Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Civil Matemática
Universidad de Chile

Apuntes del curso

Cálculo en Varias Variables (MA22A)

Profesores: Rafael Correa - Pedro Gajardo

Auxiliares: Rodolfo Gainza - Gonzalo Sánchez

2005

Índice general

CAPÍTULO 1

ESPACIOS VECTORIALES NORMADOS

1.1 Introducción

La estructura de espacio vectorial es la estructura algebraica de mayor importancia del análisis matemático. Al definir en un espacio vectorial la noción de norma, estamos introduciendo una noción fundamental, que es la de vecindad de un punto del espacio, o más intuitivamente la de cercania entre dos puntos.

Dada una norma en un espacio vectorial, definiremos en este capítulo las herramientas y propiedades básicas que nos permitirán más adelante ir construyendo las herramientas y propiedades más complejas que intervienen en los modelos matemáticos de la ingeniería, de la física y de otras ciencias.

1.2 Conceptos preliminares

Definición 1.2.1. Un espacio vectorial normado (e.v.n.) es un e.v. $\vec{E}$ sobre el cuerpo $\mathbb{R}$ (de los reales) ó $\mathbb{C}$ (de los complejos), dotado de una aplicación de $\vec{E}$ en $\mathbb{R}_+$ (conjunto de los reales ≥ 0), llamada norma y que denotamos $\|\cdot\|$, con las siguientes tres propiedades: para todo $\vec{a}, \vec{b} \in \vec{E}$ y para todo λ en el cuerpo, se tiene

$$\|\vec{a}\| = 0 \Leftrightarrow \vec{a} = \vec{0} \tag{1.2.1}$$

$$\|\lambda\vec{a}\| = |\lambda| \, \|\vec{a}\| \tag{1.2.2}$$

$$\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\| \tag{1.2.3}$$

Hablaremos entonces del e.v.n. $(\vec{E}, \|\cdot\|)$ o simplemente (si no hay confusión posible) del e.v.n. $\vec{E}$.

Nota 1.2.1. En adelante, cada vez que nos demos un e.v., supondremos que está constituído por más de un elemento, es decir, que es diferente de $\{\vec{0}\}$. Si no decimos lo contrario, supondremos también que el cuerpo es $\mathbb{R}$.

Ejemplo 1.2.1. El e.v. $\mathbb{R}^n$ dotado de alguna de las normas:

$$\|\vec{a}\|_2 := \left[\sum_{i=1}^{n} a_i^2\right]^{\frac{1}{2}} \tag{1.2.4}$$

$$\|\vec{a}\|_1 := \sum_{i=1}^{n} |a_i| \tag{1.2.5}$$

$$\|\vec{a}\|_p := \left[\sum_{i=1}^{n} |a|_i^p\right]^{\frac{1}{p}} \quad 1 \leq p < \infty \tag{1.2.6}$$

$$\|\vec{a}\|_\infty := \max_{i=1,\ldots,n} |a_i| \tag{1.2.7}$$

donde las cantidades $a_i \in \mathbb{R}$ son las componentes de $\vec{a} \in \mathbb{R}^n$, es un e.v.n.

Ejemplo 1.2.2. El e.v. $\mathcal{L}(\vec{E}, \vec{F})$ de las aplicaciones lineales de un e.v.n. $(\vec{E}, \|\cdot\|_{\vec{E}})$ a valores en un e.v.n. $(\vec{F}, \|\cdot\|_{\vec{F}})$, que cumplen la propiedad

$$l \in \mathcal{L}(\vec{E}, \vec{F}) \Leftrightarrow \exists M \geq 0 \text{ tal que } \|l(\vec{x})\|_{\vec{F}} \leq M\|\vec{x}\|_{\vec{E}} \quad \forall\, \vec{x} \in \vec{E}$$

dotado de la norma

$$\|l\| := \sup_{\vec{x} \in \vec{E} \setminus \{0\}} \frac{\|l(\vec{x})\|_{\vec{F}}}{\|\vec{x}\|_{\vec{E}}} \tag{1.2.8}$$

es un e.v.n. Más adelante veremos que $\mathcal{L}(\vec{E}, \vec{F})$ es el e.v. de las funciones lineales continuas de $\vec{E}$ en $\vec{F}$.

Ejemplo 1.2.3. El e.v. $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ dotado de alguna de las normas:

$$\|l\|_F := \left[\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right]^{1/2} \tag{1.2.9}$$

$$\|l\|_1 := \max_j \sum_{i=1}^{n} |a_{ij}| \tag{1.2.10}$$

$$\|l\|_\infty := \max_i \sum_{j=1}^{m} |a_{ij}| \tag{1.2.11}$$

$$\|l\|_{\max} := \max_{i,\ j} |a_{ij}| \tag{1.2.12}$$

donde las cantidades a_{ij} $(1 \leq i \leq m,\ 1 \leq j \leq n)$ son los elementos de la matriz que representa a la aplicación lineal $l \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, es un e.v.n. La norma en (1.2.9) es conocida como la norma de Frobenius.

Ejemplo 1.2.4. El e.v. $\mathcal{A}(A, \vec{F})$ de todas las funciones acotadas definidas en un conjunto A con valores en un e.v.n. $(\vec{F}, \|\cdot\|_{\vec{F}})$ ($f : A \to \vec{F}$ se dice acotada si existe $r \in \mathbb{R}_+$ tal que $\|f(x)\|_{\vec{F}} \leq r$ para todo $x \in A$), dotado de la norma:

$$\|f\|_\infty := \sup_{x \in A} \|f(x)\|_{\vec{F}} \tag{1.2.13}$$

es un e.v.n.

Ejemplo 1.2.5. Sean $(\vec{E}_1, \|\cdot\|_{\vec{E}_1}), ..., (\vec{E}_n, \|\cdot\|_{\vec{E}_n})$, n e.v.n. El e.v. $\vec{E}_1 \times ... \times \vec{E}_n$ dotado de alguna de las normas

$$\|\vec{a}\|_2 := \left[\sum_{i=1}^{n} \|\vec{a}_i\|_{\vec{E}_i}^2\right]^{1/2}$$

$$\|\vec{a}\|_1 := \sum_{i=1}^{n} \|\vec{a}_i\|_{\vec{E}_i}$$

$$\|\vec{a}\|_\infty := \max_{i=1,\dots,n} \|\vec{a}_i\|_{\vec{E}_i}$$

donde los elementos $\vec{a}_i \in \vec{E}_i$ son las componentes de $\vec{a} \in \vec{E}_1 \times ... \times \vec{E}_n$, es un e.v.n.

Definición 1.2.2. Dados dos elementos $\vec{a}, \vec{b}$ en un e.v.n. $\vec{E}$, se llama distancia de $\vec{a}$ a $\vec{b}$ a la cantidad $\|\vec{a} - \vec{b}\|$. De este modo, la cantidad $\|\vec{a}\|$ corresponde a la distancia de $\vec{a}$ al origen $\vec{0}$.

Definición 1.2.3. Dado un elemento $\vec{c}$ en un e.v.n. $\vec{E}$ y un real $r > 0$, se llama bola cerrada (respectivamente abierta) de centro $\vec{c}$ y radio r al conjunto

$$\begin{array}{rcl} B(\vec{c}, r) & := & \{\vec{x} \in \vec{E} \ : \ \|\vec{c} - \vec{x}\| \leq r\} \\ (\text{ resp. } \ B'(\vec{c}, r) & := & \{\vec{x} \in \vec{E} \ : \ \|\vec{c} - \vec{x}\| < r\}) \end{array}$$

Definición 1.2.4. Una parte A de un e.v.n. $\vec{E}$ se dirá acotada si existe $r > 0$ tal que $A \subset B(\vec{0}, r)$, es decir, tal que $\|\vec{x}\| \leq r$ para todo $\vec{x} \in A$.

Definición 1.2.5. Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$, definidas en un e.v. $\vec{E}$, se dirán equivalentes si existen dos constantes L_1 y L_2 tales que

$$\|\cdot\|_2 \leq L_1 \|\cdot\|_1 \ \text{ y } \ \|\cdot\|_1 \leq L_2 \|\cdot\|_2 \tag{1.2.14}$$

Nota 1.2.2. Se demuestra que la primera de las desigualdades en (1.2.14) es equivalente a la propiedad

$$\textit{para toda bola } \ B_2(\vec{0}, \varepsilon) \textit{existe una bola } \ B_1(\vec{0}, \delta) \subset B_2(\vec{0}, \varepsilon) \tag{1.2.15}$$

y la segunda es equivalente a la propiedad

$$\textit{para toda bola } \ B_1(\vec{0}, \varepsilon) \textit{existe una bola } \ B_2(\vec{0}, \delta) \subset B_1(\vec{0}, \varepsilon) \tag{1.2.16}$$

donde los subíndices indican la norma que interviene en la definición de la respectiva bola. Es fácil verificar que las propiedades (1.2.15) y (1.2.16) son las mismas si se cambia $\vec{0}$ por cualquier otro elemento $\vec{a} \in \vec{E}$.

Nota 1.2.3. Se demuestra que todas las normas que se pueden definir en un e.v. de dimensión finita son equivalentes. En particular, se demuestra fácilmente que, las normas definidas en los ejemplos 1.2.1 y 1.2.3 verifican, para todo $\vec{a} \in \mathbb{R}^n$:

$$\begin{array}{rcl} \|\vec{a}\|_\infty & \leq & \|\vec{a}\|_1 \leq n\|\vec{a}\|_\infty \\ \|\vec{a}\|_\infty & \leq & \|\vec{a}\|_2 \leq \sqrt{n}\|\vec{a}\|_\infty \\ \|\vec{a}\|_2 & \leq & \|\vec{a}\|_1 \leq \sqrt{n}\|\vec{a}\|_2 \end{array}$$

y, para todo $l \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$:

$$\begin{aligned}\frac{1}{\sqrt{n}}\|l\|_\infty &\leq \|l\|_F \leq \sqrt{nm}\|l\|_\infty \\ \frac{1}{\sqrt{m}}\|l\|_1 &\leq \|l\|_F \leq n\|l\|_1 \\ \frac{1}{n}\|l\|_\infty &\leq \|l\|_1 \leq m\|l\|_\infty\end{aligned}$$

En lo que sigue cuando hablemos del e.v.n. $\mathbb{R}^n$, sin especificar la norma, entenderemos que se trata de cualquiera de las cuatro normas del Ejemplo 1.2.1. Más adelante veremos que en un e.v. (de dimensión infinita) se pueden definir normas que no son equivalentes.

Nota 1.2.4. Se demuestra fácilmente que las tres normas definidas en el Ejemplo 1.2.5 son equivalentes (sin importar la dimensión de los espacios $\vec{E}_i$). En lo que sigue, cuando hablemos de un e.v.n. producto sin especificar la norma, entenderemos que se trata de alguna de estas tres.

1.3 Conjuntos abiertos y cerrados

Definición 1.3.1. Una parte A de un e.v.n $\vec{E}$ se dirá abierta si para todo $\vec{a} \in A$ existe una bola $B(\vec{a}, \delta) \subset A$. Una parte A de un e.v.n $\vec{E}$ se dirá cerrada si su complemento A^c es abierto.

Nota 1.3.1. Dadas dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ en un e.v. $\vec{E}$, los conjuntos abiertos en $\vec{E}$ no serán necesariamente los mismos si dotamos a $\vec{E}$ de la primera o de la segunda de estas normas, dicho de otro modo, los conjuntos abiertos del e.v.n. $(\vec{E}, \|\cdot\|_1)$ no serán necesariamente los mismos que los del e.v.n. $(\vec{E}, \|\cdot\|_2)$. Pero, si se tiene la propiedad (1.2.15) (equivalente a $\|\cdot\|_2 \leq L_1\|\cdot\|_1$) vemos que todo conjunto abierto en $\vec{E}$ con la norma $\|\cdot\|_2$, seguirá siendo abierto en $\vec{E}$ con la norma $\|\cdot\|_1$. Análogamente, si se tiene la propiedad (1.2.16) (equivalente a $\|\cdot\|_1 \leq L_2\|\cdot\|_2$) vemos que todo conjunto abierto en $\vec{E}$ con la norma $\|\cdot\|_1$ será también abierto en $\vec{E}$ con la norma $\|\cdot\|_2$. En consecuencia, si $\|\cdot\|_1$ y $\|\cdot\|_2$ son equivalentes, entonces los conjuntos abiertos son los mismos si dotamos a $\vec{E}$ de la primera o de la segunda de estas normas.

Lo mismo que hemos dicho para los conjuntos abiertos, vale para los conjuntos cerrados. De lo anterior y de lo que decíamos en la Nota 1.2.3, podemos concluir que en un e.v. de dimensión finita, cualquiera que sea la norma que definamos, los conjuntos abiertos y los conjuntos cerrados serán los mismos.

Ejemplo 1.3.1. Demostremos que en un e.v.n. $\vec{E}$ toda bola abierta $B'(\vec{c}, r)$ es un conjunto abierto. Dado $\vec{a} \in B'(\vec{c}, r)$ se tiene que $\|\vec{c} - \vec{a}\| < r$. Escojamos entonces un real $\delta > 0$ tal que $\delta < r - \|\vec{c} - \vec{a}\|$ y demostremos $B(\vec{a}, \delta) \subset B'(\vec{c}, r)$:

$$\begin{aligned} \vec{x} \in B(\vec{a}, \delta) &\Rightarrow \|\vec{a} - \vec{x}\| \leq \delta \\ &\Rightarrow \|\vec{a} - \vec{x}\| < r - \|\vec{c} - \vec{a}\| \\ &\Rightarrow \|\vec{a} - \vec{x}\| + \|\vec{c} - \vec{a}\| < r \\ &\Rightarrow \|\vec{c} - \vec{x}\| < r \\ &\Rightarrow \vec{x} \in B'(\vec{c}, r) \end{aligned}$$

con lo que queda demostrado que $B'(\vec{c}, r)$ es un conjunto abierto.

Ejemplo 1.3.2. Demostremos que en un e.v.n. $\vec{E}$ toda bola cerrada $B(\vec{c}, r)$ es un conjunto cerrado. Para esto hay que demostrar que el conjunto $B(\vec{c}, r)^c = \{\vec{x} \in \vec{E} : \|\vec{c} - \vec{x}\| > r\}$ es abierto. Dado $\vec{a} \in B(\vec{c}, r)^c$ se tiene que $\|\vec{a} - \vec{c}\| - r > 0$. Escojamos entonces un real $\delta > 0$ tal que $\delta < \|\vec{c} - \vec{a}\| - r$. Se demuestra fácilmente que $B(\vec{a}, \delta) \subset B(\vec{c}, r)^c$, con lo que queda demostrado que $B(\vec{c}, r)^c$ es un conjunto abierto.

Teorema 1.3.1. *Si denotamos por $\mathcal{O}$ la familia de todos los subconjuntos abiertos de un e.v.n. $\vec{E}$, se tendrán las tres propiedades fundamentales siguientes:*

i) *Si $\{A_i\}_{i=1}^n$ es una familia finita de elementos de $\mathcal{O}$, entonces*

$$\bigcap_{i=1}^{n} A_i \in \mathcal{O}$$

ii) *Si $\{A_t\}_{t \in T}$ es una familia cualquiera de elementos de $\mathcal{O}$, entonces*

$$\bigcup_{t \in T} A_t \in \mathcal{O}$$

iii) *$\vec{E} \in \mathcal{O}$ y $\emptyset \in \mathcal{O}$.*

Demostración. i) Debemos demostrar que el conjunto $A := \bigcap_{i=1}^{n} A_i$ es abierto. Si $A = \emptyset$ remitimos la demostración a la parte iii). Si $A \neq \emptyset$, dado $\vec{a} \in A$, se tendrá que $\vec{a} \in A_i$ para todo $i = 1, ..., n$ y, como los A_i son abiertos, existirán $\delta_1 > 0, \delta_2 > 0, ..., \delta_n > 0$ tales que $B(\vec{a}, \delta_i) \subset A_i$ para todo $i = 1, ..., n$. Definamos ahora $\delta := \text{mín}\{\delta_i,\ i = 1, ..., n\} > 0$.

