su diferencial viene dada por (2.4). Veamos que efectivamente esto es así. En efecto,

$$\begin{aligned}
Df(a) &= D\left(\sum_{i=1}^{m} u_i \circ f_i\right)(a) = \sum_{i=1}^{m} D(u_i \circ f_i)(a) \\
&= \sum_{i=1}^{m} Du_i(f_i(a)) \circ Df_i(a) = \sum_{i=1}^{m} u_i \circ Df_i(a).
\end{aligned}$$

Esta última igualdad se tiene por el Lema 2.23. □

Suponemos ahora que el espacio producto se corresponde con el espacio de salida, es decir, podemos escribir $E = E_1 \times \cdots \times E_n$. Sea $U \subset E$ un subconjunto abierto y $f : U \to F$ es una aplicación dada. Si fijamos un punto $a = (a_1, \ldots, a_n) \in E$, se puede definir una aplicación que lleva a cada elemento de E_i en E, $\lambda_i : E_i \to E$, como

$$\lambda_i(x_i) = (a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n).$$

Veremos ahora qué relación existe entre la diferenciabilidad de la aplicación f y la que se restringe a cada subespacio, lo que nos ayudará al estudio de la diferenciabilidad de un grupo más amplio de funciones de manera más sencilla.

De igual manera que antes, a partir de este momento, por simplicidad, denotaremos por $|\cdot|$ a la norma definida en E y E_i.

Proposición 2.25. *Si la función f es diferenciable en $a \in U$, entonces la composición $f \circ \lambda_i$ es diferenciable en a_i para todo $i = 1, \ldots, n$. Además*

$$Df(a)(h_1, \ldots, h_n) = \sum_{i=1}^{n} D(f \circ \lambda_i)(a_i)(h_i)$$

para cualquier $h = (h_1, \ldots, h_n) \in E$.

Demostración. Ya habíamos definido la aplicación inclusión del espacio E_i en E, entonces la aplicación λ_i se puede escribir como

$$\lambda_i(x_i) = (a_1, \ldots, a_{i-1}, a_i + x_i - a_i, a_{i+1}, \ldots, a_n) = a + u_i(x_i - a_i).$$

Comprobemos que la diferencial de λ_i en el punto x_i, se corresponde con

$$D\lambda_i(x_i) = u_i, \text{ para todo } x_i \in E_i.$$

En efecto, por la definición de diferencial en un punto se tiene que

$$\begin{aligned} &\lim_{h \to 0} \frac{|\lambda_i(x_i + h) - \lambda_i(x_i) - u_i(h)|}{|h|} \\ &= \lim_{h \to 0} \frac{|a + u_i(x_i + h - a_i) - a - u_i(x_i - a_i) - u_i(h)|}{|h|} = \lim_{h \to 0} \frac{|0|}{|h|} = 0. \end{aligned}$$

Queda así probado que $D\lambda_i(x_i) = u_i$ para todo $x_i \in E_i$.

Vemos así que λ_i es diferenciable en a_i y como f por hipótesis también lo era en el punto a, la composición $f \circ \lambda_i$ será diferenciable en a_i y se expresa como

$$D(f \circ \lambda_i)(a_i) = Df(\lambda_i(a_i)) \circ D(\lambda_i(a_i)) = Df(a) \circ u_i.$$

Por las relaciones que enunciamos antes se tiene que

$$\sum_{i=1}^{n} u_i \circ p_i = I_E$$

y si componemos esta expresión con $Df(a)$ nos da que

$$\sum_{i=1}^{n} (Df(a) \circ u_i) \circ p_i = Df(a).$$

Por tanto

$$\begin{aligned} Df(a)(h_1, \ldots, h_n) &= \left(\sum_{i=1}^{n} (Df(a) \circ u_i) \circ p_i \right) (h_1, \ldots, h_n) \\ &= \sum_{i=1}^{n} (Df(a) \circ u_i)(h_i) = \sum_{i=1}^{n} D(f \circ \lambda_i)(a_i)(h_i). \end{aligned}$$

Queda probado que $f \circ \lambda_i$ es diferenciable y se da la igualdad del enunciado. $\square$

Esta proposición nos permite introducir el concepto de derivada parcial en un punto.

Definición 2.26. A la derivada de $f \circ \lambda_i$ en el punto a_i la denotamos como la derivada parcial i-ésima de f en el punto a. A partir de este momento la denotaremos por $\frac{\partial f}{\partial x_i}(a)$ o por $\partial_i f(a)$.

Observación 3. Puede ocurrir que existan todas las derivadas parciales de la función f en un punto pero esta no sea diferenciable en él. Será necesario añadir algunas condiciones a mayores para que la diferenciabilidad de la función se pueda estudiar a través de las derivadas parciales.

En la siguiente sección introduciremos el Teorema del valor medio, ya que será necesario conocerlo para poder demostrar los resultados que relacionan las derivadas parciales con la diferenciabilidad de la función en un punto.

2.3. Teorema del valor medio

Profundizaremos en esta sección en el Teorema del valor medio, necesario a la hora de demostrar ciertos resultados. Para ello nos guiaremos también por el Capítulo 3 de [1].

Si nos encontramos en el caso en el que estamos estudiando funciones reales de variable real surge una de las primeras formas del Teorema del valor medio: Sea $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ una función continua en $[a, b]$ y diferenciable en (a, b), entonces existe $c \in (a, b)$ tal que

$$f(b) - f(a) = f'(c)(b - a).$$

Teniendo en cuenta la expresión de este resultado es fácil ver que relaciona las imágenes de la función f con sus antecedentes a través de la derivada. Además puede resultar más interesante si trabajásemos con una desigualdad, como puede ser en el caso de las funciones en las que su derivada está acotada en el intervalo (a, b), obteniendo una expresión de la forma

$$f(x_2) - f(x_1) \leq k(x_2 - x_1) \text{ para todo } x_1, x_2 \in (a, b),$$

que más adelante nos dará una condición a la hora de definir el concepto de funciones Lipschitz continuas en un intervalo.

El enunciado que nos da el Teorema del valor medio para funciones reales de una variable no se cumple en general si queremos trabajar en contextos más generales tales como serían los espacios normados arbitrarios. Un ejemplo sencillo de que esto puede ocurrir sería tomando, sin ir más lejos, el caso en el que $E = \mathbb{R}$ y $F = \mathbb{R}^2$ y una función definida como $f(t) = (\cos(t), \sin(t))$. Su matriz jacobiana vendría dada por

$$\nabla f(t) = (-\sin(t), \cos(t)).$$

Si calculamos el valor de f en los extremos del intervalo $[0, 2\pi]$, vemos fácilmente que $f(0) = f(2\pi) = (1, 0)$, pero nunca tendremos que $\nabla f(t) = (0, 0)$ para ningún $t \in (0, 2\pi)$. Por lo tanto no se cumplirá que $f(2\pi) - f(0) = 2\pi \nabla f(t)$ para ningún $t \in (0, 2\pi)$.

Ya habíamos introducido antes que había otras versiones que podrían resultar más interesantes y que como veremos a partir de ahora serán verdaderas para cualquier espacio arbitrario en el que estemos trabajando. Teniendo en cuenta lo anterior, de la expresión $f(b) - f(a) = Df(c)(b - a)$ deducimos que si f es una función real de variable real diferenciable en un intervalo $[a, b]$, entonces

$$|f(b) - f(a)| \leq \sup_{a<c<b} \|Df(c)\|_{\mathcal{L}} \cdot |b - a|.$$

Esta forma del Teorema del valor medio se podrá generalizar de manera sencilla. A la hora de demostrar esta generalización del teorema que veremos más adelante, serán necesarios una serie de conceptos previos.

Definición 2.27. Dada una aplicación $f : [a, b] \subset \mathbb{R} \to F$ diremos que admite una derivada por la derecha en un punto $x \in [a, b)$ si existe

$$Df_d(x) = \lim_{h \to 0^+} \frac{f(x + h) - f(x)}{h} \in F.$$

Df_d es lo que llamaremos la derivada por la derecha de la función f en el punto x.

De manera análoga podemos definir la derivada por la izquierda.