Se tiene entonces que $B(\vec{a},\delta) \subset B(\vec{a},\delta_i) \subset A_i$ para todo $i = 1, ..., n$, lo cual implica que $B(\vec{a},\delta) \subset A$. Con esto hemos demostrado que A es abierto.

ii) Debemos demostrar que el conjunto $A := \bigcup\limits_{t \in T} A_t$ es abierto. Sea $\vec{a} \in A$, entonces $\vec{a} \in A_{\bar{t}}$ para algún $\bar{t} \in T$ y, como $A_{\bar{t}}$ es abierto, existirá $\delta > 0$ tal que $B(\vec{a},\delta) \subset A_{\bar{t}}$. Esto implica que $B(\vec{a},\delta) \subset A$, con lo que hemos demostrado que A es abierto.

iii) Demostrar que $\vec{E}$ es abierto es trivial. Para convencerse que $\emptyset$ debe ser un conjunto abierto, basta con decir que al no tener elementos es imposible probar que no es abierto.

□

Teorema 1.3.2. *Si denotamos por $\mathcal{C}$ el conjunto de todas las partes cerradas de un e.v.n. $\vec{E}$, se tendrán las tres propiedades fundamentales siguientes:*

i) *Si $\{C_i\}_{i=1}^n$ es una familia finita de elementos de $\mathcal{C}$, entonces*

$$\bigcup_{i=1}^{n} C_i \in \mathcal{C}$$

ii) *Si $\{C_t\}_{t \in T}$ es una familia cualquiera de elementos de $\mathcal{C}$, entonces*

$$\bigcap_{t \in T} C_t \in \mathcal{C}$$

iii) *$\emptyset \in \mathcal{C}$ y $\vec{E} \in \mathcal{C}$.*

Demostración. i) Esta propiedad es consecuencia inmediata de la fórmula $[\bigcup\limits_{i=1}^{n} C_i]^c = \bigcap\limits_{i=1}^{n} C_i^c$ y de la parte i) del teorema anterior.

ii) Esta propiedad es consecuencia inmediata de la fórmula $[\bigcap\limits_{t \in T} C_t]^c = \bigcup\limits_{t \in T} C_t^c$ y de la parte ii) del teorema anterior.

iii) Esta propiedad es consecuencia inmediata del hecho que $\emptyset^c = \vec{E}$, $\vec{E}^c = \emptyset$ y de la parte iii) del teorema anterior. □

Ejemplo 1.3.3. Para convencerse que las propiedades i) en los teoremas 1.3.1 y 1.3.2 no son en general ciertas para una familia no finita de partes, es suficiente demostrar que en un e.v.n. $\bigcap_{i=1}^{\infty} B'(\vec{a}, 1/i) = \{\vec{a}\}$ y que este conjunto no es abierto y, que $\bigcup_{i=1}^{\infty} B(\vec{a}, (i-1)/i) = B'(\vec{a}, 1)$ y que este conjunto no es cerrado.

1.3.1 Interior, adherencia y frontera

Definición 1.3.2. Se llama interior de un conjunto A en un e.v.n. $\vec{E}$ al conjunto:

$$\operatorname{int} A := \{\vec{x} \in A \ : \ \text{ existe } B(\vec{x}, \delta) \subset A\} \tag{1.3.1}$$

Nota 1.3.2. De la definición anterior se deduce fácilmente que $\operatorname{int} A$ es un conjunto abierto contenido en A. Es en efecto, el mayor abierto contenido en A, esto es, la unión de todos los abiertos contenidos en A. Vemos entonces que un conjunto A es abierto si y solo si $A = \operatorname{int} A$.

Definición 1.3.3. Se llama adherencia de un conjunto A en un e.v.n. $\vec{E}$ al conjunto:

$$\bar{A} := \{\vec{x} \in \vec{E} \ : \ B(\vec{x}, \varepsilon) \cap A \neq \emptyset \text{ para todo } \varepsilon > 0\} \tag{1.3.2}$$

Nota 1.3.3. De la definición anterior se deduce que $\bar{A}$ es un conjunto cerrado que contiene a A. Es en efecto, el menor cerrado que contiene a A, esto es, la intersección de todos los cerrados que contienen a A. Vemos entonces que un conjunto A es cerrado si y solo si $A = \bar{A}$.

Nota 1.3.4. Dadas dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ en un e.v. $\vec{E}$, el interior de un conjunto $A \subset \vec{E}$ no será necesariamente el mismo si dotamos a $\vec{E}$ de la primera o de la segunda de estas normas.

En la Nota 1.3.2 decíamos que el interior de A es el mayor abierto contenido en A y en la Nota 1.3.1 decíamos que los conjuntos abiertos son los mismos para dos normas equivalentes. Concluímos entonces que el interior de A es el mismo para dos normas equivalentes.

Del mismo modo, basados en las notas 1.3.3 y 1.3.1 concluímos que la adherencia de un conjunto es le misma para dos normas equivalentes.

1.4 Sucesiones en un e.v.n.

Definición 1.4.1. Diremos que una sucesión $\{\vec{a}_k\}$ en un e.v.n. $\vec{E}$ converge a un elemento $\vec{a} \in \vec{E}$ si para toda bola $B(\vec{a}, \varepsilon)$ existe un entero k_0 tal que $\vec{a}_k \in B(\vec{a}, \varepsilon)$ para todo $k \geq k_0$. Dicho en otras palabras si: *para todo $\varepsilon > 0$, existe un entero k_0, tal que $\|\vec{a}_k - \vec{a}\| \leq \varepsilon$ para todo $k \geq k_0$.*

El elemento $\vec{a}$ se llama límite de la sucesión y escribimos

$$\lim_k \vec{a}_k = \vec{a} \quad \text{ o bien } \quad \vec{a}_k \to \vec{a}$$

Nota 1.4.1. Es fácil verificar que el límite de una sucesión convergente es único. En efecto, si $\lim \vec{a}_k = \vec{a}$ y $\lim \vec{a}_k = \vec{b}$, se tendrá que para todo $\varepsilon > 0$ existen enteros k_0 y m_0 tales que, $\|\vec{a}_k - \vec{a}\| \leq \varepsilon/2$ para todo $k \geq k_0$ y $\|\vec{a}_k - \vec{b}\| \leq \varepsilon/2$ para todo $k \geq m_0$. Si denotamos $\bar{k} := \max\{k_0,\ m_0\}$ obtenemos

$$\|\vec{a} - \vec{b}\| \leq \|\vec{a} - \vec{a}_{\bar{k}}\| + \|\vec{a}_{\bar{k}} - \vec{b}\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Como esta desigualdad se tiene para todo $\varepsilon > 0$, concluimos que $\|\vec{a} - \vec{b}\| = 0$ y, de acuerdo a la propiedad (1.2.1), que $\vec{a} = \vec{b}$.

Nota 1.4.2. De acuerdo a lo que decía la Nota 1.2.2, podemos deducir que si $\{\vec{a}_k\}$ es una sucesión convergente a un elemento $\vec{a}$ en un e.v.n. $\vec{E}$, la sucesión sigue siendo convergente a $\vec{a}$ si cambiamos la norma de $\vec{E}$ por otra equivalente.

Lema 1.4.1. $\lim_k \vec{a}_k = \vec{a} \iff \lim_k \|\vec{a}_k - \vec{a}\| = 0$

Demostración. $\lim \vec{a}_k = \vec{a} \iff \forall\, \varepsilon > 0,\ \ \exists\ k_0 \in \mathbb{N}$ tal que, $\|\vec{a}_k - \vec{a}\| \leq \varepsilon\ \ \forall k \geq k_0 \iff \lim \|a_k - \vec{a}\| = 0$. Esta última equivalencia se desprende de la definición de convergencia a 0 de una sucesión en $\mathbb{R}_+$. $\square$

Teorema 1.4.1. *Una parte A de un e.v.n. $\vec{E}$ es cerrada si y solo si toda sucesión convergente de elementos de A, tiene su límite en A.*

Demostración. En la Nota 1.3.3 veíamos que A es un conjunto cerrado si y solo si $A = \bar{A}$. Usaremos este hecho en las dos partes de la demostración.

Supongamos que A es un conjunto cerrado y que $\{\vec{a}_k\}$ es una sucesión de elementos de A convergente a un elemento $\vec{a} \in \vec{E}$. Demostremos entonces que $\vec{a} \in A$:

$$\begin{aligned}
&\vec{a}_k \in A \;\; \forall\, k \geq 0 \text{ y } \lim \vec{a}_k = \vec{a} \\
&\Rightarrow \vec{a}_k \in A \;\; \forall\, k \geq 0 \text{ y } \forall\, \varepsilon > 0 \; \exists\, k_0 \in \mathbb{N} \text{ tal que } a_{k_0} \in B(\vec{a}, \varepsilon) \\
&\Rightarrow \forall\, \varepsilon > 0 \; B(\vec{a}, \varepsilon) \cap A \neq \emptyset \Rightarrow \;\; \vec{a} \in \bar{A}.
\end{aligned}$$

Como A es un conjunto cerrado, esta propiedad implica que $\vec{a} \in A$.

Supongamos ahora que A es un conjunto tal que, toda sucesión $\{\vec{a}_k\}$ en A, convergente a un elemento $\vec{a} \in \vec{E}$, verifica que $\vec{a} \in A$. Demostremos entonces que A es cerrado:

$$\begin{aligned}
\vec{a} \in \bar{A} \;\; &\Rightarrow \;\; B(\vec{a}, \varepsilon) \cap A \neq \emptyset \quad \forall\, \varepsilon > 0 \\
&\Rightarrow \;\; B(\vec{a}, 1/k) \cap A \neq \emptyset \quad \forall\, k \in \mathbb{N} \\
&\Rightarrow \;\; \exists\, \vec{a}_k \in B(\vec{a}, 1/k) \cap A \quad \forall\, k \in \mathbb{N} \\
&\Rightarrow \;\; \{\vec{a}_k\} \text{ está en } A \text{ y } \lim \vec{a}_k = \vec{a} \\
&\Rightarrow \;\; \vec{a} \in A.
\end{aligned}$$

Hemos así demostrado que $\bar{A} \subset A$ y, por lo tanto $A = \bar{A}$ y entonces A es cerrado. □

Teorema 1.4.2. *Sea $\{\vec{a}_k\}$ una sucesión en $\mathbb{R}^n$ que denotaremos $\vec{a}_k := (a_k^1, ..., a_k^n)$. La sucesión $\{\vec{a}_k\}$ converge a un elemento $\vec{a} := (a^1, ..., a^n) \in \mathbb{R}^n$ si y solo si cada una de las n sucesiones $\{a_k^i\}$ (para $i = 1, ..., n$) converge a $a^i \in \mathbb{R}$.*

Demostración. De acuerdo a la Nota 1.4.2, para demostrar este teorema podemos usar cualquiera de las normas del Ejemplo 1.2.1

$$\begin{aligned}
\text{i) } \lim_{\mathrm{k}} \mathrm{\tilde{a}_k} = \mathrm{\tilde{a}} \;\; &\Rightarrow \;\; \forall\, \varepsilon > 0 \; \exists\, k_0 \in \mathbb{N} \text{ tal que, } \|\vec{a}_k - \vec{a}\|_\infty \leq \varepsilon \quad \forall\, k \geq k_0, \\
&\Rightarrow \;\; \forall\, \varepsilon > 0 \; \exists\, k_0 \in \mathbb{N} \text{ tal que, } |a_k^i - a^i| \leq \varepsilon \;\; \forall\, i = 1, ..., n \;\; \forall\, k \geq k_0 \\
&\Rightarrow \;\; \forall\, \varepsilon > 0 \;\; \exists\, k_0 \in \mathbb{N} \text{ tal que, } |a_k^i - a^i| \leq \varepsilon \;\; \forall\, k \geq k_0 \;\; \forall\, i = 1, ..., n \\
&\Rightarrow \;\; \lim a_k^i = a^i \;\; \forall\, i = 1, ..., n.
\end{aligned}$$

ii) $\lim a_k^i = a^i \;\; \forall\, i = 1, ..., n \Rightarrow \;\; \forall\, \varepsilon > 0 \; \exists\, k_0^i \in \mathbb{N}$ tal que, $|a_k^i - a^i| \leq \varepsilon \;\; \forall\, k \geq k_0^i \;\; \forall\, i = 1, ..., n$. Si definimos $k_0 := \max\{k_0^i, i = 1, ..., n\}$ vemos que $|a_k^i - a^i| \leq \varepsilon \;\; \forall k \geq k_0$ $\forall\, i = 1, ..., n$ lo que podemos escribir $\|\vec{a}_k - \vec{a}\|_\infty \leq \varepsilon$ para todo $k \geq k_0$. Esto demuestra que $\lim \vec{a}_k = \vec{a}$. □

Teorema 1.4.3. *Sean $\vec{E}_1, ..., \vec{E}_n$ n e.v.n. y sea $\vec{E} = \vec{E}_1 \times ... \times \vec{E}_n$ el e.v.n. producto (ver Ejemplo 1.2.5). Sea $\{\vec{a}_k\}$ una sucesión en $\vec{E}$ que escribiremos $\vec{a}_k = (\vec{a}_k{}^1, ..., \vec{a}_k{}^n)$. La sucesión $\{\vec{a}_k\}$ converge a un elemento $\vec{a} = (\vec{a}\,^1, ..., \vec{a}\,^n) \in \vec{E}$, si y solo si cada una de las n sucesiones $\{\vec{a}_k{}^i\}$ (para $i = 1, ..., n$) converge a $\vec{a}\,^i \in \vec{E}_i$.*

Demostración. De acuerdo a la Nota 1.4.2 para demostrar este teorema podemos usar cualquiera de las normas del Ejemplo 1.2.5. Usando la norma $\|\cdot\|_\infty$, la demostración es idéntica a la del teorema anterior. $\square$

Teorema 1.4.4. *Sean $\{\vec{a}_k\}$ y $\{\vec{b}_k\}$ dos sucesiones en un e.v.n. $\vec{E}$ y sea $r \in \mathbb{R}$. Si estas dos sucesiones son convergentes, entonces las sucesiones $\{\vec{a}_k + \vec{b}_k\}$ y $\{r\vec{a}_k\}$ también son convergentes y, se tienen las igualdades*

$$\lim_k (\vec{a}_k + \vec{b}_k) = \lim_k \vec{a}_k + \lim_k \vec{b}_k \tag{1.4.1}$$

$$\lim_k (r\vec{a}_k) = r \lim_k \vec{a}_k \tag{1.4.2}$$

Demostración. Denotemos $\vec{a} := \lim \vec{a}_k$ y $\vec{b} := \lim \vec{b}_k$. De acuerdo al Lema 1.4.1 esto equivale a $\lim \|\vec{a}_k - \vec{a}\| = 0$ y $\lim \|\vec{b}_k - \vec{b}\| = 0$ y, como $\|(\vec{a}_k + \vec{b}_k) - (\vec{a} + \vec{b})\| \leq \|\vec{a}_k - \vec{a}\| + \|\vec{b}_k - \vec{b}\|$, deducimos que $\lim \|(\vec{a}_k + \vec{b}_k) - (\vec{a} + \vec{b})\| = 0$, lo que de acuerdo al mismo lema es equivalente a (1.4.1)

Por otro lado, vemos que $\lim_k \vec{a}_k = \vec{a} \Rightarrow \lim \|\vec{a}_k - \vec{a}\| = 0 \Rightarrow |r| \lim \|\vec{a}_k - \vec{a}\| = 0 \Rightarrow \lim |r| \|\vec{a}_k - \vec{a}\| = 0 \Rightarrow \lim \|r\vec{a}_k - r\vec{a}\| = 0 \Rightarrow \lim r\vec{a}_k = r\vec{a}$, que es lo que deseabamos probar. $\square$

Definición 1.4.2. Diremos que una sucesión $\{\vec{a}_k\}$ de un e.v.n. $\vec{E}$ tiene al elemento $\vec{a} \in \vec{E}$ como punto de acumulación si para toda bola $B(\vec{a}, \varepsilon)$ y todo entero k_0, existe un entero $\bar{k} \geq k_0$ tal que, $\vec{a}_{\bar{k}} \in B(\vec{a}, \varepsilon)$. Dicho en otras palabras si

"*para todo $\varepsilon > 0$ y todo $k_0 \in \mathbb{N}$ existe $\bar{k} \geq k_0$ tal que $\|\vec{a}_{\bar{k}} - \vec{a}\| \leq \varepsilon$*"

Nota 1.4.3. Queda claro de la Definición 1.4.1 y la Nota 1.4.1 que si $\lim \vec{a}_k = \vec{a}$, entonces $\vec{a}$ es también punto de acumulación de $\{\vec{a}_k\}$ y es el único. Es fácil ver que una sucesión puede tener muchos puntos de acumulación o ninguno (en esos casos, de acuerdo a la Nota 1.4.1, ella no será convergente).

Teorema 1.4.5. *Dada una sucesión* $\{\vec{a}_k\}$ *en un e.v.n.* $\vec{E}$, *un elemento* $\vec{a} \in \vec{E}$ *será punto de acumulación de* $\{\vec{a}_k\}$ *si y solo si existe una subsucesión* $\{\vec{a}_{\alpha(k)}\}$ *convergente a* $\vec{a}$.

Demostración. Sea $\vec{a} \in \vec{E}$ un punto de acumulación de $\{\vec{a}_k\}$. Construyamos una subsucesión $\{\vec{a}_{\alpha(k)}\}$ convergente a $\vec{a}$. Hagámoslo en forma recurrente a partir de $\vec{a}_{\alpha(1)} := \vec{a}_1$. Para esto, definamos $\vec{a}_{\alpha(k+1)}$ a partir de $\vec{a}_{\alpha(k)}$ de la siguiente manera: "$\alpha(k+1)$ es un entero que verifica $\alpha(k+1) > \alpha(k)$ y $\vec{a}_{\alpha(k+1)} \in B(\vec{a}, \frac{1}{k+1})$". La existencia de este $\alpha(k+1)$ es una consecuencia inmediata del hecho que $\vec{a}$ es un punto de acumulación de $\{\vec{a}_k\}$. Vemos que $\{\vec{a}_{\alpha(k)}\}$ es una subsucesión de $\{\vec{a}_k\}$ y, como $\|\vec{a}_{\alpha(k)} - \vec{a}\| \leq 1/k$ para todo k, vemos que $\lim_k \vec{a}_{\alpha(k)} = \vec{a}$.

Sea ahora $\{\vec{a}_{\alpha(k)}\}$ una subsucesión de $\{\vec{a}_k\}$ tal que $\lim \vec{a}_{\alpha(k)} = \vec{a}$. Demostremos que $\vec{a}$ es punto de acumulación de $\{a_k\}$. Sea $\varepsilon > 0$ y $k_0 \in \mathbb{N}$. Como $\vec{a}_{\alpha(k)} \to \vec{a}$ existirá $k_0' \in \mathbb{N}$ tal que $\vec{a}_{\alpha(k)} \in B(\vec{a}, \varepsilon)$ para todo $\alpha(k) \geq k_0'$. Si definimos $k' = \max\{k_0, k_0'\}$, vemos que $\alpha(k') \geq k_0$ y que $\vec{a}_{\alpha(k')} \in B(\vec{a}, \varepsilon)$, lo que demuestra que $\vec{a}$ es punto de acumulación de $\{\vec{a}_k\}$. $\square$

1.4.1 Sucesiones de Cauchy

Definición 1.4.3. Una sucesión $\{\vec{a}_k\}$ en un e.v.n. $\vec{E}$ se dirá de Cauchy si para todo $\varepsilon > 0$ existe un entero k_0 tal que $\|\vec{a}_k - \vec{a}_j\| \leq \varepsilon$ para todo $k, j \geq k_0$.

Nota 1.4.4. De las desigualdades (1.2.14), podemos deducir que si $\{\vec{a}_k\}$ es una sucesión de Cauchy en un e.v.n. $\vec{E}$ ella sigue siendo de Cauchy si cambiamos la norma de $\vec{E}$ por otra equivalente.

Teorema 1.4.6. *Toda sucesión convergente en un e.v.n. es de Cauchy.*

Demostración. Sea $\vec{a}_k \to \vec{a}$. Dado $\varepsilon > 0$ existirá $k_0 \in \mathbb{N}$ tal que $\|\vec{a}_k - \vec{a}\| \leq \varepsilon/2$ para todo $k \geq k_0$. Entonces $\|\vec{a}_k - \vec{a}_j\| \leq \|\vec{a}_k - \vec{a}\| + \|\vec{a} - \vec{a}_j\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ para todo $k, j \geq k_0$. Con esto queda demostrado que $\{\vec{a}_k\}$ es de Cauchy. $\square$

Nota 1.4.5. Corresponde ahora hacer la pregunta: ¿Es toda sucesión de Cauchy convergente? Los dos teoremas que siguen responderán afirmativamente a esta pregunta en dos casos particulares: i) cuando el e.v.n. $\vec{E}$ es de dimensión finita y, ii) cuando la sucesión de Cauchy tiene un punto de acumulación. Pero en general, la respuesta a esta pregunta no es afirmativa. Veremos más adelante que existen e.v.n. con sucesiones de Cauchy que no convergen. Naturalmente serán e.v.n. de dimensión infinita y la sucesión de Cauchy no convergente no tendrá ningún punto de acumulación.

Teorema 1.4.7. *Toda sucesión de Cauchy en $\mathbb{R}^n$ es convergente.*

Demostración. Sea $\{\vec{a}_k\}$ una sucesión de Cauchy en $\mathbb{R}^n$ y denotemos $\vec{a}_k := (a_k^1, ..., a_k^n)$. Según la Nota 1.4.4, podemos usar cualquier norma en $\mathbb{R}^n$. Entonces, puesto que para todo $i = 1, ..., n$ se tiene que $|a_k^i - a_j^i| \leq \|\vec{a}_k - \vec{a}_j\|_\infty$, es fácil deducir que las n sucesiones $\{a_k^i\}$ son de Cauchy en $\mathbb{R}$ y por lo tanto convergentes (esto último fue demostrado en el curso de Cálculo). Del Teorema 1.4.2 concluímos que $\{\vec{a}_k\}$ es convergente. $\square$

Teorema 1.4.8. *Si una sucesión de Cauchy en un e.v.n. $\vec{E}$ tiene un punto de acumulación, entonces ella converge a ese punto.*

Demostración. Sea $\{a_k\}$ una sucesión de Cauchy. Dado $\varepsilon > 0$, existirá $k_0 \in \mathbb{N}$ tal que $\|\vec{a}_k - \vec{a}_j\| \leq \varepsilon/2$ para todo $k, j \geq k_0$. Si además $\vec{a}$ es punto de acumulación de $\{a_k\}$, existirá $\bar{k} \geq k_0$ tal que $\|\vec{a}_{\bar{k}} - \vec{a}\| \leq \frac{\varepsilon}{2}$. De lo anterior deducimos que $\|\vec{a}_k - \vec{a}\| \leq \|\vec{a}_k - \vec{a}_{\bar{k}}\| + \|\vec{a}_{\bar{k}} - \vec{a}\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ para todo $k \geq k_0$. Con esto concluimos que $\vec{a}_k \to \vec{a}$. $\square$

Definición 1.4.4. Se llama espacio de *Banach* a todo e.v.n $\vec{E}$ cuyas sucesiones de Cauchy son siempre convergentes.

Nota 1.4.6. Casi todos los e.v.n. que se usan en los modelos matemáticos de la ingeniería, son espacios de Banach. En particular, de acuerdo al Teorema 1.4.7, todo e.v.n. de dimensión finita es de Banach.

1.5 Conjuntos compactos

Definición 1.5.1. Un conjunto A en un e.v.n. $\vec{E}$ se dirá compacto si toda sucesión en A tiene una subsucesión convergente a un elemento de A.

Nota 1.5.1. Del Teorema 1.4.5 se desprende que A es compacto si y solo si toda sucesión en A tiene un punto de acumulación en el conjunto A.

Teorema 1.5.1. *Toda sucesión de Cauchy en un conjunto compacto de un e.v.n. es convergente.*

Demostración. Es una consecuencia inmediata de la nota anterior y del Teorema 1.5.3
□
.

Teorema 1.5.2. *Todo conjunto compacto en un e.v.n. es cerrado y acotado.*

Demostración. Para demostrar que si A es compacto entonces es cerrado, usaremos el Teorema 1.4.1. Si $\{\vec{a}_k\}$ es una sucesión de elementos de A convergente a $\vec{a} \in \vec{E}$, como toda subsucesión de $\{\vec{a}_k\}$ también converge a $\vec{a}$, se tendrá obligatoriamente que $\vec{a} \in A$, lo que nos permite concluir que A es cerrado.

Ahora demostraremos que si A es compacto entonces es acotado. Razonemos por contradicción. Si A no fuera acotado se tendría que $A \subsetneq B(\vec{0}, k)$ para todo $k \in \mathbb{N}$, es decir que para todo $k \in \mathbb{N}$ existe un elemento $\vec{a}_k \in A \cap B(\vec{0}, k)^c$. Como $\|a_k\| > k$ para todo $k \in \mathbb{N}$ es evidente que la sucesión $\{a_k\}$ no puede tener una subsucesión convergente. En efecto, cualquiera sea $\vec{a} \in \vec{E}$, si nos damos $k_0 > \|\vec{a}\|$ y $\varepsilon := \frac{k_0 - \|\vec{a}\|}{2}$, vemos que $a_k \notin B(\vec{a}, \varepsilon)$ para todo $k \geq k_0$. La sucesión $\{\vec{a_k}\}$ contradice el hecho que A sea compacto. □

Teorema 1.5.3. *Si A es una parte cerrada y acotada de $\mathbb{R}^n$, entonces A es compacta.*

Demostración. Sea A un conjunto cerrado y acotado en $\mathbb{R}^n$. Sea $\{\vec{a}_k\}$ una sucesión en A. Consideremos en $\mathbb{R}^n$ la norma $\|\cdot\|_\infty$. Sean $\vec{c}_0 \in \mathbb{R}^n$ y $r > 0$ tales que $A \subset B(\vec{c}_0, r)$. Dividamos esta bola en 2^n bolas de radio $r/2$ y elijamos aquella que contiene una infinidad de términos de la sucesión $\{\vec{a}_k\}$, denotemosla $B(\vec{c}_1, r/2)$. Dividamos esta nueva bola en 2^n bolas de radio $r/4$ y elijamos aquella que contiene una infinidad de términos de la sucesión $\{\vec{a}_k\}$, denotemosla $B(\vec{c}_2, r/4)$. Construimos así una sucesión encajonada de bolas $B(\vec{c}_k, r/2^k)$ cada una de ellas con una infinidad de términos de la sucesión $\{\vec{a}_k\}$.
Para concluir vamos a demostrar que la sucesión $\{\vec{a}_k\}$ tiene una subsucesión $\{\vec{a}_{\alpha(k)}\}$ que es de Cauchy. Esto nos permitirá, de acuerdo al Teorema 1.4.7, concluir que $\{\vec{a}_k\}$ es convergente a un elemento $\vec{a} \in \mathbb{R}^n$ y, de acuerdo al Teorema 1.4.1, que $\vec{a} \in A$.

Definamos el entero $\alpha(1)$ de modo que $\vec{a}_{\alpha(1)} \in B(\vec{c}_1, r/2)$ y, en general, definamos para cada $k > 1$ el entero $\alpha(k) > \alpha(k-1)$ de modo que $\vec{a}_{\alpha(k)} \in B(\vec{c}_k, r/2^k)$. Demostremos entonces que $\{\vec{a}_{\alpha(k)}\}$ es de Cauchy. Dado $\varepsilon > 0$ sea $k_0 \in \mathbb{N}$ tal que $\varepsilon > r/2^{k_0-1}$, como la bola $B(\vec{c}_{k_0}, r/2^{k_0})$ contiene a todos los $\vec{a}_{\alpha(k)}$ para $k \geq k_0$ y como su diámetro es $r/2^{k_0-1}$, concluimos que $\|\vec{a}_{\alpha(k)} - \vec{a}_{\alpha(j)}\| \leq \varepsilon$ para todo $k, j \geq k_0$. $\square$

Teorema 1.5.4. *Toda bola cerrada en $\mathbb{R}^n$ es compacta.*

Demostración. Puesto que toda bola cerrada en un e.v.n. es un conjunto cerrado (ver Ejemplo 1.3.2) y acotado, del teorema anterior, concluimos que es compacta. $\square$

Teorema 1.5.5. *Toda sucesión acotada en $\mathbb{R}^n$ tiene una subsucesión convergente.*

Demostración. Si $\{\vec{a}_k\}$ es una sucesión acotada, ella debe estar contenida en una bola $B(\vec{0}, r)$ y, como por el teorema anterior sabemos que esta bola es compacta, concluimos que $\{\vec{a}_k\}$ debe tener una subsucesión convergente. $\square$

Nota 1.5.2. Más adelante veremos que los tres últimos teoremas no son válidos en un e.v.n. de dimensión infinita. Mostraremos que existen e.v.n. donde la bola $B(\vec{0}, 1)$ no es compacta, lo que equivale a decir que existen sucesiones acotadas sin ningún punto de acumulación.

Para cerrar este capítulo, veremos una aplicación del Teorema 1.5.5 la que nos dirá que todo s.e.v. de dimensión finita en un e.v.n. es cerrado.

Definición 1.5.2. Sea $\vec{E}$ un e.v.n. un conjunto finito $\{\vec{v}_1, ..., \vec{v}_n\} \subset \vec{E}$ se dice linealmente independiente (l.i.) si

$$\sum_{i=1}^{n} \lambda_i \vec{v}_i = 0 \Rightarrow \ \lambda_i = 0 \quad \forall\ i = 1, ..., n.$$

Teorema 1.5.6. *Sea $\vec{E}$ un e.v.n. y $\vec{F}$ un s.e.v. de $\vec{E}$ de dimensión finita generado por el conjunto l.i. $\{\vec{v}_1, ..., \vec{v}_n\}$, entonces $\vec{F}$ es cerrado.*

Demostración. Sea $\vec{w}_k = \sum_{i=1}^{n} \lambda_i^k \vec{v}_i \to \vec{w}$. Debemos demostrar que $\vec{w} \in \vec{F}$. Considere $\vec{\lambda}_k = (\lambda_1^k, ..., \lambda_n^k) \in \mathbb{R}^n$. Si $\|\vec{\lambda}_k\|_\infty \to \infty$ entonces la sucesión $\frac{\vec{\lambda}_k}{\|\vec{\lambda}_k\|_\infty}$ es acotada, y por el Teorema 1.5.5 se obtiene que existe una subsucesión

$$\frac{\vec{\lambda}_{k_j}}{\|\vec{\lambda}_{k_j}\|_\infty} \to \vec{\mu} = (\mu_1, ..., \mu_n) \neq 0. \tag{1.5.1}$$

Como $\vec{w}_k$ es una sucesión convergente y por lo tanto acotada, notamos que $\frac{\vec{w}_{k_j}}{\|\vec{\lambda}_{k_j}\|_\infty} \to 0$ lo que implica

$$\sum_{i=1}^{n} \mu_i \vec{v}_i = 0.$$

Por la independencia lineal de $\{\vec{v}_1, ..., \vec{v}_n\}$ resulta $\mu_i = 0$ para todo $i = 1, ..., n$ lo que es una contradicción con (1.5.1). Por lo tanto existe $M \geq 0$ tal que $\|\vec{\lambda}_k\|_\infty \leq M$. Nuevamente utilizando el Teorema 1.5.5 concluímos la existencia de una subsucesión $\vec{\lambda}_{k_j} \to \vec{\lambda} = (\lambda_1, ..., \lambda_n)$ y por lo tanto $\vec{w}_{k_j} \to \sum_{i=1}^{n} \lambda_i \vec{v}_i = \vec{w}$ lo que muestra la cerradura de $\vec{F}$.