Definición 2.28. Dada una aplicación $f : [a,b] \subset \mathbb{R} \to F$ diremos que admite una derivada por la izquierda en un punto $x \in (a,b]$ si existe

$$Df_i(x) = \lim_{h \to 0^-} \frac{f(x+h) - f(x)}{h} \in F.$$

Df_i es lo que llamaremos la derivada por la izquierda de la función f en el punto x.

Una vez hemos definido estas dos derivadas, cabe destacar que se tiene el siguiente resultado.

Proposición 2.29. *Sea $f : [a,b] \subset \mathbb{R} \to F$ y $x \in [a,b]$. Se cumple que:*

1. *Si $x \in (a,b)$, entonces f es diferenciable en x si y solo si existen las derivadas por la derecha y por la izquierda en x y son iguales. En tal caso se tiene que*

$$Df_d(x) = Df_i(x) = Df(x)(1)$$

o, lo que es lo mismo, $Df(x)(h) = h \cdot Df_d(x) = h \cdot Df_i(x)$.

2. *Si $x = a$, entonces f es diferenciable en x si y solo si existe la derivada por la derecha en x. En tal caso se tiene que*

$$Df_d(x) = Df(x)(1)$$

o, lo que es lo mismo, $Df(x)(h) = h \cdot Df_d(x)$.

3. *Si $x = b$, entonces f es diferenciable en x si y solo si existe la derivada por la izquierda en x. En tal caso se tiene que*

$$Df_i(x) = Df(x)(1)$$

o, lo que es lo mismo, $Df(x)(h) = h \cdot Df_i(x)$.

Demostración. Lo demostraremos únicamente para *1*, ya que el resto de casos son análogos tomando el único límite lateral que tenga sentido.

"⇒" Suponemos que nuestra aplicación es diferenciable en x y por lo tanto se tiene que

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Df(x)(h)}{|h|} = 0 \in F. \tag{2.5}$$

Veamos que tanto la derivada por la derecha como por la izquierda se corresponden con $Df(x)(1)$ para todo $x \in (a,b)$. Lo probaremos para el caso de la derivada por la derecha, ya que el otro se prueba de manera análoga. Se tiene que

$$\lim_{h \to 0^+} \frac{f(x+h) - f(x) - Df(x)(h)}{h} = \lim_{h \to 0^+} \frac{f(x+h) - f(x) - h \cdot Df(x)(1)}{h}$$
$$= \lim_{h \to 0^+} \frac{f(x+h) - f(x)}{h} - Df(x)(1).$$

Entonces por (2.5) se deduce que existe

$$\lim_{h\to 0^+} \frac{f(x+h)-f(x)}{h} = Df(x)(1),$$

es decir, existe la derivada por la derecha de f en x y cumple que $Df_d(x) = Df(x)(1)$ para todo $x \in (a,b)$.

"$\Leftarrow$" Suponemos que existen las derivadas por la derecha y por la izquierda en x y que son iguales. Entonces f es diferenciable en x y su diferencial es $Df(x)(h) = h \cdot Df_d(x)$. En efecto, veamos que

$$\lim_{h\to 0} \frac{f(x+h)-f(x)-h\cdot Df_d(x)}{|h|} = 0. \tag{2.6}$$

Tomando el límite lateral por la derecha se tiene que

$$\lim_{h\to 0^+} \frac{f(x+h)-f(x)-h\cdot Df_d(x)}{|h|} = \lim_{h\to 0^+} \frac{f(x+h)-f(x)}{|h|} - Df_d(x) = 0.$$

De manera análoga, si tomamos el límite lateral por la izquierda se tiene que

$$\begin{aligned}\lim_{h\to 0^-} \frac{f(x+h)-f(x)-h\cdot Df_d(x)}{|h|} &= \lim_{h\to 0^-} \frac{f(x+h)-f(x)}{|h|} - Df_d(x) \\ = \lim_{h\to 0^-} \frac{f(x+h)-f(x)}{|h|} - Df_i(x) = 0.\end{aligned}$$

Por tanto se cumple (2.6) y concluímos que f es una función diferenciable en x y $Df(x)(h) = h \cdot Df_d(x) = h \cdot Df_i(x)$. □

El resultado anterior es de gran importancia, ya que caracteriza la diferencial de una función en un punto a partir de sus derivadas por la derecha e izquierda en ese punto.

A continuación veremos también que gracias a la derivada por la derecha podemos obtener una expresión que nos recuerda al Teorema del valor medio y por lo tanto será importante a la hora de su prueba.

Proposición 2.30. *Sean $f : [a,b] \subset \mathbb{R} \to F$ una aplicación continua, $k > 0$ una constante y supongamos que f admite una derivada por la derecha para todo $x \in (a,b)$ que cumple que*

$$\|Df_d(x)\| \le k \quad \textit{para todo} \quad x \in (a,b).$$

Entonces se tiene que

$$\|f(b)-f(a)\| \le k(b-a),$$

y, de manera general,

$$\|f(x_2)-f(x_1)\| \le k|x_2-x_1| \quad \textit{para todo} \quad x_1, x_2 \in [a,b].$$

Un último comentario ahora que tenemos las herramientas suficientes para poder demostrar el siguiente resultado, es que a lo largo de la demostración trabajaremos tanto en esapcios normados como en $\mathbb{R}$. Para que no haya confusión con las normas, en caso de trabajar en E, la norma la denotaremos por $|\cdot|_E$ y en el caso de $\mathbb{R}$ utilizaremos $|\cdot|$.

Teorema 2.31 (Teorema del valor medio). *Sean E y F espacios vectoriales normados y $U \subset E$ un subconjunto abierto. Sean $f : U \to F$ una función diferenciable en U y $a, b \in U$ tales que $L[a, b] \subset U$. Entonces*

$$\|f(b) - f(a)\| \leq |b - a|_E \sup_{x \in L[a,b]} \|Df(x)\|_{\mathcal{L}}.$$

Demostración. Por hipótesis sabemos que nuestra función es diferenciable en el segmento $L[a, b]$, por lo tanto podemos definir una función

$$\alpha : [0, 1] \subset \mathbb{R} \to L[a, b] \subset U \subset E,$$

de la forma $\alpha(t) = a + t(b - a)$ para todo $t \in [0, 1]$. Definimos ahora $h(t) = f(\alpha(t))$, que es diferenciable en $[0, 1]$ por ser composición de dos funciones diferenciables. Aplicando la regla de la cadena se tiene que

$$Dh(t) = Df(\alpha(t)) \circ D\alpha(t).$$

Para aligerar notación, denotamos $D\alpha(t_0) = \beta$, que por el Ejemplo 2.11 sabemos que es de la forma $\beta(t) = (b - a)t$, y escribimos

$$Dh(t) = Df(\alpha(t)) \circ \beta. \tag{2.7}$$

Entonces, aplicando normas en (2.7) y por la Proposición 1.12, se tiene que

$$\|Dh(t)\|_{\mathcal{L}} \leq \|Df(\alpha(t))\|_{\mathcal{L}} \cdot \|\beta\|_{\mathcal{L}}.$$

Además, como teníamos que $\beta(t) = t(b - a)$ y por como habíamos definido la norma $\|\cdot\|_{\mathcal{L}}$, vemos que

$$\|\beta\|_{\mathcal{L}} = \sup_{t \neq 0} \frac{|\beta(t)|_E}{|t|} = \sup_{t \neq 0} \frac{|t(b - a)|_E}{|t|} = \sup_{t \neq 0} \frac{|t| \cdot |b - a|_E}{|t|} = |b - a|_E.$$

Es claro entonces, teniendo todo esto en cuenta, que

$$\|Dh(t)\|_{\mathcal{L}} \leq \|Df(\alpha(t))\|_{\mathcal{L}} \cdot |b - a|_E.$$

Por como habíamos definido h sabemos que se trata de una aplicación $h : [0, 1] \to F$ y por lo tanto $Dh(t) : [0, 1] \to F$. Por la Proposición 2.29 se cumple que

$$Dh(t)(k) = k \cdot Dh_d(t) \quad \text{para todo} \quad x \in [0, 1].$$

Entonces,

$$\|Dh(t)\|_{\mathcal{L}} = \sup_{k\in[0,1]} \frac{\|Dh(t)(k)\|}{|k|} = \sup_{k\in[0,1]} \frac{\|k \cdot Dh_d(t)\|}{|k|} = \sup_{k\in[0,1]} \frac{|k| \cdot \|Dh_d(t)\|}{|k|} = \|Dh_d(t)\|.$$

Por lo tanto

$$\|Dh_d(t)\| \leq \|Df(\alpha(t))\|_{\mathcal{L}} \cdot |b-a|_E \leq \sup_{x\in L[a,b]} \|Df(x)\|_{\mathcal{L}} \cdot |b-a|_E.$$

Ahora, por la Proposición 2.30 se deduce que

$$\|h(1)-h(0)\| \leq \sup_{x\in L[a,b]} \|Df(x))\|_{\mathcal{L}} \cdot |b-a|_E \cdot (1-0)$$

o, lo que es lo mismo,

$$\|f(b)-f(a)\| \leq \sup_{x\in L[a,b]} \|Df(x)\|_{\mathcal{L}} \cdot |b-a|_E.$$

□

El concepto de una función Lipschitz continua en un intervalo ya apareció previamente. Veremos algún resultado relacionado con estas.