$\square$

1.6 Ejercicios

1. *a*) Dada una norma $\|\cdot\|$ en $\mathbb{R}^n$ y una matriz no singular P de $n \times n$ demuestre que la función $\|\vec{x}\|_P = \|P\vec{x}\|$ es también una norma en $\mathbb{R}^n$.

   b) Demuestre que en el e.v. $\mathbb{R}^2$, la función $\|\vec{x}\|_e = \left[\frac{x_1^2}{9} + 4x_2^2\right]^{1/2}$ es una norma. Haga un dibujo de los conjuntos $B(\vec{0}, 1)$ y $B'(\vec{0}, 1)$.

2. Demuestre que en el e.v. $\mathcal{C}([a, b], \mathbb{R})$ de las funciones continuas de $[a, b]$ en $\mathbb{R}$, la función $\|f\|_1 = \int_a^b |f(t)|dt$ es una norma. Demuestre que en el e.v. $I([a, b], \mathbb{R})$ de las funciones integrables de $[a, b]$ en $\mathbb{R}$, la función $\|\cdot\|$, no es una norma.

3. Demuestre que en el e.v. $\mathcal{C}([a, b], \mathbb{R})$ la función $\|f\|_2 = \left[\int_a^b f(t)^2 dt\right]^{1/2}$ es una norma.

4. Demuestre (usando el teorema fundamental del álgebra) que en el e.v. $\mathcal{P}$ de los polinomios de una variable real, las siguientes funciones son normas

$$\begin{aligned} \|p\| &= \max_{x \in [0,1]} |p(x)| \\ \|p\|' &= \max\{|\alpha_0|, |\alpha_1|, ..., |\alpha_n|\} \quad (\text{ donde } \quad p(x) = \sum_{i=0}^{n} \alpha_i x^i) \end{aligned}$$

5. Demuestre que en el e.v. l_∞ de las sucesiones reales acotadas, la función $\|\{r_k\}\|_\infty = \sup_{k \in \mathbb{N}} |r_k|$ es una norma.

6. Demuestre que en el e.v. l_1 de las sucesiones reales que verifican "$\sum |r_k|$ convergente", la función $\|\{r_k\}\|_1 = \sum |r_k|$ es una norma.

7. Demuestre que en un e.v.n. todo conjunto formado por un solo elemento es cerrado.

8. Demuestre que en un e.v.n. los únicos conjuntos que son abiertos y cerrados al mismo tiempo, son el conjunto vacío y el espacio entero.

9. Determine el interior y la adherencia de cada uno de los siguientes subconjuntos de $\mathbb{R}^2$:
$$\{\vec{x} : x_2 > x_1\}, \{\vec{x} : 0 < \|\vec{x}\|_2 \leq 1\}$$
$$\{\vec{x} : x_1 = x_2 \text{ y } x_1 > 0\}, \{\vec{x} : x_1 \in \mathbb{Q}\}, \{(\frac{1}{k}, (-1)^k) : k \in \mathbb{N}\}$$

10. Dados dos conjuntos A, B en un e.v.n. $\vec{E}$, demuestre que

    a) $\operatorname{int}(A \cap B) = \operatorname{int} A \cap \operatorname{int} B$.
    b) $\operatorname{int}(A \cup B) \supset \operatorname{int} A \cup \operatorname{int} B$ (dé un ejemplo donde no hay igualdad).
    c) $\overline{A \cup B} = \bar{A} \cup \bar{B}$
    d) $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$ (dé un ejemplo donde no hay igualdad)
    e) $\bar{A} = \operatorname{int} A \cup \{\vec{x} \in \vec{E} : B(\vec{x}, \varepsilon) \cap A \neq \emptyset \text{ y } \ B(\vec{x}, \varepsilon) \cap A^c \neq \emptyset \text{ para todo } \ \varepsilon > 0\}$
    f) $A \subset B \ \Rightarrow \ \operatorname{int} A \subset \operatorname{int} B$ y $\bar{A} \subset \bar{B}$.
    g) $\operatorname{int} A \cap B = \emptyset \Leftrightarrow \operatorname{int} A \cap \bar{B} = \emptyset$
    h) $\bar{A} = \vec{E}$ y $\operatorname{int} B \cap A = \emptyset \Rightarrow \operatorname{int} B = \emptyset$
    i) $\operatorname{int}(A^c) = (\bar{A})^c$
    j) $(\overline{A^c}) = (\operatorname{int} A)^c$

11. Si definimos la "distancia" de un punto $\vec{x}$ de un e.v.n. $\vec{E}$ a un conjunto $A \subset \vec{E}$ como la cantidad
$$d_A(\vec{x}) = \inf\{\|\vec{x} - \vec{y}\| : \vec{y} \in A\}$$
demuestre que $\bar{A} = \{\vec{x} \in \vec{E} : d_A(\vec{x}) = 0\}$.

12. Si A es un conjunto en un e.v.n. $\vec{E}$ con la propiedad "$\bar{A} = \vec{E}$ y $\operatorname{int} A = \phi$", entonces el conjunto A^c (complemento de A) tiene la misma propiedad.

13. En el e.v. $\mathcal{C}([0,1], \mathbb{R})$ dotado de la norma $\|\cdot\|_1$ (ver Problema 2) demuestre que la sucesión $\{f_k\}$ definida por
$$f_k(x) = \begin{cases} 1 & \text{si } \ x \in [0, \frac{1}{2}] \\ -2^k(x - \frac{1}{2}) + 1 & \text{si } \ x \in [\frac{1}{2}, \frac{1}{2} + \frac{1}{2^k}] \\ 0 & \text{si } x \in [\frac{1}{2} + \frac{1}{2^k}, 1] \end{cases}$$
es de Cauchy y no es convergente.

14. Demuestre que en el e.v. $\mathcal{A}([0,1],\mathbb{R})$ de las funciones acotadas de $[0,1]$ en $\mathbb{R}$, dotado de la norma $\|\cdot\|_\infty$, la sucesión $\{f_k\}$ definida en el problema anterior no es convergente.

15. Demuestre que en el e.v.n. l_1, definido en el Problema 6, el conjunto $P = \{\{x_k\} \ : \ x_k \geq 0 \quad \forall\, k \in \mathbb{N}\}$ tiene interior vacío.

16. Demuestre que en el e.v.n. l_1, definido en el Problema 6, la bola $B(\vec{0},1)$ no es compacta.

17. Demuestre que un subconjunto cerrado de un conjunto compacto, es también compacto.

18. Demuestre que la intersección de un conjunto compacto con un conjunto cerrado, es un conjunto compacto.

19. Si $\{A_i\}_{i\in\mathbb{N}}$ es una familia de conjuntos compactos en un e.v.n. tal que $\bigcap\limits_{i\in\mathbb{N}} A_i = \emptyset$, entonces existe un número finito de conjuntos de la familia: A_{i_1}, A_{i_2}, ..., A_{i_n}, con la propiedad $\bigcap\limits_{j=1}^{n} A_{i_j} = \emptyset$.

20. Si A es un conjunto compacto en un e.v.n. y si $\{A_i\}_{i\in\mathbb{N}}$ es una familia de conjuntos abiertos tales que $\bigcup\limits_{i\in\mathbb{N}} A_i \supset A$, entonces existe un número finito de conjuntos de la familia: $A_{i_1}, A_{i_2}, ..., A_{i_n}$, con la propiedad $\bigcup\limits_{j=1}^{n} A_{i_j} \supset A$.

21. Dado el conjunto $A = [0,1[$ en $\mathbb{R}$, encuentre una familia de abiertos $\{A_i\}_{i\in\mathbb{N}}$ cuya unión contenga al conjunto A y tal que no exista un número finito de ellos cuya unión contenga a A.

22. Si $\{A_i\}_{i\in\mathbb{N}}$ es una familia decreciente de conjuntos compactos (no vacíos) en un e.v.n., demuestre que $\bigcap\limits_{i\in\mathbb{N}} A_i \neq \emptyset$.

23. *a*) De un ejemplo en $\mathbb{R}$ de una familia decreciente de conjuntos acotados (no vacíos) cuya intersección sea vacía.

    b) De un ejemplo en $\mathbb{R}$ de una familia decreciente de conjuntos cerrados (no vacíos) cuya intersección sea vacía.

CAPÍTULO 2

FUNCIONES DEFINIDAS EN UN E.V.N. $\vec{E}$ CON VALORES EN UN E.V.N. $\vec{F}$

2.1 Introducción

Este capítulo está completamente dedicado a estudiar la continuidad de las funciones definidas en un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Tal como se vió en el caso de funciones de una variable real, la continudad seguirá siendo el punto de partida del estudio de funciones, ahora definidas en un e.v.n.

Veremos también la noción de límite de una función que está íntimamente relacionada con la de continuidad. Luego estudiaremos algunas de las propiedades fundamentales de las funciones continuas definidas en un conjunto compacto. El Teorema 2.5.2 es el de mayor importancia en esa sección.

En la Sección 2.7 estudiaremos las funciones lineales continuas y analizaremos el primer espacio vectorial normado de dimensión no finita. El Teorema 2.7.1 es el de mayor importancia en dicha sección.

Para terminar el capítulo, demostraremos una propiedad de las funciones continuas (el teorema de punto fijo) que constituye uno de los resultados más útiles de las matemáticas aplicadas.

2.2 Continuidad

Definición 2.2.1. Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, se dirá continua en $\vec{a} \in A$ si para toda bola $B(f(\vec{a}), \varepsilon)$ existe una bola $B(\vec{a}, \delta)$ tal que

$$f[B(\vec{a}, \delta) \cap A] \subset B(f(\vec{a}), \varepsilon) \tag{2.2.1}$$

dicho en otras palabras, si

$$''\forall\ \varepsilon > 0 \ \ \exists \ \ \delta > 0 \text{ tal que, } \ \vec{x} \in A \text{ y } \|\vec{x} - \vec{a}\| \leq \delta \ \Rightarrow \ \|f(\vec{x}) - f(\vec{a})\| \leq \varepsilon''.$$

(se da por entendido que las normas escritas anteriormente corresponden a espacios normados diferentes) La función f se dirá continua si ella es continua en todo elemento de A.

Nota 2.2.1. Es fácil ver que si $f : A \to \vec{F}$ es continua en $\vec{a} \in A \subset \vec{E}$, entonces f sigue siendo continua en $\vec{a}$ si cambiamos la norma de $\vec{E}$ o la de $\vec{F}$ por otra equivalente.

Nota 2.2.2. Cuando $A = \vec{E}$ la inclusión en (2.2.1) se reduce a

$$f[B(\vec{a}, \delta)] \subset B(f(\vec{a}), \varepsilon) \tag{2.2.2}$$

y la continuidad de f en $\vec{a}$ se expresa diciendo

$$''\forall\ \varepsilon > 0 \ \ \exists\, \delta > 0 \text{ tal que, } \|\vec{x} - \vec{a}\| \leq \delta \ \Rightarrow \ \|f(\vec{x}) - f(\vec{a})\| \leq \varepsilon''$$

Teorema 2.2.1. *Si f y g son dos funciones definidas en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, continuas en $\vec{a} \in A$, entonces las funciones $f + g$ y λf (con $\lambda \in \mathbb{R}$) también son continuas en $\vec{a}$.*

Demostración. i) Dado $\varepsilon > 0$, si f y g son continuas en $\vec{a}$, existen $\delta_1, \delta_2 > 0$ tales que

$$\vec{x} \in A, \ \|\vec{x} - \vec{a}\| \leq \delta_1 \ \Rightarrow \ \|f(\vec{x}) - f(\vec{a})\| \leq \frac{\varepsilon}{2} \quad \text{y}$$

$$\vec{x} \in A, \ \|\vec{x} - \vec{a}\| \leq \delta_2 \ \Rightarrow \ \|g(\vec{x}) - g(\vec{a})\| \leq \frac{\varepsilon}{2}.$$

Definiendo $\delta := \min\{\delta_1, \delta_2\}$ se obtiene:

$$\begin{aligned} \vec{x} \in A, \ \|\vec{x} - \vec{a}\| \leq \delta \quad \Rightarrow \quad & \|(f+g)(\vec{x}) - (f+g)(\vec{a})\| \leq \\ & \|f(\vec{x}) - f(\vec{a})\| + \|g(\vec{x}) - g(\vec{a})\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Concluímos así que $f + g$ es una función continua en $\vec{a}$.

ii) Si $\lambda = 0$ el resultado es evidente. Supondremos entonces que $\lambda \neq 0$. Dado $\varepsilon > 0$ si f es continua en $\vec{a}$, existe $\delta > 0$ tal que: $\vec{x} \in A$, $\|\vec{x} - \vec{a}\| \leq \delta \Rightarrow \|f(\vec{x}) - f(\vec{a})\| \leq \varepsilon/|\lambda|$, que es equivalente a $\|\lambda f(\vec{x}) - \lambda f(\vec{a})\| \leq \varepsilon$. Concluímos así que la función λf es continua en $\vec{a}$. $\square$

Teorema 2.2.2. *Dados tres e.v.n. $\vec{E}, \vec{F}, \vec{G}$, una parte A de $\vec{E}$ y dos funciones $f : A \to \vec{F}$ y $h : \vec{F} \to \vec{G}$ continuas en $\vec{a} \in A$ y $f(\vec{a}) \in \vec{F}$ respectivamente. Entonces la función $h \circ f$ también es continua en $\vec{a}$.*

Demostración. Dado $\varepsilon > 0$, por ser h continua en $f(\vec{a})$, existe $\eta > 0$ tal que

$$\|f(\vec{a}) - \vec{y}\| \leq \eta \Rightarrow \|h(f(\vec{a})) - h(\vec{y})\| \leq \varepsilon$$

y como f es continua en $\vec{a}$, existe $\delta > 0$ tal que

$$\vec{x} \in A, \|\vec{a} - \vec{x}\| \leq \delta \Rightarrow \|f(\vec{a}) - f(\vec{x})\| \leq \eta$$

De estas dos implicaciones concluimos que

$$\vec{x} \in A, \|\vec{a} - \vec{x}\| \leq \delta \Rightarrow \|h(f(\vec{a})) - h(f(\vec{x}))\| \leq \varepsilon$$

es decir, que $h \circ f$ es continua en $\vec{a}$. $\square$

Teorema 2.2.3. *Si f y g son dos funciones definidas en una parte A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$, continuas en $\vec{a} \in A$, entonces las funciones $fg, 1/f$ (suponiendo $f(\vec{a}) \neq 0$) y $\max\{f, g\}$ siguen siendo continuas en $\vec{a}$.*

Demostración. Como $fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$, para demostrar la continuidad de $f \cdot g$ en $\vec{a}$, usando el Teorema 2.2.1, sólo hay que demostrar que el cuadrado de una función continua en $\vec{a}$ sigue siendo una función continua en $\vec{a}$.

Mostremos que f^2 es continua en $\vec{a}$. Como f^2 es la composición de $f : A \to \mathbb{R}$ con $h : \mathbb{R} \to \mathbb{R}$, definida por $h(y) = y^2$, y como ambas funciones son continuas en $\vec{a}$ y $f(\vec{a})$ respectivamente, el teorema anterior nos muestra que f^2 es continua en $\vec{a}$.

La continuidad de $1/f$ en $\vec{a}$, es también consecuencia del teorema anterior puesto que $1/f$ es la composición de $f : A \to \mathbb{R}$ con $h : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, definida por $h(y) = y^{-1}$, que son continuas en $\vec{a}$ y $f(\vec{a})$ respectivamente.