Definición 2.32. Decimos que un conjunto es convexo si dados dos puntos cualesquiera de él, el segmento que los une está contenido en el conjunto.

Corolario 2.33. *Si $U \subset E$ es un conjunto abierto convexo y $f : U \to F$ es diferenciable en todos los puntos de U y cumple que*

$$\|Df(x)\|_{\mathcal{L}} \leq k \quad \forall x \in U,$$

entonces es una función Lipschitz continua.

Corolario 2.34. *Sea $f : U \subset E \to F$ una función diferenciable en E. Sea U un conjunto conexo tal que $Df(x) = 0$ para todo $x \in U$. Entonces se tiene que f es una función constante.*

Demostración. Tomamos un punto $x_0 \in U$ y sea $B(x_0; r) \subset U$ la bola abierta en U de centro x_0 y radio $r > 0$. La bola que habíamos tomado $B(x_0; r)$ es un conjunto convexo y como por hipótesis $Df(x) = 0$ para todo x en $B(x_0; r)$, aplicando el Corolario 2.33 tenemos que f es una función Lipschitz continua. Tomando la constante de Lipschitz $k = 0$ se tiene que $f(x) = f(x_0)$ para todo $x \in B(x_0; r)$, lo que nos muestra que f es localmente constante.

Sea $b \in F$, como f es localmente constante sabemos que $f^{-1}(b)$ es abierto en U. Sabemos también que F es un espacio Hausdorff, por lo que $\{b\}$ será cerrado y como f es continua, $f^{-1}(b)$ es cerrado en U. Por hipótesis se tiene que U es conexo, es decir, que los únicos abiertos y cerrados del espacio son U o $\emptyset$, entonces es claro que $f^{-1}(b) = \emptyset$ o U. Si tomamos $f(x_0) = b$ para algún $x_0 \in U$, tenemos entonces que $f(x) = b$ para todo $x \in U$. Queda probado entonces que f es una función constante. □

Teniendo en cuenta todos estos nuevos conceptos podemos demostrar ahora el resultado que relaciona la diferenciabilidad de una función con sus derivadas parciales, que solo existirán si estamos trabajando en espacios producto. Será una condición suficiente para estudiar la diferenciabilidad de una función, que en muchas ocasiones nos facilitará las cosas.

Proposición 2.35. *Sean E y F dos espacios vectoriales normados, con $E = E_1 \times \cdots \times E_n$, y $f : U \subset E \to F$ una función. Si sus derivadas parciales $\frac{\partial f}{\partial x_i}$ existen en todos los puntos de U para cada $i = 1, \ldots, n$ y son continuas en el punto $a \in U$, entonces f es diferenciable en a.*

Demostración. Sabemos, por la Proposición 2.25, que si f es diferenciable en a, entonces se cumple que

$$Df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(h_i). \tag{2.8}$$

Vamos a ver que esta expresión cumple que

$$\lim_{h \to 0} \frac{\|f(h+a) - f(a) - Df(a)(h)\|}{|h|} = 0.$$

Si donde escribíamos h ahora ponemos $x-a$, desarrollando el numerador de la expresión anterior se tendría

$$\begin{aligned}
&f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \\
={}& f(x_1, \ldots, x_n) - f(a_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) \\
&+ f(a_1, x_2, \ldots, x_n) - f(a_1, a_2, x_3, \ldots, x_n) - \frac{\partial f}{\partial x_2}(a)(x_2 - a_2) \\
&+ \cdots + f(a_1, \ldots, a_{n-1}, x_n) - f(a_1, \ldots, a_n) - \frac{\partial f}{\partial x_n}(a)(x_n - a_n).
\end{aligned} \tag{2.9}$$

Bastará demostrar que, para un $\varepsilon > 0$ dado, existe un $\eta > 0$ tal que si se cumplen las siguientes desigualdades

$$|x_1 - a_1| \leq \eta, \ldots, \ |x_n - a_n| \leq \eta, \tag{2.10}$$

entonces, cada fila de la ecuación (2.9) podemos acotarla en norma por $\varepsilon|x_i - a_i|$. Para la desigualdad de cada fila de (2.9) podemos elegir un η adecuado, y si lo tomamos lo suficientemente pequeño, tenemos que se cumplirá para las n desigualdades.

Veamos entonces el caso de la primera desigualdad, ya que sería análogo para el resto de casos. Tomaremos ξ_1, un elemento variable del espacio E_1 lo suficientemente cercano a a_1, y definimos una nueva función g que vendrá dada por

$$g(\xi_1) = f(\xi_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a)(\xi_1 - a_1).$$

Queremos mayorar $\|g(x_1)-g(a_1)\|$. Esta función g será diferenciable por como está definida y es claro que su diferencial es

$$Dg(\xi_1) = \frac{\partial f}{\partial x_1}(\xi_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a_1, \ldots, a_n).$$

Por hipótesis $\frac{\partial f}{\partial x_1}$ es continua, y por la definición de continuidad en un punto a, existe $\eta > 0$ tal que si $|x - a| < \eta$, entones

$$\left\| \frac{\partial f}{\partial x_1}(x_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots, a_n) \right\|_{\mathcal{L}} \leq \varepsilon.$$

Si esto se cumple y definimos $\xi_1 = (1-t)a_1 + tx_1$, como un punto del segmento de origen a_1 y fin x_1, se tiene que

$$\left\| \frac{\partial f}{\partial x_1}(\xi_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots, a_n) \right\|_{\mathcal{L}} \leq \varepsilon,$$

ya que $|\xi_1 - a_1| \leq |x_1 - a_1| \leq \eta$. Aplicando el Teorema del valor medio a la función g se tiene que

$$\|g(x_1) - g(a_1)\| \leq \varepsilon|x_1 - a_1|.$$

Por como está definida g

$$g(x_1) - g(a_1) = f(x_1, x_2, \ldots, x_n) - f(a_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a)(x_1 - a_1),$$

expresión que prueba que cada fila de la ecuación (2.9) se puede acotar superiormente por $\varepsilon|x_i - a_i|$ y, consecuentemente, por $\varepsilon|x - a|$. Queda probado de esta manera que

$$\left\| f(x) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \right\| \leq n\varepsilon|x - a|.$$

Y por lo tanto

$$\lim_{x \to a} \frac{\left\| f(x) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \right\|}{|x - a|} \leq n\varepsilon \quad \text{para todo} \quad \varepsilon \geq 0.$$

Consecuentemente,

$$\lim_{x \to a} \frac{\left\| f(x) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) \right\|}{|x - a|} = 0,$$

es decir, f es diferenciable en a y $Df(a)(h)$ viene dada por (2.8). □

Para finalizar esta sección veremos una aplicación del Teorema del valor medio.

Definición 2.36. Dada una función $f : U \subset E \to F$ se dice que es estrictamente tangente a cero en un punto $a \in U$, si cumple las condiciones siguientes:

1. $f(a) = 0$.

2. Para todo $\varepsilon > 0$ existe $r > 0$ tal que f es ε-lipschitziana en la bola $B[a, r]$.

Partiendo de esta definición introduciremos cuando, dadas dos funciones, estas serán estrictamente tangentes entre ellas en un punto del dominio.