Finalmente, dado $\varepsilon > 0$, existen δ_1, $\delta_2 > 0$ tales que $\vec{x} \in A, \|\vec{x} - \vec{a}\| \leq \delta_1 \Rightarrow |f(\vec{x}) - f(\vec{a})| \leq \varepsilon$ y, $\vec{x} \in A$, $\|\vec{x} - \vec{a}\| \leq \delta_2 \Rightarrow |g(\vec{x}) - g(\vec{a})| \leq \varepsilon$. De la desigualdad

$$|\max\{f(\vec{x}), g(\vec{x})\} - \max\{f(\vec{a}), g(\vec{a})\}| \leq \max\{|f(\vec{x}) - f(\vec{a})|, |g(\vec{x}) - g(\vec{a})|\}$$

definiendo $\delta = \text{mín}\{\delta_1, \delta_2\}$ obtenemos que

$$\vec{x} \in A,\ \|x - \vec{a}\| \leq \delta \ \Rightarrow\ |\,\text{máx}\{f(\vec{x}), g(\vec{x})\} - \text{máx}\{f(\vec{a}), g(\vec{a})\}| \leq \varepsilon$$

lo que demuestra que la función $\text{máx}\{f, g\}$ es continua en $\vec{a}$. $\square$

Teorema 2.2.4. *La norma de un e.v.n. $\vec{E}$ es una función continua de $\vec{E}$ en $\mathbb{R}$.*

Demostración. Es una consecuencia inmediata de la desigualdad

$$|\|\vec{x}\| - \|\vec{a}\|| \leq \|\vec{x} - \vec{a}\|$$

que es válida para todo $\vec{x}$ y todo $\vec{a}$ en $\vec{E}$. $\square$

Teorema 2.2.5. *Sea A una parte de un e.v.n. $\vec{E}$ y d_A la función de $\vec{E}$ en $\mathbb{R}$ definida por*

$$d_A(\vec{x}) := \text{ínf}\{\|\vec{x} - \vec{z}\| : \vec{z} \in A\}. \tag{2.2.3}$$

Entonces d_A es una función continua, más aún, ella verifica

$$|d_A(\vec{x}) - d_A(\vec{x}\,')| \leq \|\vec{x} - \vec{x}\,'\| \ \ \textit{para todo } \vec{x}, \vec{x}\,' \in \vec{E}. \tag{2.2.4}$$

A la cantidad $d_A(\vec{x})$ se le llama usualmente distancia de $\vec{x}$ al conjunto A y a d_A función distancia al conjunto A.

Demostración. Sea $\varepsilon > 0$. Por definición de ínfimo, para todo $\vec{x}\,' \in \vec{E}$ debe existir $\vec{z} \in A$ tal que $\|\vec{x}\,' - \vec{z}\| \leq d_A(\vec{x}\,') + \varepsilon$. Por lo tanto, para todo $\vec{x}$ y todo $\vec{x}\,'$ se tendrá

$$d_A(\vec{x}) \leq \|\vec{x} - \vec{z}\| \leq \|\vec{x} - \vec{x}\,'\| + \|\vec{x}\,' - \vec{z}\| \leq \|\vec{x} - \vec{x}\,'\| + d_A(\vec{x}\,') + \varepsilon.$$

Como $\varepsilon > 0$ es cualquiera (tan pequeño como uno quiera) se tendrá

$$d_A(\vec{x}) - d_A(\vec{x}\,') \leq \|\vec{x} - \vec{x}\,'\| \quad \text{para todo } \vec{x},\ \vec{x}\,' \in \vec{E}$$

e intercambiando $\vec{x}\,'$ con $\vec{x}$ se tiene

$$-(d_A(\vec{x}) - d_A(\vec{x}\,')) \leq \|\vec{x} - \vec{x}\,'\| \quad \text{para todo } \vec{x},\ \vec{x}\,' \in \vec{E}.$$

Estas dos últimas desigualdades son equivalentes a la desigualdad (2.2.4). $\square$

2.3 Límite de funciones y caracterización de la continuidad

Definición 2.3.1. Sea f una función definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$ y sea $\vec{a} \in \bar{A}$. Diremos que f tiende a $\vec{b} \in \vec{F}$ cuando $\vec{x}$ tiende a $\vec{a}$ en A si para toda bola $B(\vec{b}, \varepsilon)$ existe una bola $B(\vec{a}, \delta)$ tal que

$$f[B(\vec{a}, \delta) \cap A] \subset B(\vec{b}, \varepsilon) \tag{2.3.1}$$

dicho en otras palabras, si

$$''\forall\ \varepsilon > 0 \ \ \exists\ \delta > 0 \text{ tal que , } \vec{x} \in A \ \text{ y } \ \|\vec{x} - \vec{a}\| \leq \delta \ \Rightarrow \ \|f(\vec{x}) - \vec{b}\| \leq \varepsilon''$$

Escribiremos entonces $\lim\limits_{\substack{\vec{x} \to \vec{a} \\ x \in A}} f(\vec{x}) = \vec{b}$ y, en el caso en que $\vec{a} \in \operatorname{int} A$, escribiremos simplemente $\lim\limits_{\vec{x} \to \vec{a}} f(\vec{x}) = \vec{b}$.

Teorema 2.3.1. *Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$ será continua en $\vec{a} \in A$ si y solo si*

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} f(\vec{x}) = f(\vec{a}) \tag{2.3.2}$$

Demostración. Es una consecuencia inmediata de las inclusiones en (2.2.1) y (2.3.1). □

Teorema 2.3.2. *Sean f, g dos funciones definidas en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Considere $\vec{a} \in \bar{A}$ y $\lambda \in \mathbb{R}$. Si los límites de f y g cuando $\vec{x}$ tiende a $\vec{a}$ en A existen, entonces los límites de $f + g$ y λf cuando $\vec{x}$ tiende a $\vec{a}$ en A también existen y se tienen las igualdades*

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} (f + g)(\vec{x}) = \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} f(\vec{x}) + \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} g(\vec{x}) \tag{2.3.3}$$

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} \lambda f(\vec{x}) = \lambda \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} f(\vec{x}). \tag{2.3.4}$$

Demostración. La demostración es casi idéntica a la del Teorema 2.2.1 □

Teorema 2.3.3. *Dados tres e.v.n. $\vec{E}, \vec{F}, \vec{G}$, un subconjunto A en $\vec{E}, \vec{a} \in \bar{A}$, y dos funciones $f : A \to \vec{F}$ y $h : f(A) \to \vec{G}$. Si existen los límites*

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} f(\vec{x}) = \vec{l} \text{ y } \lim_{\substack{\vec{y} \to \vec{l} \\ \vec{y} \in f(A)}} h(\vec{y}) = \vec{m},$$

entonces el límite de $h \circ f$ cuando $\vec{x}$ tiende $\vec{a}$ en A también existe y se tiene

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} (h \circ f)(\vec{x}) = \vec{m} \tag{2.3.5}$$

Demostración. La demostración es casi idéntica a la del Teorema 2.2.2. □

Teorema 2.3.4. *Sean f, g dos funciones definidas en una parte A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$ y $\vec{a} \in \bar{A}$. Si los límites de f y g cuando $\vec{x}$ tiende a $\vec{a}$ en A existen, entonces los límites de $f \cdot g$ y de $1/f$ (suponiendo $\lim_{x \to \vec{a}} f(\vec{x}) \neq 0$) cuando x tiende a $\vec{a}$ en A también existen y, se tienen las igualdades*

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} (f \cdot g)(\vec{x}) = \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} f(\vec{x}) \cdot \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} g(\vec{x}) \tag{2.3.6}$$

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} (1/f)(\vec{x}) = 1/ \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in A}} f(\vec{x}) \tag{2.3.7}$$

Demostración. La demostración es casi idéntica a la del Teorema 2.2.3 □

2.3.1 Caracterización de la continuidad y el límite de una función mediante sucesiones

Teorema 2.3.5. *Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$ será continua en $\vec{a} \in A$ si y solo si para toda sucesión $\{\vec{x}_k\}$ en A convergente a $\vec{a}$, la sucesión $\{f(\vec{x}_k)\}$ converge a $f(\vec{a})$, es decir*

$$\lim_k f(\vec{x}_k) = f(\vec{a}) \quad \text{para toda sucesión } \vec{x}_k \to \vec{a} \text{ en } A. \tag{2.3.8}$$

Demostración. Supongamos que f es continua en $\vec{a}$ y demostremos que se tiene (2.3.8). Dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\vec{x} \in A, \|\vec{x} - \vec{a}\| \leq \delta \Rightarrow \|f(\vec{x}) - f(\vec{a})\| \leq \varepsilon$$

y por otra parte, para toda sucesión $\{\vec{x}_k\}$ en A convergente a $\vec{a}$ debe existir $k_0 \in \mathbb{N}$ tal que, $\|\vec{x}_k - \vec{a}\| \leq \delta \; \forall \; k \geq k_0$. Concluímos entonces que

$$\|f(\vec{x}_k) - f(\vec{a})\| \leq \varepsilon \; \forall \; k \geq k_0$$

lo que muestra que $\{f(\vec{x}_k)\}$ converge a $f(\vec{a})$.

Supongamos ahora que para toda sucesión $\vec{x}_k \to \vec{a}$ en A se tiene que $f(\vec{x}_k) \to f(\vec{a})$. Si f no fuera continua en $\vec{a}$, existiría $\varepsilon > 0$ tal que para todo $\delta > 0$ existe $\vec{x}_\delta \in A$ que verifica

$$\|\vec{x}_\delta - \vec{a}\| \leq \delta \;\text{ y }\; \|f(\vec{x}_\delta) - f(\vec{a})\| > \varepsilon.$$

Aplicando este razonamiento sucesivamente para $\delta = 1, 1/2, 1/3, ..., 1/k, ...$ concluímos que existiría una sucesión $\{\vec{x}_k\}$ en A que verifica:

$$\|\vec{x}_k - \vec{a}\| \leq 1/k \;\text{ y }\; \|f(\vec{x}_k) - f(\vec{a})\| > \varepsilon \quad \text{para todo} \quad k$$

lo que contradice la hipótesis pues $\vec{x}_k \to \vec{a}$ y $f(x_k)$ no converge a $f(\vec{a})$. Esto muestra que f debe ser continua en $\vec{a}$. $\square$

Teorema 2.3.6. *Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, tiende a $\vec{b} \in \vec{F}$ cuando $\vec{x}$ tiende a $\vec{a} \in \bar{A}$ en A, si y solo si para toda sucesión $\{\vec{x}_k\}$ en A convergente a $\vec{a}$, la sucesión $\{f(\vec{x}_k)\}$ converge a $\vec{b}$.*

Demostración. La demostración es casi identica a la del teorema anterior. $\square$

2.4 Funciones continuas con valores en $\mathbb{R}^m$

Dada una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}^m$, es usual caracterizar f por sus m funciones componentes $f_1, ..., f_m$ de A en $\mathbb{R}$, definidas de modo que para todo $\vec{x} \in A$ se tenga

$$f(\vec{x}) = (f_1(\vec{x}), ..., f_m(\vec{x})) \qquad (\in \mathbb{R}^m)$$

La misma caracterización se usa cuando f toma sus valores en un e.v.n producto $\vec{F}_1 \times ... \times \vec{F}_m$. En ese caso las m funciones componentes $f_1, ..., f_m$ tendrán sus valores en $\vec{F}_1, ..., \vec{F}_m$ respectivamente.

Teorema 2.4.1. *Una función f definida en una parte de A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}^m$ será continua en $\vec{a} \in A$ si y solo si cada una de sus m funciones componentes $f_i : A \to \mathbb{R}$ es continua en $\vec{a} \in A$.*

Demostración. De acuerdo a las Notas 1.2.3 y 2.2.1, podemos usar en $\mathbb{R}^m$ cualquier norma. Usaremos entonces la norma $\|\cdot\|_\infty$.

Supongamos f continua en $\vec{a}$. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\vec{x} \in A, \|\vec{x} - \vec{a}\| \leq \delta \Rightarrow \|f(\vec{x}) - f(\vec{a})\|_\infty \leq \varepsilon$$

lo que equivale a decir que para todo $i = 1...m$ se tiene

$$\vec{x} \in A, \|\vec{x} - \vec{a}\| \leq \delta \Rightarrow |f_i(\vec{x}) - f_i(\vec{a})| \leq \varepsilon.$$

Esto demuestra que las funciones $f_1, ..., f_m$ son todas continuas en $\vec{a}$.

Supongamos ahora que las funciones $f_1, ..., f_m$ son todas continuas en $\vec{a}$. Dado $\varepsilon > 0$, existen $\delta_1, ..., \delta_m > 0$ tales que para todo $i = 1, ..., m$

$$\vec{x} \in A, \|\vec{x} - \vec{a}\| \leq \delta_i \Rightarrow |f_i(\vec{x}) - f_i(\vec{a})| \leq \varepsilon.$$

Definiendo $\delta := \min\{\delta_1, ..., \delta_m\} > 0$ obtenemos para todo $i = 1, ..., m$

$$\|\vec{x} - \vec{a}\| \leq \delta \Rightarrow |f_i(\vec{x}) - f_i(\vec{a})| \leq \varepsilon$$

lo que equivale a decir que

$$\|\vec{x} - \vec{a}\| \leq \delta \Rightarrow \|f(\vec{x}) - f(\vec{a})\|_\infty \leq \varepsilon.$$

Esto demuestra la continuidad de f en $\vec{a}$. □

Teorema 2.4.2. *Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}^m$, tiende a $\vec{b} \in \mathbb{R}^m$ cuando $\vec{x}$ tiende a $\vec{a} \in \bar{A}$ en A, si y solo si cada una de las m componentes $f_i : A \to \mathbb{R}$ de la función f, tiende a la componente $b_i \in \mathbb{R}$ de $\vec{b}$ cuando $\vec{x}$ tiende a $\vec{a}$ en A. Esto lo escribimos*

$$\lim_{\substack{\vec{x}\to\vec{a} \\ \vec{x}\in A}} f(\vec{x}) = (\lim_{\substack{\vec{x}\to\vec{a} \\ \vec{x}\in A}} f_1(\vec{x}), ..., \lim_{\substack{\vec{x}\to\vec{a} \\ \vec{x}\in A}} f_m(\vec{x})) \tag{2.4.1}$$

Demostración. La demostración es casi idéntica a la del teorema anterior. □

Nota 2.4.1. Dados m e.v.n. $\vec{F}_1, ..., \vec{F}_m$, los dos teoremas anteriores se generalizan fácilmente al caso en que f toma sus valores en el e.v.n. producto $\vec{F} := \vec{F}_1 \times ... \times \vec{F}_m$ (ver Ejemplo 1.2.5 y Nota 1.2.4).

2.5 Funciones continuas definidas en un compacto

Teorema 2.5.1. *Si f es una función continua definida en una parte compacta A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, entonces $f(A)$ es una parte compacta de $\vec{F}$.*

Demostración. Sea $\{\vec{y}_k\}$ una sucesión en $f(A)$. Existe entonces para cada $\vec{y}_k$ un $\vec{x}_k \in A$ tal que $\vec{y}_k = f(\vec{x}_k)$. Como A es compacto, existe una subsucesión $\{\vec{x}_{\alpha(k)}\}$ convergente a un elemento $\vec{x}_0 \in A$ y, como f es continua en $\vec{x}_0$, la subsucesión $\{\vec{y}_{\alpha(k)}\}$ será convergente a $\vec{y}_0 := f(\vec{x}_0) \in f(A)$. Esto muestra que $f(A)$ es compacto. □

Teorema 2.5.2. *Si f es una función continua definida en una parte compacta A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$, entonces f alcanza su máximo y su mínimo sobre A. Esto significa que existen $\vec{a}_m, \vec{a}_M \in A$ tales que*

$$f(\vec{a}_m) \leq f(\vec{x}) \leq f(\vec{a}_M) \quad \text{para todo} \;\; \vec{x} \in A. \tag{2.5.1}$$

Demostración. De acuerdo al teorema anterior, el conjunto $f(A)$ es compacto en $\mathbb{R}$ y, de acuerdo al Teorema 1.5.2, $f(A)$ será cerrado y acotado. Por ser acotado podemos definir los reales $M := \sup\{f(A)\}$ y $m := \inf\{f(A)\}$ y, por ser cerrado $M, m \in f(A)$. Existirán entonces elementos $\vec{a}_M, \vec{a}_m \in A$ tales que $f(\vec{a}_M) = M$ y $f(\vec{a}_m) = m$. Es evidente entonces que $\vec{a}_m$ y $\vec{a}_M$ verifican (2.5.1). □

Nota 2.5.1. El hecho que toda función continua alcance su máximo y su mínimo en un compacto, constituye una propiedad importante. Numerosos modelos matemáticos usados por las ciencias de la ingeniería y la física se plantean en términos de maximización o minimización de funciones. Estar entonces seguros de la existencia de soluciones para estos modelos, es crucial.