Definición 2.37. Dadas dos funciones $f_1, f_2 : U \subset E \to F$, se dice que son estrictamente tangentes en un punto $a \in U$ si $f_1 - f_2$ es estrictamente tangente a 0 en a.

Podemos ahora definir lo que es una función estrictamente diferenciable.

Definición 2.38. Una función $f : U \subset E \to F$ se dice que es estrictamente diferenciable en un punto $a \in U$ si existe una aplicación lineal y continua, $g : E \to F$, tal que las aplicaciones dadas por

$$x \mapsto f(x) - f(a) \quad \text{y} \quad x \mapsto g(x - a)$$

son estrictamente tangentes en el punto a.

Esta definición nos indica que estas dos aplicaciones son a la fuerza tangentes entre ellas en a. Por lo tanto, si f es estrictamente diferenciable en a entonces f es diferenciable en a y $g = Df(a)$.

Lema 2.39. *Una función $f : U \subset E \to F$ es estrictamente diferenciable en un punto $a \in U$ si y solo si f es diferenciable en a y para todo $\varepsilon > 0$, existe $r > 0$ tal que la aplicación*

$$g(x) = f(x) - f(a) - Df(a)(x - a)$$

es ε-lipschitziana en la bola $B[a, r]$.

Tenemos el siguiente resultado que se sigue del Teorema del valor medio.

Teorema 2.40. *Si $f : U \subset E \to F$ es diferenciable en U y la aplicación $Df : U \to \mathcal{L}(E, F)$ es continua en un punto $a \in U$, entonces f es estrictamente diferenciable en a.*

Demostración. Lo probaremos utilizando el Teorema del valor medio. Si definimos la función

$$g(x) = f(x) - f(a) - Df(a)(x - a)$$

es claro que g es diferenciable y su diferencial viene dada por

$$Dg(x) = Df(x) - Df(a),$$

por lo que $\lim_{x\to a} \|Dg(x)\|_{\mathcal{L}} = 0$ por hipótesis. Entonces para $\varepsilon > 0$ dado, existe $r > 0$ tal que, para $|x - a| \leq r$ se tiene que $\|Dg(x)\|_{\mathcal{L}} \leq \varepsilon$. Por lo tanto, de acuerdo con el Teorema del valor medio

$$\|g(x_1) - g(x_2)\| \leq \varepsilon |x_1 - x_2| \quad \text{para todo } x_1, x_2 \in B[a, r],$$

es decir, g es ε-lipschitziana en $B[a, r]$.

□

Capítulo 3

Diferenciabilidad de orden superior

Con el propósito de obtener mejores resultados a la hora de aproximar una función en un punto nace el concepto de diferenciabilidad de orden superior, el cual introduciremos en este capítulo. Estudiaremos con más profundidad la diferencial segunda y a partir de esta, generalizando, llegaremos a hablar de diferenciales n-ésimas. Además, a lo largo del capítulo veremos también un resultado de gran importancia como es la Fórmula de Taylor.

Nos guiaremos de nuevo por el Capítulo 1 de [4], el Capítulo 4 de [7] y el Capítulo 5 de [1] para explicar todos estos conceptos.

3.1. Diferencial segunda

Hablar de la diferencial segunda es la situación más sencilla para introducir el concepto de diferencial de orden superior y sus propiedades pueden extenderse sin demasiados problemas a casos más complicados. Empezaremos introduciendo la definición de la aplicación diferencial.

Definición 3.1. Si tenemos una función $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable en $\mathbb{R}^n$, entonces está definida la aplicación diferencial

$$\begin{aligned} Df :\ \mathbb{R}^n &\to \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \\ x &\mapsto Df(x). \end{aligned}$$

Recordaremos que una función $f : \mathbb{R}^n \to \mathbb{R}$ es dos veces diferenciable en el punto $a \in \mathbb{R}^n$ si es diferenciable en un entorno del punto a y, a su vez, la aplicación diferencial es también diferenciable en a. En caso de que lo fuese, existirán $\mathcal{H} \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ y una aplicación $\rho : \mathbb{R} \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ que esté definida en un entorno de 0 y además

1. $\lim\limits_{k \to 0} \rho(k) = 0$ en $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

2. $Df(a+k) = Df(a) + \mathcal{H}(k) + \|k\|\rho(k)$.

Por lo tanto, diremos así que la aplicación f es dos veces diferenciable en a si es diferenciable en un entorno A del punto y se tiene que la aplicación

$$x \in A \mapsto Df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$$

es diferenciable en a.

Entonces, volviendo al caso en el que f esté definida entre dos espacios vectoriales normados arbitrarios, E y F, y U sea un subconjunto abierto de E, podríamos definir de manera análoga la aplicación diferencial como

$$\begin{aligned} Df: \ & E \to \mathcal{L}(E, F) \\ & x \mapsto Df(x) \end{aligned}$$

y tendríamos la definición siguiente.

Definición 3.2. Dada una función $f : U \subset E \to F$ diferenciable, diremos que es dos veces diferenciable en $a \in U$ si

$$x \mapsto Df(x) \in \mathcal{L}(E, F)$$

es diferenciable en el punto. Representaremos $D(Df)(a)$ por $D^2f(a)$, que es lo que se conoce como diferencial segunda de f en a.

Observación 4. Diremos que la función f es dos veces diferenciable en U si es dos veces diferenciable en todo $a \in U$ y en este caso la aplicación $x \mapsto D^2f(x)$ es una aplicación

$$D^2f : U \to \mathcal{L}(E, \mathcal{L}(E, F)).$$

Cabe destacar que la aplicación D^2f pertenece al espacio $\mathcal{L}(E, \mathcal{L}(E, F))$, que es isomorfo a $\mathcal{L}(E, F)$, y se trata de una aplicación bilineal continua $E \times E \to F$, que se puede definir como

$$D^2f(a)(h, k) = (D^2f(a)h)k.$$

Esto podemos explicarlo de la siguiente manera: se tiene que h y k designan vectores de E y como $D^2f(a)$ es una aplicación lineal continua $E \to \mathcal{L}(E, F)$, entonces $D^2f(a)h \in \mathcal{L}(E, F)$. Entonces, $D^2f(a)h$ es una aplicación lineal continua $E \to F$, y evaluada en $k \in E$ se denota por

$$(D^2f(a)h)k.$$

Si de nuevo nos volvemos a centrar en funciones más sencillas como serían las de dos variables reales, nos encontramos con dos teoremas a través de los que llegaremos a la conclusión de la simetría de la diferencial segunda. Cabe destacar que para poder hablar de derivadas parciales será necesario que nuestro dominio sea un espacio producto, por lo tanto en ambos resultados trabajaremos en un espacio A de este tipo.

Teorema 3.3 (Schwarz). *Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ una función tal que:*

1. *Las derivadas parciales $\frac{\partial f}{\partial x_1}$ y $\frac{\partial f}{\partial x_2} f$ están bien definidas en $\mathbb{R}^2$.*

2. *La derivada de segundo orden $\frac{\partial f}{\partial x_1 x_2} f$ está bien definida en $\mathbb{R}^2$ y es continua en a.*

Entonces, $\frac{\partial f}{\partial x_2 x_1}(a)$ existe y $\frac{\partial f}{\partial x_1 x_2}(a) = \frac{\partial f}{\partial x_2 x_1}(a)$.

Teorema 3.4 (Young). *Dada una función $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ y $a \in \mathbb{R}^2$. Si $\frac{\partial f}{\partial x_1}$ y $\frac{\partial f}{\partial x_2} f$ existen y son diferenciables en un entorno de a, entonces*

$$\frac{\partial f}{\partial x_1 x_2}(a) = \frac{\partial f}{\partial x_2 x_1}(a).$$

Volviendo de nuevo a espacios normados arbitrarios, tenemos el siguiente resultado, que sigue la línea de los anteriores.

Teorema 3.5. *Sea $f : U \subset E \to F$ dos veces diferenciable en un punto $a \in U$. Entonces la derivada segunda es una aplicación bilineal simétrica, es decir*

$$(D^2 f(a)h)k = (D^2 f(a)k)h \tag{3.1}$$

para todo $(h, k) \in E \times E$.

Demostración. Para llevar a cabo la demostración será necesario definir una nueva función

$$A(h, k) = f(a + h + k) - f(a + h) - f(a + k) + f(a),$$

que evidentemente es simétrica, ya que $A(h, k) = A(k, h)$. Utilizaremos la siguiente igualdad

$$\|A(h, k) - (D^2 f(a)k)h\| = o((|h| + |k|)^2), \tag{3.2}$$

cuya demostración se puede encontrar en el Capítulo 5 de la referencia [1].

Si en la ecuación (3.2) cambiamos h por k, entonces obtendríamos

$$\|A(h, k) - (D^2 f(a)h)k\| = o((|h| + |k|)^2)$$

y de esta manera obtendríamos la relación

$$\|(D^2 f(a)k)h - (D^2 f(a)h)k\| = o((|h| + |k|)^2) \tag{3.3}$$

ya que

$$\|(D^2 f(a)k)h - (D^2 f(a)h)k\| \leq \|(D^2 f(a)k)h - A(h, k)\| + \|A(h, k) - (D^2 f(a)h)k\|.$$

Por (3.3) sabemos que para un $\varepsilon > 0$ dado, entonces existe un $\eta > 0$ tal que si $|h| + |k| \leq \eta$, se cumple que

$$\|(D^2f(a)k)h - (D^2f(a)h)k\| \leq \varepsilon((|h| + |k|)^2).$$

Sabemos además que para todo λ escalar se tiene que

$$\|(D^2f(a)\lambda k)(\lambda h) - (D^2f(a)\lambda h)(\lambda k)\| = |\lambda|^2\|(D^2f(a)k)h - (D^2f(a)h)k\|$$

donde, en este caso, $|\lambda|$ denota el valor absoluto.

Por lo tanto, para cualesquiera que sean h y k vectores de E, podemos encontrar un $\lambda \neq 0$ que cumpla que $|\lambda h| + |\lambda k| \leq \eta$. Entonces, dados h y k arbitrarios se tiene que existe un $\lambda \neq 0$ tal que

$$|\lambda|^2 \cdot \|(D^2f(a)k)h - (D^2f(a)h)k\| \leq \varepsilon|\lambda|^2(|h| + |k|)^2,$$

lo cual implica que

$$\|(D^2f(a)k)h - (D^2f(a)h)k\| \leq \varepsilon(|h| + |k|)^2$$

para todo $(h, k) \in E \times E$ y queda, por tanto, probado el teorema. $\square$

Puede ocurrir que nuestro dominio se pueda escribir como producto de espacios vectoriales normados, $E = E_1 \times \cdots \times E_n$. En estos casos es cuando hablamos de las derivadas parciales y podemos generalizar el Teorema de Schwarz a espacios de Banach. Como se tiene que la forma bilineal tiene que ser simétrica, entonces, para cualesquiera i, j

$$D_{ij}f(a) = D_{ji}f(a).$$

3.2. Diferencial n-ésima

Los resultados vistos para la diferencial de segundo orden pueden ser generalizados para las diferenciales de orden n teniendo en cuenta, únicamente, las posibles dificultades de notación con las que nos podremos encontrar.

Para llegar al concepto general de diferencial n-ésima lo haremos por inducción. Si se tiene que la aplicación $f : U \subset E \to F$ es $n - 1$ veces diferenciable en cada punto de U, definimos entonces una aplicación

$$x \longmapsto D^{n-1}f(x) \in \mathcal{L}_{n-1}(E, F),$$

donde $\mathcal{L}_{n-1}(E, F)$ denota el subespacio de las aplicaciones $(n-1)$-lineales simétricas. Si esta aplicación es diferenciable en el punto $a \in U$, entonces diremos que la aplicación f es n veces diferenciable en a y a la diferencial n-ésima en el punto la denotaremos por $D^nf(a)$.

Definición 3.6. Se dice que la función f es de clase $\mathcal{C}^n$, $n \geq 1$, cuando la aplicación $D^p f(x) \in \mathcal{L}(E, F)$ esté definida para todos los $x \in U$ y sea continua para todo $p = 1, \ldots, n$. Diremos que f es de clase $\mathcal{C}^\infty$ si es de clase $\mathcal{C}^n$ para todos los enteros positivos n.

Ejemplo 3.7. Toda aplicación bilineal continua

$$f : E_1 \times E_2 \to F$$

es de clase $\mathcal{C}^\infty$. Esto se debe a que $D^2 f$ es una aplicación constante y, por lo tanto, las derivadas $D^n f$ son nulas para todo $n > 2$.

Observación 5. Si la función f es continua diremos que es de clase $\mathcal{C}^0$.

Teorema 3.8. *Sea $f : E \times \cdots \times E \to F$ una función n-veces diferenciable en un punto $a \in E$. Entonces la diferencial $D^n f(a) \in \mathcal{L}(E, F)$, es una aplicación multilineal simétrica. Dicho de otra manera, si tenemos σ una permutación de los símbolos $\{1, 2, \ldots, n\}$ y consideramos $h_1, \ldots, h_n$ n vectores de E, entonces*

$$D^n f(a)(h_1, \ldots, h_n) = D^n f(a)(h_{\sigma(1)}, \ldots, h_{\sigma(n)}).$$

Demostración. Lo probaremos para los $n \geq 3$, ya que el caso de $n = 2$ lo habíamos visto en el Teorema 3.5. Suponemos que es cierto para $n - 1$, entonces $D^n f(a)$ es la derivada de

$$D^{n-1} f : V \to \mathcal{L}_{n-1}(E, F),$$

siendo V un entorno de a donde f es $n - 1$ veces diferenciable. Para un vector $h_1 \in E$, $D^n f(a) h_1$ es un elemento del espacio $\mathcal{L}_{n-1}(E, F)$, es decir, que

$$(D^n f(a) h_1) \cdot (h_2, \ldots, h_n)$$

es una función simétrica de $h_2, \ldots, h_n$. Esto es lo mismo que

$$D^n f(a) \cdot (h_1, h_2, \ldots, h_n) \tag{3.4}$$

y vemos que la aplicación multilineal $D^n f(a) : E^n \to F$ es una función simétrica de las $n-1$ últimas variables. Bastará probar que (3.4) no cambia su valor aunque cambiemos h_1 y h_2, y sabemos además que cualquier permutación de n elementos está formada por un número finito de transposiciones que consisten en permutar dos elementos consecutivos. Por lo tanto nada cambiará si los elementos que permutamos son h_i y h_{i+1}, para $i = 2, \ldots, n - 1$. Es claro que $D^n f(a)$ es la derivada segunda de $D^{n-2} f$, y entonces

$$(D^n f(a) h_1) h_2 \in \mathcal{L}_{n-2}(E, F)$$

es simétrica en h_1 y h_2 por el Teorema 3.5. □

3.3. Generalización Fórmula de Taylor

Como ya hemos dicho, el concepto de diferencial de orden n nos ofrece la posibilidad de obtener una aproximación local de la función, que lograremos a través de un polinomio de grado menor o igual a n. Será la fórmula de Taylor la que nos proporcionará una estimación de esta aproximación.

Teorema 3.9. *(Fórmula de Taylor) Dada una función $f : U \subset E \to F$, donde E y F son espacios vectoriales normados y U es un subconjunto abierto, si f es $(n-1)$-veces diferenciable en U y además es n-veces diferenciable en un punto $a \in U$, entonces se tiene que*

$$\left\| f(a+h) - f(a) - Df(a)(h) - \cdots - \frac{1}{n!} D^n f(a)(h)^n \right\| = o(|h|^n) \quad \forall h. \tag{3.5}$$

Observación 6. Puede ocurrir que una función $f : U \subset \mathbb{R} \to \mathbb{R}$ tenga expresión de Taylor de orden n en un punto de su dominio, pero que no sea n-veces diferenciable en ese punto. En este caso la expresión vendría dada por

$$f(a+h) = a_0 + a_1 h + \frac{a_2}{2!} h^2 + \cdots + \frac{a_n}{n!} h^n + o(|h|^n). \tag{3.6}$$

En el caso de ser n-veces diferenciable en un punto y admitir la expresión de Taylor (3.6), entonces esta debe cumplir que

$$a_0 = f(a), \; a_i = D^i f(a), \quad 1 \leq i \leq n.$$

De aquí en adelante trabajaremos con la fórmula de Taylor que expresa el resto con la forma integral, pero para ello es necesario hacer la siguiente anotación. Si trabajamos con una función continua definida como $f : [a,b] \subset \mathbb{R} \to F$, donde el conjunto F es un espacio de Banach, podemos determinar como un vector en F a la siguiente integral

$$y = \int_a^b f(t)dt.$$

Es claro que el valor de esta integral se puede escribir como una aproximación dada por lo que se conoce como sumas de Riemann cuando hacemos tender $n \to \infty$. Este valor quedará bien determinado, ya que las sumas de Riemann forman una sucesión de Cauchy debido a la continuidad de la función f, y como F por ser Banach, es completo, la sucesión es convergente.