Teorema 2.5.3. *Sea f una función continua definida en una parte compacta A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}_+^*$, (conjunto de los reales > 0). Entonces existe $\alpha > 0$ tal que $f(\vec{x}) \geq \alpha$ para todo $\vec{x} \in A$.*

Demostración. De acuerdo al teorema anterior, existirá $\vec{a}_m \in A$ tal que $f(\vec{x}) \geq f(\vec{a}_m)$ para todo $\vec{x} \in A$. Como por hipótesis $f(\vec{a}_m) > 0$, si definimos $\alpha := f(\vec{a}_m)$ obtenemos que $f(\vec{x}) \geq \alpha$ para todo $\vec{x} \in A$. □

Teorema 2.5.4. *Sea A una parte compacta de un e.v.n. $\vec{E}$, y d_A la función distancia al conjunto A definida por (2.2.3). Para todo $\vec{a} \in \vec{E}$ existirá $\vec{p} \in A$ tal que*

$$d_A(\vec{a}) = \|\vec{a} - \vec{p}\| \tag{2.5.2}$$

Este elemento $\vec{p} \in A$ se llama proyección de $\vec{a}$ sobre A.

Demostración. Como $\|\vec{a} - \cdot\| : \vec{E} \to \mathbb{R}$ es la composición de las funciones $\vec{x} \in \vec{E} \to \vec{a} - \vec{x} \in \vec{E}$ y $\| \cdot \| : \vec{E} \to \mathbb{R}$ y, como estas dos funciones son continuas, el Teorema 2.2.2 nos dice que $\|\vec{a} - \cdot\|$ será también continua. Del Teorema 2.5.2 concluímos entonces que esta función alcanza su mínimo sobre A en un punto $\vec{p} \in A$ y se tendrá por lo tanto $d_A(\vec{a}) = \|\vec{a} - \vec{p}\|$. $\square$

Nota 2.5.2. Es importante tener claro que la proyección depende de la norma. Así por ejemplo, en $\vec{E} = \mathbb{R}^2$, el conjunto de las proyecciones del origen sobre el triángulo de vértices (1,2)(1,-2) y (2,0) será:

i) usando la norma $\| \cdot \|_\infty$, igual al conjunto $\{(1,t) : t \in [-1,1]\}$.

ii) usando la norma $\| \cdot \|_2$, igual al conjunto $\{(1,0)\}$.

Nota 2.5.3. Si en el teorema anterior consideramos $\vec{E} = \mathbb{R}^n$, podemos hacer menos hipótesis sobre el conjunto A para asegurar la existencia de la proyección. Basta con suponer que A es cerrado (recordemos que de acuerdo al Teorema 1.5.2 todo conjunto compacto es cerrado y acotado). En efecto, dado $r > d_A(\vec{a})$ es fácil ver que $d_A(\vec{a}) = d_{A\cap B(\vec{a},r)}(\vec{a})$. Como el conjunto $A \cap B(\vec{a}, r)$ es cerrado y acotado, de acuerdo al Teorema 1.5.3 será compacto. Aplicando el teorema anterior a la función $d_{A\cap B(\vec{a},r)}$, concluimos que existe $\vec{p} \in A$ tal que $d_A(\vec{a}) = \|\vec{a} - \vec{p}\|$.

Teorema 2.5.5. *Sean A, B dos conjuntos cerrados disjuntos en un e.v.n. $\vec{E}$. Definamos la distancia entre A y B por*

$$\delta(A, B) := \inf\{d_A(\vec{x}) : \vec{x} \in B\} \tag{2.5.3}$$

donde d_A es la función distancia al conjunto A definida por (2.2.3)*. Si B es compacto, entonces existirá $\vec{b} \in B$ tal que $\delta(A, B) = d_A(\vec{b}) > 0$. Si suponemos además que A es compacto, existirá también $\vec{a} \in A$ tal que*

$$\delta(A, B) = \|\vec{b} - \vec{a}\| \tag{2.5.4}$$

Demostración. Del Teorema 2.2.5 sabemos que d_A es una función continua y, como B es compacto, del Teorema 2.5.2 concluímos que d_A alcanza su mínimo sobre B en un punto $\vec{b} \in B$. Se tiene así que $\delta(A, B) = d_A(\vec{b})$.

La desigualdad $d_A(\vec{b}) > 0$ es consecuencia del hecho que $\vec{b} \notin \bar{A}$ (es fácil probar que todo conjunto A en un e.v.n. verifica la equivalencia $d_A(\vec{x}) = 0 \Leftrightarrow \vec{x} \in \bar{A}$).

Supongamos ahora que A es compacto. De acuerdo al teorema anterior existe $\vec{a} \in A$ tal que $d_A(\vec{b}) = \|\vec{b} - \vec{a}\|$, que es lo que queríamos probar. $\square$

2.6 Continuidad uniforme y Lipschitzianidad

Definición 2.6.1. Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$ se dice uniformemente continua en A si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\vec{x}, \vec{y} \in A,\ \|\vec{x} - \vec{y}\| \leq \delta \Rightarrow \|f(\vec{x}) - f(\vec{y})\| \leq \varepsilon \tag{2.6.1}$$

Nota 2.6.1. Es evidente que una función $f : A \to \vec{F}$ uniformemente continua en A, será continua en A. Por el contrario, si f es continua en A, no será necesariamente uniformemente continua en A. Un ejemplo simple que muestra este hecho es el de la función $f(x) = x^2$ de $\mathbb{R}$ en $\mathbb{R}$, que siendo continua no es uniformemente continua en $\mathbb{R}$. En efecto, para cualquier $\delta > 0$, basta con definir $x = \frac{1}{\delta} + \delta$, e $y = \frac{1}{\delta}$ y constatar que $|x - y| \leq \delta$ y que $|x^2 - y^2| = 2 + \delta^2 > 2$.

Teorema 2.6.1. *Si f es una función continua definida en una parte compacta A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, entonces ella es uniformemente continua.*

Demostración. Razonemos por contradicción. Si f no es uniformemente continua en A, existirá $\varepsilon > 0$ tal que para todo $\delta > 0$ existen $\vec{x}_\delta$, $\vec{y}_\delta \in A$ que verifican $\|\vec{x}_\delta - \vec{y}_\delta\| \leq \delta$ y $\|f(\vec{x}_\delta) - f(\vec{y}_\delta)\| > \varepsilon$. Aplicando éste hecho sucesivamente para $\delta = 1, 1/2, .., 1/k, ...$ obtenemos en A un par de sucesiones $\{\vec{x}_k\}$ e $\{\vec{y}_k\}$ que verifican $\|\vec{x}_k - \vec{y}_k\| \leq 1/k$ y $\|f(\vec{x}_k) - f(\vec{y}_k)\| > \varepsilon$. Como A es compacto, existirá una subsucesión $\{x_{\alpha(k)}\}$ convergente a $\vec{a} \in A$. Por otra parte, como $\|\vec{y}_{\alpha(k)} - \vec{a}\| \leq \|\vec{y}_{\alpha(k)} - \vec{x}_{\alpha(k)}\| + \|\vec{x}_{\alpha(k)} - \vec{a}\| \leq 1/k + \|\vec{x}_{\alpha(k)} - \vec{a}\|$, vemos que $\|y_{\alpha(k)} - \vec{a}\| \to 0$ lo que implica, de acuerdo al Lema 1.4.1, que la subsucesión $\{\vec{y}_{\alpha(k)}\}$ también converge a $\vec{a}$. Finalmente, la desigualdad $\varepsilon < \|f(\vec{x}_{\alpha(k)}) - f(\vec{y}_{\alpha(k)})\| \leq \|f(\vec{x}_{\alpha(k)}) - f(\vec{a})\| + \|f(\vec{a}) - f(\vec{y}_{\alpha(k)})\|$ contradice (de acuerdo al Teorema 2.3.5) la continuidad de f en $\vec{a}$. $\square$

Definición 2.6.2. Una función f definida en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n $\vec{F}$ se dirá Lipschitziana en A si existe una constante L, llamada constante de Lipschitz, tal que

$$\|f(\vec{x}) - f(\vec{y})\| \leq L\|\vec{x} - \vec{y}\| \quad \text{para todo } \ \vec{x}, \vec{y} \in A.$$

Nota 2.6.2. Es evidente que si f es una función Lipschitziana en A, ella sigue siendo Lipschitziana si cambiamos la norma de $\vec{E}$ y la de $\vec{F}$ por otras equivalentes. También es evidente que una función Lipschitziana es uniformemente continua.

2.7 El e.v.n. de las funciones lineales continuas

Teorema 2.7.1. *Una función lineal l de un e.v.n. $\vec{E}$ en un e.v.n. $\vec{F}$ es continua si y solo si ella es continua en $\vec{0} \in \vec{E}$ y esto se tiene, si y solo si existe una constante L tal que*

$$\|l(\vec{x})\| \leq L\|\vec{x}\| \quad \textit{para todo } \vec{x} \in \vec{E}. \tag{2.7.1}$$

Demostración. Si l es continua, ella es en particular continua en $\vec{0}$.

Demostremos ahora que si l es continua en $\vec{0}$, entonces se tiene la desigualdad (2.7.1). Dado $\varepsilon = 1$ debe existir $\delta > 0$ tal que

$$\|\vec{z}\| \leq \delta \Rightarrow \|l(\vec{z})\| \leq 1$$

y como $\vec{z} := \frac{\delta}{\|\vec{x}\|}\vec{x}$ verifica $\|\vec{z}\| \leq \delta$ para todo $\vec{x} \neq \vec{0}$, se tendrá que $\|l(\vec{x})\| \leq \frac{1}{\delta}\|\vec{x}\|$ para todo $\vec{x} \in \vec{E}$, que equivale a (2.7.1) con $L := \frac{1}{\delta}$.

Para terminar demostremos que (2.7.1) implica que l es continua, como

$$\|l(\vec{x}) - l(\vec{x}')\| = \|l(\vec{x} - \vec{x}')\| \leq L\|\vec{x} - \vec{x}'\| \quad \text{para todo } \vec{x}, \vec{x}' \in \vec{E}$$

concluímos que l es lipschitziana en $\vec{E}$ y a fortiori continua. $\square$

Nota 2.7.1. La propiedad (2.7.1) constituye una caracterización fundamental de la continuidad de las funciones lineales que nos permitirá, entre otras cosas, definir una norma en el e.v. de las funciones lineales continuas. Ella nos muestra también que toda función lineal continua es Lipschitziana.

Teorema 2.7.2. *En el e.v. $\mathcal{L}(\vec{E}, \vec{F})$ de las funciones lineales continuas del e.v.n. $\vec{E}$ en el e.v.n. $\vec{F}$, la aplicación que a $l \in \mathcal{L}(\vec{E}, \vec{F})$ le hace corresponder*

$$\|l\| := \sup_{\vec{x} \neq \vec{0}} \frac{\|l(\vec{x})\|}{\|\vec{x}\|} \tag{2.7.2}$$

verifica las propiedades de la Definición 1.2.1 y constituye así una norma en $\mathcal{L}(\vec{E}, \vec{F})$.

Demostración. De acuerdo al teorema anterior, la continuidad de l implica la existencia de una constante L que verifica (2.7.1), lo que muestra que el supremo en (2.7.2) es finito (el cuociente está acotado superiormente por L). Demostremos ahora cada una de las tres propiedades de la Definición 1.2.1. i) Es evidente. ii) Dado $\lambda \in \mathbb{R}$ y $l \in \mathcal{L}(\vec{E}, \vec{F})$ se tiene

$$\|\lambda l\| = \sup \frac{\|\lambda l(\vec{x})\|}{\|\vec{x}\|} = \sup |\lambda| \frac{\|l(\vec{x})\|}{\|\vec{x}\|} = |\lambda| \sup \frac{\|l(\vec{x})\|}{\|\vec{x}\|} = |\lambda| \|l\|.$$

iii) Dados $l,\ m\ \in \mathcal{L}(\vec{E}, \vec{F})$ se tiene

$$\|l+m\| = \sup \frac{\|(l+m)(\vec{x})\|}{\|\vec{x}\|} \leq \sup \frac{\|l(\vec{x})\| + \|m(\vec{x})\|}{\|\vec{x}\|}$$

$$\leq \sup \frac{\|l(\vec{x})\|}{\|\vec{x}\|} + \sup \frac{\|m(\vec{x})\|}{\|\vec{x}\|} = \|l\| + \|m\|.$$

□

Nota 2.7.2. Es fácil ver que para $l \in \mathcal{L}(\vec{E}, \vec{F})$ la cantidad $\|l\|$ es la menor constante que verifica la desigualdad (2.7.1). Se tendrá entonces que

$$\|l(\vec{x})\| \leq \|l\| \|\vec{x}\| \quad \text{para todo } \vec{x} \in \vec{E} \tag{2.7.3}$$

Nota 2.7.3. $\mathcal{L}(\vec{E}, \vec{F})$ dotado de la norma (2.7.2) representa el primer e.v.n. de dimensión infinita (cuando $\vec{E}$ o $\vec{F}$ es de dimensión infinita) que estudiaremos en este curso. Más adelante estudiaremos otros.

Dada la importancia que tienen en un e.v.n. las sucesiones de Cauchy, una de las primeras preguntas que debemos hacernos frente a un e.v.n. de dimensión infinita, es si será o no un espacio de Banach (ver definición 1.4.4). El teorema que sigue da una respuesta satisfactoria a esta pregunta.

Teorema 2.7.3. *Si $\vec{E}$ es un e.v.n. y $\vec{F}$ un espacio de Banach entonces $\mathcal{L}(\vec{E}, \vec{F})$ dotado de la norma* (2.7.2) *también es un espacio de Banach.*

Demostración. Sea $\{l_k\}$ una sucesión de Cauchy en $\mathcal{L}(\vec{E}, \vec{F})$. Para cada $\vec{x} \in \vec{E}$, de la desigualdad (2.7.3), aplicada a las funciones lineales $(l_k - l_j)$, se tendrá

$$\|l_k(\vec{x}) - l_j(\vec{x})\| = \|(l_k - l_j)(\vec{x})\| \leq \|l_k - l_j\| \|\vec{x}\| \text{ para todo } k, j \in \mathbb{N}$$

lo que nos permite deducir que $\{l_k(\vec{x})\}$ es una sucesión de Cauchy en $\vec{F}$ y, como $\vec{F}$ es un espacio de Banach, concluímos que $\{l_k(\vec{x})\}$ es una sucesión convergente. Definamos entonces la aplicación $l : \vec{E} \to \vec{F}$ por la relación $l(x) := \lim l_k(\vec{x})$ e intentemos demostrar que $\{l_k\}$ converge a l en el e.v.n. $\mathcal{L}(\vec{E}, \vec{F})$.

Para empezar debemos probar que $l \in \mathcal{L}(\vec{E}, \vec{F})$, es decir, que l es lineal y continua. Dados $\vec{x}, \vec{y} \in \vec{E}$ y $\lambda \in \mathbb{R}$, del Teorema 1.4.4 deducimos que

$$\begin{aligned} l(\vec{x} + \vec{y}) &= \lim l_k(\vec{x} + \vec{y}) = \lim[l_k(\vec{x}) + l_k(\vec{y})] \\ &= \lim l_k(\vec{x}) + \lim l_k(\vec{y}) = l(\vec{x}) + l(\vec{y}) \\ rl(\lambda\vec{x}) &= \lim l_k(\lambda\vec{x}) = \lim \lambda l_k(\vec{x}) = \lambda \lim l_k(\vec{x}) = \lambda l(\vec{x}) \end{aligned}$$

lo que muestra que l es lineal.