Podemos establecer ahora una relación entre la función y la aproximación por la fórmula de Taylor utilizando el resto integral, pero para ello notemos que las hipótesis que debe cumplir la función tienen que ser más fuertes. Lo vemos en el siguiente resultado.

Teorema 3.10 (fórmula de Taylor con resto integral). *Dada una función $f : U \subset E \to F$ de clase $\mathcal{C}^{n+1}$, donde E y F son espacios vectoriales normados y además F es Banach, U es un subconjunto abierto y sea $L[a, a+h] \subset U$. Se tiene entonces que,*

$$\begin{aligned} f(a+h) =& f(a) + Df(a)(h) + \cdots + \frac{1}{n!}D^n f(a)(h)^n \\ &+ \int_0^1 \frac{(1-t)^n}{n!} D^{n+1} f(a+th)(h)^{n+1} dt. \end{aligned} \tag{3.7}$$

Observando la expresión (3.7) vemos que podemos desgranarla y escribirla como la suma de dos términos claramente diferenciados que definiremos a continuación.

Definición 3.11. Denominaremos polinomio de Taylor de grado n de la función el punto $a \in U$ a la expresión dada por

$$P^f_{n,a}(a) = f(a) + Df(a)(h) + \cdots + \frac{1}{n!}D^n f(a)(h)^n.$$

Definición 3.12. Definimos el resto integral del polinomio de Taylor como el error que se comete al realizar la aproximación de la función por el polinomio y viene dado por

$$R^f_{n,a}(a) = \int_0^1 \frac{(1-t)^n}{n!} D^{n+1} f(a+th)(h)^{n+1} dt.$$

Observación 7. Existen más expresiones para representar el resto de la fórmula de Taylor. Uno de ellos será la forma de Lagrange, de la que hablaremos más adelante.

Ya habíamos hablado con anterioridad sobre el Teorema del valor medio y lo habíamos generalizado de manera que su enunciado fuese válido para situaciones arbitrarias en los que los conjuntos con los que trabajamos fuesen espacios vectoriales normados. Ahora, dando un nuevo resultado, veremos la relación que podemos establecer entre él y la fórmula de Taylor.

Teorema 3.13 (fórmula de Taylor con resto de Lagrange). *Sea $f : U \in E \to F$ una función $(n+1)$-veces diferenciable en U y que cumple que*

$$\|D^{n+1} f(x)\| \leq M \textit{ para todo } x \in U.$$

Entonces

$$\left\| f(a+h) - f(a) - Df(a)(h) - \cdots - \frac{1}{n!}D^n f(a)(h)^n \right\| \leq \frac{M|h|^{n+1}}{(n+1)!}. \tag{3.8}$$

Este teorema nos ofrece también una expresión con la que poder aproximar una función a través de un polinomio cuyos coeficientes se corresponden con las diferenciales de orden 1 hasta n de la función que queremos aproximar. En la ecuación (3.8) nos encontramos con la fórmula de Lagrange del resto, que definimos como

$$R^f_{n,a}(a) = \frac{M|h|^{n+1}}{(n+1)!}.$$

Capítulo 4

Los Teoremas de la Función Inversa e Implícita

A lo largo de este capítulo veremos dos grandes teoremas del Análisis Matemático, que son los teoremas de la Función Inversa e Implícita. El primero de estos resultados nos proporcionará las condiciones suficientes para que una función tenga inversa en un punto de su dominio teniendo en cuenta su diferencial en dicho punto. El segundo nace con la necesidad de querer despejar unas variables en función de otras a partir de una ecuación que las relaciona, y nos proporciona las condiciones suficientes bajo las que podemos hacer esto.

Trabajaremos durante todo el capítulo en espacios de Banach, que por su definición sabemos que se tratan de espacios normados. Por tanto podemos utilizar todos los resultados que hemos visto en los capítulos previos.

Para el desarrollo de este capítulo nos guiaremos por el Capítulo 4 de [1] y por el Capítulo 1 de [4].

Cabe destacar que ambos teoremas son equivalentes, es decir, se puede probar cada uno como una consecuencia del otro. En nuestro caso probaremos en primer lugar el de la inversa y a partir de este el de la implícita, pero en otras referencias, como puede ser en [4] o [5], realizan este proceso al revés.

4.1. El Teorema de la Función Inversa

Para poder desarrollar el primero de los teoremas que veremos introduciremos una serie de conceptos previos como son la definición de difeomorfismo y algunos resultados relacionados con estos.

Definición 4.1. Sean E y F dos espacios de Banach y V y W dos abiertos de estos, respectivamente. Se dice que $f : V \to W$ es un difeomorfismo de clase $\mathcal{C}^1$ si es de clase $\mathcal{C}^1$, biyectiva y además su inversa $g = f^{-1} : W \to V$ es también de clase $\mathcal{C}^1$.

Cuando la aplicación inversa no es de clase $\mathcal{C}^1$ en lugar de tener un difeomorfismo diremos que f es un homeomorfismo, como habíamos visto en la Definición 1.17.

Veremos a continuación un ejemplo donde se muestra que no todos los homeomorfismos son difeomorfismos.

Ejemplo 4.2. La función $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$ define un homeomorfismo, ya que es biyectiva y de clase $\mathcal{C}^1$, pero su inversa, definida como

$$g(y) = y^{\frac{1}{3}},$$

no es diferenciable en el origen, y por lo tanto no es de clase $\mathcal{C}^1$. Por lo tanto no será un difeomorfismo, pero como g sí es continua, f es un homeomorfismo.

Teniendo en cuenta este ejemplo, enunciaremos un resultado que nos da una condición general que han de cumplir los homeomorfismos para que pasen a ser difeomorfismos de clase $\mathcal{C}^1$. Primero veremos un lema necesario a la hora de demostrarlo.

Lema 4.3. *Sean E y F dos espacios de Banach, $V \subset E$ y $W \subset F$ dos abiertos y $f : V \to W$ un homeomorfismo diferenciable en un punto $a \in V$. Entonces $g = f^{-1}$ es diferenciable en un punto $b = f(a) \in W$ si y solo si se tiene que $Df(a) \in \mathrm{Isom}(E, F)$ y en tal caso*

$$Dg(b) = (Df(a))^{-1}.$$

Demostración. "$\Rightarrow$" Suponemos que g es diferenciable en un punto $b = f(a) \in W$. Además, por hipótesis f es diferenciable en a. Vemos que se cumplen las hipótesis de la Proposición 2.18 y por lo tanto se tiene que

$$Dg(f(a)) \circ Df(a) = I_E \;\text{ y }\; Df(a) \circ Dg(f(a)) = I_F.$$

Entonces, tanto $Df(a)$ como $Dg(f(a))$ son dos aplicaciones biyectivas y por lo tanto isomorfismos, la primera de E en F y la segunda de F en E.

"$\Leftarrow$" Para llevar a cabo esta implicación suponemos que $Df(a) \in \mathrm{Isom}(E, F)$. Tomamos $y = f(x)$ para un x cercano a a y, como f es diferenciable en a, se tiene que, existe una función ϕ que cumple que $\lim_{x \to a} \phi(x - a) = 0$ y

$$y - b = Df(a)(x - a) + |x - a| \cdot \phi(x - a).$$

Efectuando en ambos miembros la transformación lineal $(Df(a))^{-1}$ llegamos a la expresión siguiente

$$x - a = (Df(a))^{-1}(y - b) - (Df(a))^{-1}\Big(|x - a|\phi(x - a)\Big). \tag{4.1}$$

Si, para aligerar notación, escribimos

$$(Df(a))^{-1}\Big(\phi(x - a)\Big) = \psi(x - a)$$

como $(Df(a))^{-1}$ es una aplicación lineal y continua de F en E, $\psi(x-a)$ tiende a 0 cuando $x \to a$.

Entonces, reescribiendo (4.1) se tiene que

$$g(y) - g(b) - (Df(a))^{-1}(y - b) = |x - a|\psi(x - a)$$

por lo que si probamos que

$$\lim_{y \to b} \frac{|x - a|\psi(x - a)}{\|y - b\|} = 0$$

habremos probado que g es diferenciable en b y que $Dg(b) = (Df(a))^{-1}$.

Si partimos de (4.1) llegamos a que

$$\|(Df(a))^{-1}(y - b)\| \geq |x - a| \left(1 - \|\psi(x - a)\|\right),$$

de donde, como habíamos tomado un x lo suficientemente próximo a a, podemos garantizar que $1 - \|\psi(x - a)\| > 0$ y obtener

$$|x - a| \leq \|y - b\| \cdot \frac{\|(Df(a))^{-1}\|_{\mathcal{L}}}{1 - \|\psi(x - a)\|},$$

y multiplicando en ambos lados por $\|\psi(x - a)\|$ se tiene que

$$|x - a| \cdot \|\psi(x - a)\| \leq \|y - b\| \cdot \|(Df(a))^{-1}\|_{\mathcal{L}} \cdot \frac{\|\psi(x - a)\|}{1 - \|\psi(x - a)\|}.$$

Consecuentemente

$$\lim_{y \to b} \frac{|x - a|\psi(x - a)}{\|y - b\|} = \lim_{y \to b} \|(Df(a))^{-1}\|_{\mathcal{L}} \cdot \frac{\|\psi(x - a)\|}{1 - \|\psi(x - a)\|} = 0$$

ya que si $y \to b$, entonces $x = g(y)$ tiende a $a = g(b)$.

□

Enunciamos y probamos ahora la proposición.

Proposición 4.4. *Dados E y F dos espacios de Banach, $V \subset E$ y $W \subset F$ dos abiertos y $f : V \to W$ un homeomorfismo de clase $\mathcal{C}^1$. Para que f sea un difeomorfismo de clase $\mathcal{C}^1$, será necesario que para todo $x \in V$, $Df(x)$ pertenezca a* $\mathrm{Isom}(E, F)$.

Demostración. Por el Lema 4.3 sabemos que si $Df(x) \in \text{Isom}(E, F)$ para todo $x \in V$ la función g es diferenciable en todo punto $y \in W$ y que además

$$Dg(y) = (Df(g(y)))^{-1}.$$

Para probar que g es de clase $\mathcal{C}^1$ bastará ver que

$$Dg : W \to \mathcal{L}(F, E)$$

es continua. Por como está definida Dg sabemos que se trata de la composición de tres aplicaciones y por lo tanto estudiaremos cada una por separado.

Empezaremos viendo la función $g : W \to V$, que es continua ya que f es un homeomorfismo. Por otro lado, la aplicación Df es también continua ya que supusimos que f es de clase $\mathcal{C}^1$. En último lugar sabemos que la aplicación $u \mapsto u^{-1}$ que va de $\text{Isom}(E, F)$ a $\mathcal{L}(E, F)$ es continua, esto se puede ver en el Teorema 1.7.3 del Capítulo 1 de [1]. Por lo tanto, como la composición de aplicaciones continuas es continua, queda probado entonces que Dg es una aplicación continua para todo $y \in W$. □

Con estos dos últimos resultados queda introducida la primera parte de difeomorfismos de clase $\mathcal{C}^1$. Introduciendo una última proposición, estaremos en condiciones de enunciar el Teorema de la Función Inversa.

Proposición 4.5. *Sean dos espacios de Banach E y F, un abierto $U \subset E$ y una función diferenciable $f : U \to F$. Si Df es continua en $a \in U$ y $Df(a) \in \text{Isom}(E, F)$, entonces existen dos entornos abiertos, $V' \subset U$ de a y W' de $b = f(a)$, tales que f es un homeomorfismo de V' en W'.*

Demostración. Tomamos la aplicación lineal y continua $(Df(a))^{-1}$ que va de F en E y consideramos la aplicación composición definida como

$$f_1 : (Df(a))^{-1} \circ f : U \to E.$$

Probaremos que esta aplicación es estrictamente diferenciable en a. Partimos de que f es diferenciable en U. Además, como $(Df(a))^{-1}$ es lineal, por la Proposición 2.21 será diferenciable en todo punto y en particular lo será en $f(a)$. Por lo tanto, como la composición de dos aplicaciones diferenciables es diferenciable, aplicando la Regla de la cadena, f_1 será diferenciable en a. Usando la Proposición 2.21 y la Regla de la cadena, calcularemos la diferencial de f_1

$$Df_1(c) = D\left((Df(a))^{-1}\right)(f(c)) \circ Df(c) = (Df(a))^{-1} \circ Df(c).$$

Se tiene entonces que $Df_1 = (Df(a))^{-1} \circ Df$ y, en particular, $Df_1(a) = I_E$.

Para poder aplicar el Teorema 2.40 faltaría ver que Df_1 es continua en a. Esto es inmediato ya que por hipótesis Df es continua en a y $(Df(a))^{-1}$ también lo es. Por lo tanto Df_1 es continua en a. Queda probado que f_1 es estrictamente diferenciable en a.

Como f_1 es estrictamente diferenciable, a cada $k > 0$ le podemos asociar un $r > 0$ tal que la aplicación

$$\begin{aligned} \varphi : U &\longrightarrow E \\ x &\longmapsto \varphi(x) = x - f_1(x) \end{aligned}$$

sea k-lipschitziana en $B[a,r]$. Si elegimos un $k \in (0,1)$, la aplicación φ será k-contractiva en la bola $B[a,r]$ y por lo tanto, el Teorema 1.19, nos permite afirmar que existe un entorno abierto de a, $V' \subset B[a,r]$, tal que f_1 es un homeomorfismo de V' en un entorno W_1 de $b_1 = f_1(a)$. Como también tenemos que $Df(a)$ es un homeomorfismo de E en F, llegamos a que

$$f = Df(a) \circ f_1$$

es un homeomorfismo de V' en W', con $W' \subset F$ entorno abierto de $f(a) = b$. $\square$

Teniendo en cuenta todos estos preliminares, podemos ahora enunciar y demostrar el Teorema de la Función Inversa.

Teorema 4.6 (de la Función Inversa)**.** *Sean E, F espacios de Banach, $U \subset E$ un subconjunto abierto y $f : U \to F$ una aplicación de clase $\mathcal{C}^1$. Supongamos además que $Df(a) \in \mathrm{Isom}(E,F)$ para un punto $a \in U$. Entonces, existen entornos abiertos $V \subset U$ de a y W de $f(a) = b$, tales que f es un difeomorfismo de clase $\mathcal{C}^1$ de V en W.*

Demostración. Sabemos que, bajo nuestras hipótesis, $Df(x)$ existe para todo $x \in V'$. Además, sabemos que existe un entorno $V \subset V'$ de a tal que $Df(x) \in \mathrm{Isom}(E,F)$, ya que por ser $\mathrm{Isom}(E,F)$ un abierto de $\mathcal{L}(E,F)$, Teorema 1.7.3 del Capítulo 1 de [1], la imagen inversa de la aplicación continua Df es un abierto de V' que contiene a a. Como f es un homeomorfismo de V' en W', por la Proposición 4.5 tenemos un abierto W de W' tal que $W = f(V)$ y además f es un homeomorfismo de V en W. Teniendo esto en cuenta, nos encontramos en las hipótesis de la Proposición 4.4, lo que nos permite concluir que f es un difeomorfismo de clase $\mathcal{C}^1$ de V en W. $\square$

Un resultado que obtenemos a partir de este teorema sería el siguiente.