Dado $\varepsilon > 0$, por ser $\{l_k\}$ de Cauchy, debe existir $k_0 \in \mathbb{N}$ tal que $\|l_k - l_j\| \leq \varepsilon$ para todo $k, j \geq k_0$, lo que implica que $\|l_k\| = \|l_k - l_j + l_j\| \leq \varepsilon + \|l_j\|$ para todo $k, j \geq k_0$. Usando entonces (2.7.3), de las dos desigualdades anteriores deducimos que

$$\begin{aligned} \|l_k(\vec{x}) - l_j(\vec{x})\| &\leq \varepsilon\|\vec{x}\| \quad \text{para todo} \quad k, j \geq k_0 \\ &\Rightarrow \|l_k(\vec{x})\| \leq (\varepsilon + \|l_j\|)\|\vec{x}\| \quad \text{para todo} \quad k, j \geq k_0. \end{aligned}$$

Fijando $j \geq k_0$ y $\vec{x} \in \vec{E}$ en estas dos desigualdades, teniendo en cuenta la continuidad de la norma (Teorema 2.2.4 y el Teorema 2.3.1) y, tomando límite sobre k obtenemos

$$\begin{aligned} \|l(\vec{x}) - l_j(\vec{x})\| &\leq \varepsilon\|\vec{x}\| \\ &\Rightarrow \|l(\vec{x})\| \leq (\varepsilon + \|l_j\|)\|\vec{x}\| \end{aligned}$$

que serán válidas para todo $\vec{x} \in \vec{E}$ y todo $j \geq k_0$.

Usando el Teorema 2.7.1, la segunda de estas desigualdades nos muestra que la función lineal l es continua.

Finalmente, de la definición de la norma en $\mathcal{L}(\vec{E}, \vec{F})$, la primera de las dos desigualdades anteriores nos muestra que

$$\|l - l_j\| \leq \varepsilon \quad \text{para todo} \quad j \geq k_0$$

lo que significa que $\{l_k\}$ converge a l en $\mathcal{L}(\vec{E}, \vec{F})$. □

Teorema 2.7.4. *Toda función lineal l de $\mathbb{R}^n$ en un e.v.n. $\vec{F}$ es continua.*

Demostración. Como todas las normas en $\mathbb{R}^n$ son equivalentes, de acuerdo a la Nota 2.2.1 podemos usar cualquiera, en particular, la norma $\|\cdot\|_\infty$. Sea $\vec{e}_1, ..., \vec{e}_n$ la base canónica de $\mathbb{R}^n$. Para todo $\vec{x} \in \mathbb{R}^n$ se tendrá

$$\|l(\vec{x})\| = \left\| l(\sum_{i=1}^n x_i\vec{e}_i) \right\| = \left\| \sum_{i=1}^n x_i l(\vec{e}_i) \right\| \leq \sum_{i=1}^n |x_i|\|l(\vec{e}_i)\| \leq \left[\sum_{i=1}^n \|l(\vec{e}_i)\| \right] \|\vec{x}\|_\infty.$$

Esto muestra que l verifica la desigualdad (2.7.1) con $L = \sum_{i=1}^{n} \|l(\vec{e}_i)\|$ y, de acuerdo al Teorema 2.7.1, será continua. $\square$

Nota 2.7.4. Más adelante veremos que existen aplicaciones lineales que no son continuas. Naturalmente, esto se tendrá cuando el dominio es un e.v.n. de dimensión infinita.

2.8 Teorema del punto fijo

Definición 2.8.1. Una aplicación f de un e.v.n. $\vec{E}$ en un e.v.n. $\vec{F}$ se dirá contractante en una parte A de $\vec{E}$, si es Lipschitziana en A con constante de Lipchitz $L < 1$.

Teorema 2.8.1. *Si $f : A \longrightarrow A$ es contractante en una parte cerrada A de un espacio de Banach $\vec{E}$ (ver Definición 1.4.4), entonces existe un único $\vec{a} \in A$ tal que $f(\vec{a}) = \vec{a}$. El elemento $\vec{a}$ se llama punto fijo de f en A.*

Demostración. Demostremos primero la unicidad. Si $f(\vec{a}) = \vec{a}$ y $f(\vec{b}) = \vec{b}$ entonces

$$\|\vec{a} - \vec{b}\| = \|f(\vec{a}) - f(\vec{b})\| \leq L\|\vec{a} - \vec{b}\|.$$

Como $L < 1$, vemos que esta desigualdad implica que $\vec{a} = \vec{b}$.

Para demostrar la existencia vamos a construir una sucesión $\vec{x}_k$ convergente a un punto fijo de f. A partir de $\vec{x}_1 \in A$, construímos $\{\vec{x}_k\}$ en A mediante la fórmula de recurrencia $\vec{x}_k = f(\vec{x}_{k-1})$. Esta sucesión verifica

$$\begin{aligned}
\|\vec{x}_3 - \vec{x}_2\| &= \|f(\vec{x}_2) - f(\vec{x}_1)\| \leq L\|\vec{x}_2 - \vec{x}_1\| \\
\|\vec{x}_4 - \vec{x}_3\| &= \|f(\vec{x}_3) - f(\vec{x}_2)\| \leq L\|\vec{x}_3 - \vec{x}_2\| \leq L^2\|\vec{x}_2 - \vec{x}_1\| \\
&\vdots \\
\|\vec{x}_{k+1} - \vec{x}_k\| &= \|f(\vec{x}_k) - f(\vec{x}_{k-1})\| \leq L\|\vec{x}_k - \vec{x}_{k-1}\| \leq ... \leq L^{k-1}\|\vec{x}_2 - \vec{x}_1\|
\end{aligned}$$

por lo tanto para todo $k, p \in \mathbb{N}$ se tiene

$$\begin{aligned}
\|\vec{x}_{k+p} - \vec{x}_k\| &\leq \|\vec{x}_{k+p} - \vec{x}_{k+p-1}\| + \|\vec{x}_{k+p-1} - \vec{x}_{k+p-2}\| + ... + \|\vec{x}_{k+1} - \vec{x}_k\| \\
&\leq [L^{k+p-2} + L^{k+p-3} + ... + L^{k-1}]\|\vec{x}_2 - \vec{x}_1\| \\
&\leq \frac{L^{k-1}}{1-L}\|\vec{x}_2 - \vec{x}_1\|
\end{aligned}$$

Como $L^{k-1} \to 0$, es fácil verificar que $\{\vec{x}_k\}$ es una sucesión de Cauchy. En efecto, dado $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que $\frac{L^{k-1}\|\vec{x}_2 - \vec{x}_1\|}{1-L} \leq \varepsilon$ para todo $k \geq k_0$, por lo tanto $\|\vec{x}_{k+p} - \vec{x}_k\| \leq \varepsilon$ para todo $p \geq 0$ y todo $k \geq k_0$ y definiendo $j = k + p$ obtenemos $\|\vec{x}_j - \vec{x}_k\| \leq \varepsilon$ para todo $j, k \geq k_0$, que corresponde exactamente a la definición de sucesión de Cauchy. Como $\vec{E}$ es un espacio de Banach concluímos que $\{\vec{x}_k\}$ es convergente y, por ser A cerrado el límite $\vec{a}$ de esta sucesión esta en A y verifica

$$\vec{a} = \lim \vec{x}_{k+1} = \lim f(\vec{x}_k) = f(\vec{a})$$

donde la última igualdad se desprende de (2.3.8) por ser f continua en $\vec{a}$. Esto muestra que $\vec{a}$ es punto fijo de la función f. $\square$

2.9 Ejercicios

1. Sea f una función continua definida en un e.v.n. $\vec{E}$ con valores en un e.v.n $\vec{F}$. Demuestre que

   a) Si $B \subset \vec{F}$ es cerrado, entonces $f^{-1}[B] \subset \vec{E}$ también es cerrado.

   b) Si $B \subset \vec{F}$ es abierto, entonces $f^{-1}[B] \subset \vec{E}$ también es abierto.

   c) $f(\bar{A}) \subset \overline{f(A)}$ para todo $A \subset \vec{E}$.

   d) $f^{-1}[\operatorname{int} B] \subset \operatorname{int}(f^{-1}[B])$ para todo $B \subset \vec{F}$.

   e) $\overline{f^{-1}[B]} \subset f^{-1}[\bar{B}]$ para todo $B \subset \vec{F}$.

2. Demuestre que la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(\vec{x}) = \frac{x_1 x_2^2}{x_1^2 + x_2^4}$ si $\vec{x} \neq \vec{0}$ y $f(\vec{0}) = 0$ es continua en todo punto de $\mathbb{R}^2$ salvo en $\vec{0}$.

3. Demuestre que la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(\vec{x}) = \frac{x_1 x_2}{\|\vec{x}\|_2}$ si $\vec{x} \neq \vec{0}$ y $f(\vec{0}) = 0$, es continua.

4. Demuestre que la función $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(\vec{x}) = \frac{1}{\|\vec{x}\|_2}(x_1^2, x_2^2)$ si $\vec{x} \neq \vec{0}$ y $f(\vec{0}) = \vec{0}$, es continua.

5. Considere la función $f : \mathbb{R}^2 \setminus \{\vec{0}\} \to \mathbb{R}$ definida por $f(\vec{x}) = \frac{x_1 x_2}{\|\vec{x}\|_2^2}$. Demuestre que $\lim\limits_{x_1 \to 0}(\lim\limits_{x_2 \to 0} f(\vec{x})) = \lim\limits_{x_2 \to 0}(\lim\limits_{x_1 \to 0} f(\vec{x})) = 0$ y que $\lim\limits_{\vec{x} \to \vec{0}} f(\vec{x})$ no existe.

6. Considere la función $f : \mathbb{R}^2 \setminus \{\vec{0}\} \to \mathbb{R}$ definida por $f(\vec{x}) = \frac{x_1^2 - x_2^2}{\|\vec{x}\|_2^2}$. Demuestre que $\lim\limits_{x_2 \to 0}(\lim\limits_{x_1 \to 0} f(\vec{x})) = -1$, $\lim\limits_{x_1 \to 0}(\lim\limits_{x_2 \to 0} f(\vec{x})) = 1$ y, que $\lim\limits_{\vec{x} \to \vec{0}} f(\vec{x})$ no existe.

7. Sea f una función continua definida en un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Demuestre que el conjunto $G(f) = \{(\vec{x}, \vec{y}) \in \vec{E} \times \vec{F} : \vec{y} = f(\vec{x})\}$ es cerrado en el e.v.n. $\vec{E} \times \vec{F}$.

8. Sean f, g dos funciones continuas definidas en un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$ y, sea $A \subset \vec{E}$ tal que $f(\vec{x}) = g(\vec{x})$ para todo $\vec{x} \in A$. Demuestre que $f(\vec{x}) = g(\vec{x})$ para todo $\vec{x} \in \bar{A}$.

9. Sean $\vec{E}_1, ..., \vec{E}_n$ espacios vectoriales normados y sea p_i la función del e.v.n. $\vec{E} = \vec{E}_1 \times ... \times \vec{E}_n$ en $\vec{E}_i$, que a cada $\vec{x} \in \vec{E}$ le asocia su componente $\vec{x}_i \in \vec{E}_i$. Demuestre que esta función es continua.

10. Sea f una función continua definida en un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Demuestre que

    a) $S_\lambda = \{\vec{x} \in \vec{E} : f(\vec{x}) \leq \lambda\}$ es un conjunto cerrado para todo $\lambda \in \mathbb{R}$.

    b) $I_\lambda = \{\vec{x} \in \vec{E} : f(\vec{x}) = \lambda\}$ es un conjunto cerrado para todo $\lambda \in \mathbb{R}$.

    c) $A_\lambda = \{\vec{x} \in \vec{E} : f(\vec{x}) < \lambda\}$ es un conjunto abierto para todo $\lambda \in \mathbb{R}$.

11. Sea f una función continua y biyectiva, de un compacto A de un e.v.n. $\vec{E}$ en un conjunto B de un e.v.n. $\vec{F}$. Demuestre que la función $f^{-1} : B \to A$ es continua.

12. Sea f una función continua definida en $\mathbb{R}^n$ con valores en $\mathbb{R}$. Si f verifica la propiedad: para todo $L \in \mathbb{R}_+$ existe $k_0 \in \mathbb{N}$ tal que $\|\vec{x}\| \geq k_0 \;\Rightarrow\; f(\vec{x}) \geq L$, demuestre que el conjunto S_λ definido en el Ejercicio 10 es un compacto y que la función f alcanza su mínimo en $\mathbb{R}^n$.

13. Dado un e.v.n. $\vec{E}$ y dos puntos $\vec{a}, \vec{b} \in \vec{E}$ demuestre que la función $\lambda \in \mathbb{R} \mapsto \vec{a} + \lambda(\vec{b} - \vec{a}) \in \vec{E}$ es continua. Con esto, demuestre que el intervalo $[\vec{a}, \vec{b}] = \{\vec{a} + \lambda(\vec{b} - \vec{a}) \in \vec{E} : \lambda \in [0, 1]\}$ es compacto.

14. Demuestre que la función $f : \mathbb{R}_+ \to \mathbb{R}$ definida por $f(x) = \sqrt{x}$ es uniformemente continua pero no es Lipschitziana.

15. Demuestre que toda función Lipschitziana es uniformemente continua.

16. Sea f una función uniformemente continua en una parte A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Demuestre que si $\{\vec{x}_k\}$ es una sucesión de Cauchy en A, entonces $\{f(\vec{x}_k)\}$ es una sucesión de Cauchy en $\vec{F}$. Demuestre que esta propiedad no se tiene necesariamente si f no es uniformemente continua en A.

17. Sea l una función lineal definida en un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Demostrar que l es continua si y solo si el conjunto $N(l) := \{\vec{x} \in \vec{E} : l(\vec{x}) = 0\}$ es cerrado en $\vec{E}$.

18. Sea l una función lineal definida en el e.v.n. $(\mathcal{C}([a,b],\mathbb{R}),\|\cdot\|_1)$ con valores en $\mathbb{R}$, por $l(f)=\int\limits_a^b f(x)dx$. Demuestre que l es continua.

19. Sea l una función lineal definida en el e.v.n. $(\mathcal{C}([a,b],\mathbb{R}),\|\cdot\|_\infty)$ con valores en $\mathbb{R}$, por $l(f)=\int\limits_a^b g(x)f(x)dx$, donde g es una función fija en $\mathcal{C}([a,b],\mathbb{R})$. Demuestre que l es continua.

20. Sea l una función lineal definida en el e.v.n. $(\mathcal{C}([0,1],\mathbb{R}),\|\cdot\|_1)$ con valores en $\mathbb{R}$, por $l(f)=f(0)$. Considere en $\mathcal{C}([0,1],\mathbb{R})$ la sucesión $\{f_n\}$ definida por $f_n(x)=1-nx$ si $x\in[0,1/n]$, $f(x)=0$ si $x\in[1/n,1]$. Usando esta sucesión demuestre que l no es continua.

21. Usando el teorema del punto fijo de Banach demuestre que dado $\alpha>0$, la sucesión $\{x_k\}$ construída por la fórmula recurrente $x_{k+1}=\frac{1}{2}\left(x_k+\frac{\alpha}{x_k}\right)$, a partir de un valor $x_1\geq\sqrt{\alpha}$, converge a $\sqrt{\alpha}$.

CAPÍTULO 3

ESPACIOS DE FUNCIONES

3.1 Introducción

Todos los e.v. de dimensión infinita que intervienen en los modelos matemáticos de la ingeniería y de la física, son espacios de funciones: el e.v. de las funciones acotadas, el e.v. de las funciones continuas, el e.v. de las funciones diferenciables, etc.

En este capítulo veremos las propiedades más elementales de este tipo de espacios vectoriales, dotados de la norma $\|\cdot\|_\infty$ que definimos en (1.2.13).

Los tres resultados importantes del capítulo están en el Teorema 3.2.1 el cual muestra que el e.v. $\mathcal{A}(A,\vec{F})$ (de las funciones acotadas de A en $\vec{F}$) dotado de la norma $\|\cdot\|_\infty$ es un espacio de Banach, los Teoremas 3.4.1 y 3.4.2 que establecen la cerradura del subespacio vectorial $\mathcal{C}(A,\vec{F})$ (de las funciones continuas de A en $\vec{F}$) en el espacio $\mathcal{A}(A,\vec{F})$ dotado de la norma $\|\cdot\|_\infty$ y, finalmente en el Teorema 3.6.1 que nos dice que para toda función en $\mathcal{C}([a,b],\mathbb{R})$ se puede construir una sucesión de polinomios convergente a esa función para la norma $\|\cdot\|_\infty$.