Corolario 4.7. *Para que una función $f : U \to F$, de clase $\mathcal{C}^1$, sea un difeomorfismo de clase $\mathcal{C}^1$ de U sobre un abierto de F, será necesario y suficiente que:*

1. *f sea una función inyectiva.*

2. *$Df(x) \in \mathrm{Isom}(E, F)$ para todo $x \in U$.*

4.2. El Teorema de la Función Implícita

El resultado que veremos en esta sección está relacionado con el conjunto de soluciones de ecuaciones de la forma

$$f(x, y) = 0, \tag{4.2}$$

donde $f : E \times F \to G$ es una función continua entre espacios vectoriales normados. Es complicado tener un resultado general y preciso sobre la estructura global del conjunto de soluciones de las funciones que cumplen (4.2). A pesar de esto, si tenemos una solución (a, b) de la ecuación, bajo algunas condiciones podemos escribir el conjunto local de soluciones. Tenemos el siguiente resultado donde se recogen dichas condiciones.

Teorema 4.8 (de la Función Implícita)**.** *Sean E, F y G tres espacios de Banach, un abierto $U \subset E \times F$ y $f : U \to G$ una aplicación tal que:*

1. *f es continua.*

2. *Para todo punto $(x_1, x_2) \in U$, $\frac{\partial f}{\partial x_2}(x_1, x_2)$ existe y es continua en U.*

3. *Existe $(a, b) \in U$ tal que $f(a, b) = 0$, y $A = \frac{\partial f}{\partial x_2}(a, b)$ es inversible y su inversa es continua.*

Entonces, existen un entorno abierto $V \subset U$ de (a, b), un entorno abierto $W \subset E$ de a y una aplicación de clase $\mathcal{C}^1$

$$\varphi : W \to F$$

tal que

$$(x, y) \in V \ \text{ y } f(x, y) = 0 \tag{4.3}$$

es equivalente a

$$x \in W \ \text{ e } y = \varphi(x). \tag{4.4}$$

Demostración. Para la demostración necesitaremos el Teorema de la Función Inversa. Consideremos la aplicación

$$\begin{aligned} g : U &\longrightarrow E \times G \\ (x_1, x_2) &\longmapsto g(x_1, x_2) = (x_1, f(x_1, x_2)), \end{aligned} \tag{4.5}$$

donde $x_1 \in E$ y $x_2 \in F$. Es claro que g es una función de clase $\mathcal{C}^1$ en U, ya que sus dos componentes lo son. Su aplicación diferencial viene dada por la siguiente matriz

$$Jg(a,b) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

donde se tiene que $\alpha \in \mathcal{L}(E,E)$, $\beta \in \mathcal{L}(F,E)$, $\gamma \in \mathcal{L}(E,G)$ y $\delta \in \mathcal{L}(F,G)$. De hecho, si calculamos las derivadas parciales de g, llegamos a que

$$\begin{cases} \alpha = I_E, \quad \beta = 0 \\ \gamma = \frac{\partial f}{\partial x_1}(a,b), \quad \delta = \frac{\partial f}{\partial x_2}(a,b). \end{cases}$$

De esta manera, llegamos a la siguiente aplicación lineal

$$\begin{aligned} Dg(a,b) : E \times F &\longrightarrow E \times G \\ (h,k) &\longmapsto \left(h, \frac{\partial f}{\partial x_1}(a,b)(h) + \frac{\partial f}{\partial x_2}(a,b)(k) \right). \end{aligned} \tag{4.6}$$

Como $\frac{\partial f}{\partial x_2}(a,b) \in \mathrm{Isom}(E,F)$, entonces (4.6) es un isomorfismo de $E \times F$ en $E \times G$, donde el isomorfismo inverso viene dado por

$$(h',k') \longmapsto \left(h', \left(\frac{\partial f}{\partial x_2}(a,b) \right)^{-1} (k') - \left(\frac{\partial f}{\partial x_1}(a,b) \right)^{-1} (h') \right).$$

Podemos aplicar a g, en un entorno del punto $(a,b) \in U$, el Teorema 4.6. Por lo tanto sabemos que en $E \times F$ existe un entorno abierto $V \subset U$ de (a,b) y que en $E \times G$ existe otro entorno abierto W_1 de $(a,0) = g(a,b)$, tales que g es un difeomorfismo de clase $\mathcal{C}^1$ de V en W_1.

Denotemos por g_1 al difeomorfismo inverso, el cual es de la forma

$$g_1(x,z) = (x, g(x,z)), \text{ para } x \in E, \ z \in G$$

con $(x,z) \in W_1$, y

$$\widetilde{\varphi} : W_1 \to F$$

una función de clase $\mathcal{C}^1$. Como g y g_1 son aplicaciones inversas entre si, existirá una equivalencia entre estas dos condiciones:

1. $(x,y) \in V$ y $f(x,y) = z$.

2. $(x,z) \in W_1$ y $\widetilde{\varphi}(x,z) = y$.

Si en las relaciones anteriores hacemos $z = 0$, la primera se convierte en (4.3) y veamos que pasa con 2. Si identificamos E con un subespacio vectorial de $E \times F$, entonces podemos escribir un punto $x \in E$ como $(x, 0) \in E \times F$ y como $(x, 0) \in W_1$, podemos afirmar que x pertenece a la intersección de E y W_1, que es un abierto $W \subset E$. Por otro lado

$$\widetilde{\varphi}(x, 0) = \varphi(x)$$

es una función de clase $\mathcal{C}^1$ en W, por lo tanto la expresión 2, para $z = 0$ queda de la forma

$$x \in W \text{ y } y = \varphi(x).$$

Esta expresión es exactamente (4.4), lo que prueba que la equivalencia con (4.3) se cumple. □

Observación 8. Además, ya que por hipótesis, para $(a, b) \in V$ se tiene que $f(a, b) = 0$, las equivalencias anteriores nos muestran que $\varphi(a) = b$.

Veremos un ejemplo que nos ayuda a clarificar el resultado anterior.

Ejemplo 4.9. Tomemos una función $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida como

$$f(x, y, z) = (x + y + z, x + 2y + 2z).$$

Por como está definida f sabemos que es de clase $\mathcal{C}^\infty$.

Veamos si podemos resolver el siguiente sistema

$$\begin{cases} x + y + z &= 0, \\ x + 2y + 2z &= 0, \end{cases}$$

tomando x como parámetro.

Teniendo en cuenta la notación que usábamos al enunciar y demostrar el Teorema de la Función Implícita, se tiene que $E_1 = \mathbb{R}$, $E_2 = \mathbb{R}^2$ y $F = \mathbb{R}^2$. La derivada parcial de f respecto de (y, z) en el punto (x, y, z) viene dada por la matriz

$$Jf(x, y, z) = \begin{pmatrix} D_1 f_1(x, y, z) & D_2 f_1(x, y, z) \\ D_1 f_2(x, y, z) & D_2 f_2(x, y, z) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

que no es inversible, ya que su determinante es 0. Por lo tanto no se puede despejar (y, z) en función de x. Sin embargo sí que podemos despejar (x, z) en función de y, pues

$$Jf(x, y, z) = \begin{pmatrix} D_1 f_1(x, y, z) & D_2 f_1(x, y, z) \\ D_1 f_2(x, y, z) & D_2 f_2(x, y, z) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

y $Jf(x, y, z)$ tiene determinante distinto de 0.

Bibliografía

[1] H. Cartan, *Calcul Différentiel*, Hermann, 1971.

[2] J. B. Conway, *A Course in Functional Analysis*, Springer-Verlag New York, 1990.

[3] J. Dieudonné, *Fundamentos de Análisis Moderno*, Reverté, 1966.

[4] S. Kesavan, *Nonlinear Functional Analysis: A First Course*, Hindustan Book Agency, 2004.

[5] S. Krantz y H. Parks, *The Implicit Function Theorem. History, Theory, and Applications*, Springer New York Heidelberg Dordrecht London, 2013.

[6] L. Nachbin, *Introduction to Functional Analysis: Banach Spaces and Differential Calculus*, Marcel Dekker Inc, 1981.

[7] G. Rodríguez, *Diferenciación de funciones de varias variables reales*, Universidade de Santiago de Compostela, 2003.