3.2 Espacio vectorial normado de las funciones acotadas

Definición 3.2.1. Dado un conjunto A y un e.v.n. $\vec{F}$ denotaremos por $\mathcal{A}(A,\vec{F})$ al espacio vectorial de las funciones acotadas de A en $\vec{F}$ ($f \in \mathcal{A}(A,\vec{F})$ si existe r tal que $\|f(x)\| \leq r$

para todo $x \in A$). A este espacio lo dotamos de la norma $\|\cdot\|_\infty$ definida por

$$\|f\|_\infty := \sup_{x\in A} \|f(x)\| \tag{3.2.1}$$

Nota 3.2.1. Como decíamos cuando estudiábamos el e.v.n. $\mathcal{L}(\vec{E},\vec{F})$, dada la importancia que tienen en un e.v.n. las sucesiones de Cauchy, una pregunta fundamental que se debe hacer frente a un e.v.n. de dimensión infinita, es si será o no un espacio de Banach (ver Definición 1.4.4). El teorema que sigue da una respuesta satisfactoria a esta pregunta.

Teorema 3.2.1. *Si $\vec{F}$ es un espacio de Banach, entonces $\mathcal{A}(A,\vec{F})$ dotado de la norma $\|\cdot\|_\infty$ también lo es.*

Demostración. Sea $\{f_k\}$ una sucesión de Cauchy en $\mathcal{A}(A,\vec{F})$. Como para todo $x \in A$ se tiene la desigualdad $\|f_k(x) - f_j(x)\| \leq \|f_k - f_j\|_\infty$, es evidente que para cada $x \in A$ la sucesión $\{f_k(x)\}$ es de Cauchy en $\vec{F}$ y, como por hipótesis $\vec{F}$ es un espacio de Banach, ella será convergente. Definamos entonces la aplicación $f : A \to \vec{F}$ por la relación $f(x) = \lim f_k(x)$ e intentemos demostrar que $\{f_k\}$ converge a f en el e.v.n. $(\mathcal{A}(A,\vec{F}), \|\cdot\|_\infty)$. Lo primero que debemos hacer es demostrar que $f \in \mathcal{A}(A,\vec{F})$. Dado $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que $\|f_k - f_j\|_\infty \leq \varepsilon$ para todo $k, j \geq k_0$, lo que equivale a decir que:

$$\|f_k(x) - f_j(x)\| \leq \varepsilon \;\text{ para todo }\; x \in A \text{ y todo }\; k, j \geq k_0 \tag{*}$$

Como para k y x fijos, la función $y \in \vec{F} \to \|f_k(x) - y\|$ es continua, del Teorema 2.3.5 obtenemos

$$\lim_j \|f_k(x) - f_j(x)\| = \|f_k(x) - \lim_j f_j(x)\| = \|f_k(x) - f(x)\|.$$

Tomando límite sobre j en (*), la igualdad anterior nos permite concluir que

$$\|f_k(x) - f(x)\| \leq \varepsilon \;\text{ para todo } x \in A \text{ y todo } k \geq k_0 \tag{**}$$

y como $f_k \in \mathcal{A}(A,\vec{F})$, para todo $x \in A$ tendremos

$$\|f(x)\| - \|f_k\|_\infty \leq \|f(x)\| - \|f_k(x)\| \leq \|f(x) - f_k(x)\| \leq \varepsilon,$$

es decir,

$$\|f(x)\| \leq \varepsilon + \|f_k\|_\infty \text{ para todo }\; x \in A$$

lo que muestra que f es acotada.

El resto de la demostración esta ya practicamente hecha, en efecto, (**) es equivalente a

$$\|f_k - f\|_\infty \leq \varepsilon \;\text{ para todo } k \geq k_0$$

lo que muestra que $\{f_k\}$ converge a f. $\square$

Nota 3.2.2. Un espacio de funciones particularmente importante es el de las funciones continuas definidas en una parte compacta A de un e.v.n. $\vec{E}$, con valores en $\mathbb{R}^m$. Este espacio de funciones lo denotaremos $\mathcal{C}(A, \mathbb{R}^m)$. Del Teorema 2.5.2 concluímos que $\mathcal{C}(A, \mathbb{R}^m)$ es un subespacio vectorial del e.v. $\mathcal{A}(A, \mathbb{R}^m)$. Usando entonces el teorema anterior y el Teorema 3.4.1 que veremos más adelante, concluiremos que $\mathcal{C}(A, \mathbb{R}^m)$, dotado de la norma $\|\cdot\|_\infty$, es también un espacio de Banach.

3.3 Convergencia uniforme y convergencia simple de una sucesión de funciones

Definición 3.3.1. La convergencia de una sucesión de funciones en el e.v.n. $\mathcal{A}(A, \vec{F})$, definida en la sección anterior, se llama convergencia uniforme de la sucesión. Así entonces, decimos que $\{f_k\}$ converge uniformemente a la función f, cuando $\|f_k - f\|_\infty \to 0$.

Nota 3.3.1. El propósito de esta sección es ver qué relación existe entre la convergencia de una sucesión $\{f_k\}$ en $\mathcal{A}(A, \vec{F})$ dotado de la norma $\|\cdot\|_\infty$ (que hemos llamado convergencia uniforme de $\{f_k\}$) y, la convergencia en el e.v.n $\vec{F}$ de la sucesión $\{f_k(x)\}$ donde x está fijo (que llamaremos convergencia simple o puntual de $\{f_k\}$).

Definición 3.3.2. Dada una sucesión $\{f_k\}$ de funciones definidas en un conjunto A con valores en un e.v.n. $\vec{F}$, diremos que $\{f_k\}$ converge simplemente o puntualmente a una función $f : A \to \vec{F}$ si para todo $x \in A$ la sucesión $\{f_k(x)\}$ converge a $f(x)$ en $\vec{F}$.

Nota 3.3.2. Dada una sucesión $\{f_k\}$ en $\mathcal{A}(A, \vec{F})$ convergente uniformemente a f, puesto que para todo $k \in \mathbb{N}$ y todo $x \in A$ se tiene que $\|f_k(x) - f(x)\| \leq \|f_k - f\|_\infty$, es evidente que $\{f_k\}$ converge simplemente a f. Decimos entonces que la convergencia uniforme implica la convergencia simple. Lo contrario no es verdadero. Los dos ejemplos que siguen ilustran bien este hecho negativo y, permiten al mismo tiempo comprender mejor la convergencia uniforme.

Ejemplo 3.3.1. En el e.v. $\mathcal{A}(\mathbb{R}, \mathbb{R})$ definamos la sucesión $\{f_k\}$ por

$$f_k(x) = \frac{1}{1 + (x-k)^2} \tag{3.3.1}$$

Entonces, como para todo $x \in \mathbb{R}$ se tiene que $\lim f_k(x) = 0$ concluímos que la sucesión $\{f_k\}$ converge simplemente a la función nula $f(\cdot) = 0$. En la nota anterior veíamos que la convergencia uniforme de una sucesión de funciones implica su convergencia simple, por lo tanto si $\{f_k\}$ fuera uniformemente convergente debería serlo a la función nula. Veamos si esto es cierto, para lo cual hay que estudiar la sucesión $\|f_k - f\|_\infty$ (donde f es la función nula). Como

$$\|f_k - f\|_\infty = \sup_{x\in\mathbb{R}} |f_k(x)| = 1 \qquad \text{para todo} \quad k \in \mathbb{N}$$

concluímos que $\{f_k\}$ no converge uniformemente.

Ejemplo 3.3.2. En el e.v. $\mathcal{A}(\mathbb{R},\mathbb{R})$ definamos la sucesión $\{f_k\}$ por

$$f_k(x) = \begin{cases} k^2x & \text{si} \quad x \in [0, 1/k] \\ -k^2x + 2k & \text{si} \quad x \in [1/k, 2/k] \\ 0 & \text{si} \quad x \notin [0, 2/k]. \end{cases} \tag{3.3.2}$$

Como para todo $x \in \mathbb{R}$ se tiene que $\lim_k f_k(x) = 0$, concluímos que la sucesión $\{f_k\}$ converge simplemente a la función nula. Siguiendo entonces el mismo razonamiento del ejemplo anterior para ver si $\{f_k\}$ converge uniformemente, obtenemos

$$\|f_k - f\|_\infty = \sup_{x\in\mathbb{R}} |f_k(x)| = k$$

concluyendo así que $\{f_k\}$ no converge uniformemente.

Nota 3.3.3. Observando bien los gráficos de las funciones f_k en cada uno de los dos ejemplos anteriores, vemos que la convergencia uniforme corresponde bien a nuestra noción intuitiva de convergencia de una sucesión y que lo que hemos llamado convergencia simple no es una “verdadera convergencia”.

3.4 Continuidad del límite de una sucesión de funciones continuas

Teorema 3.4.1. *Sean $\vec{E}$ y $\vec{F}$ dos e.v.n. y sea A un subconjunto de $\vec{E}$. Si $\{f_k\}$ es una sucesión de funciones continuas en $\mathcal{A}(A, \vec{F})$ convergente uniformemente a una función f, entonces f también es continua.*

Demostración. Dado $\vec{a} \in A$, debemos demostrar que el límite f de la sucesión $\{f_k\}$ es una función continua en $\vec{a}$. Dado $\varepsilon > 0$, puesto que $\{f_k\}$ converge uniformemente a f existe $k_0 \in \mathbb{N}$ tal que

$$\|f_k(\vec{x}) - f(\vec{x})\| \leq \varepsilon/3 \qquad \text{para todo} \quad \vec{x} \in A \quad \text{y todo } k \geq k_0. \tag{*}$$

Por otra parte, puesto que f_{k_0} es una función continua en $\vec{a}$, existe $\delta > 0$ tal que

$$\|f_{k_0}(\vec{x}) - f_{k_0}(\vec{a})\| \leq \varepsilon/3 \quad \text{para todo} \quad x \in B(\vec{a}, \delta) \cap A \tag{**}$$

De las desigualdades (*) y (**) concluímos que

$$\|f(\vec{x}) - f(\vec{a})\| \leq \|f(\vec{x}) - f_{k_0}(\vec{x})\| + \|f_{k_0}(\vec{x}) - f_{k_0}(\vec{a})\| + \|f_{k_0}(\vec{a}) - f(\vec{a})\| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

para todo $x \in B(\vec{a}, \delta) \cap A$. Queda así demostrada la continuidad de f en $\vec{a}$. □

Nota 3.4.1. El Teorema 3.4.1 se enuncia usualmente diciendo que el límite uniforme de una sucesión de funciones continuas es una función continua.

Nota 3.4.2. El ejemplo que sigue muestra que el límite simple de una sucesión de funciones continuas, no es necesariamente una función continua. En ese caso, de acuerdo al teorema anterior, la sucesión no converge uniformemente.

Ejemplo 3.4.1. Consideremos en $\mathcal{A}([0,1], \mathbb{R})$ la sucesión de funciones continuas $f_k(x) = x^k$. Es fácil verificar que esta sucesión converge simplemente a la función $f(x) = 0$ si $x \in [0, 1[$ y $f(1) = 1$, que no es continua. Usando el teorema anterior deducimos también que esta sucesión $\{f_k\}$ no puede converger uniformemente.

Nota 3.4.3. El siguiente teorema repite casi exactamente lo que decíamos al final de la Nota 3.2.2.

Teorema 3.4.2. *El e.v. $\mathcal{C}(A, \vec{F})$ de las funciones continuas definidas en una parte compacta A de un e.v.n., con valores en un espacio de Banach $\vec{F}$, es un subespacio vectorial cerrado del e.v.n. $(\mathcal{A}(A, \vec{F}), \|\cdot\|_\infty)$ y, dotado de la misma norma $\|\cdot\|_\infty$, es un espacio de Banach.*

Demostración. Gracias al Teorema 2.5.2 se obtiene que $\mathcal{C}(A, \vec{F}) \subset \mathcal{A}(A, \vec{F})$ y por el Teorema 3.4.1 se concluye la cerradura de $\mathcal{C}(A, \vec{F})$ con la norma $\|\cdot\|_\infty$. Por otro lado, el Teorema 3.2.1 nos dice que $(\mathcal{A}(A, \vec{F}), \|\cdot\|_\infty)$ es un espacio de Banach y como $\mathcal{C}(A, \vec{F})$ es cerrado, se decuce que también será Banach. □

3.5 Cuatro contraejemplos interesantes

Contraejemplo 3.5.1. En la Nota 1.2.3 decíamos que todas las normas que se pueden definir en un e.v.n. de dimensión finita son equivalentes. Veremos aquí dos normas definidas en un mismo e.v. que no son equivalentes. Consideremos el e.v. $\mathcal{C}([a,b],\mathbb{R})$ de las funciones continuas en $[a,b]$ con valores en $\mathbb{R}$, la norma $\|\cdot\|_\infty$ que ya conocemos bien y, la norma:

$$\|f\|_1 := \int_a^b |f(x)|dx. \tag{3.5.1}$$

Verifiquemos que estas dos normas no son equivalentes. Si bien es cierto que

$$\|f\|_1 \leq (b-a)\|f\|_\infty \quad \text{para todo} \quad f \in \mathcal{C}([a,b],\mathbb{R}), \tag{3.5.2}$$

probaremos que no existe L tal que $\|\cdot\|_\infty \leq L\|\cdot\|_1$. Lo que es equivalente a demostrar que para todo k existe una función $f_k \in \mathcal{C}([a,b],\mathbb{R})$ que verifica $\|f_k\|_\infty \geq k\|f_k\|_1$. Si consideramos las funciones definidas en (3.3.2), vemos que $\|f_k\|_\infty = k$ y $\|f_k\|_1 = 1$ y por lo tanto, ellas verifican la última desigualdad.

Para entender cuan diferentes son estas dos normas en el e.v. $\mathcal{C}([a,b],\mathbb{R})$ será interesante ver el Contraejemplo 3.5.3 donde estudiaremos una sucesión $\{f_k\}$ que converge para la norma $\|\cdot\|_1$ y no converge para la norma $\|\cdot\|_\infty$. Otra diferencia importante entre estas dos normas la constituye el hecho que, de acuerdo al Teorema 3.4.2, $\mathcal{C}([a,b],\mathbb{R})$ dotado de la norma $\|\cdot\|_\infty$ es un espacio de Banach y, como veremos al final del contraejemplo siguiente, dotado de la norma $\|\cdot\|_1$ no es un espacio de Banach.

Contraejemplo 3.5.2. El Teorema 1.4.7 nos decía que en un e.v.n. de dimensión finita toda sucesión de Cauchy es convergente. Veremos aquí un e.v.n. (de dimensión infinita) donde hay sucesiones de Cauchy que no convergen. Consideremos el e.v. $\mathcal{C}_0([-1,1],\mathbb{R})$ de las funciones continuas de $[-1,1]$ a valores en $\mathbb{R}$, que se anulan en -1 y 1 dotado de la norma $\|\cdot\|_1$ definida en (3.5.1) y consideremos en este e.v.n. la sucesión $f_k(x) = 1 - x^{2k}$. Puesto que $\|f_k - f_j\|_1 = |\frac{2}{2k+1} - \frac{2}{2j+1}|$, dado $\varepsilon > 0$, si definimos $k_0 \geq \frac{4-\varepsilon}{2\varepsilon}$, obtenemos que $\|f_k - f_j\|_1 \leq \varepsilon$ para todo $k, j \geq k_0$ lo que muestra que $\{f_k\}$ es una sucesión de Cauchy. Para demostrar que esta sucesión no es convergente consideremos el e.v. $\mathcal{C}([-1,1],\mathbb{R})$ del cual $\mathcal{C}_0([-1,1],\mathbb{R})$ es un sub e.v. y dotémoslo de la misma norma. Es evidente entonces que si $\{f_k\}$ converge a una función f en el e.v.n. $\mathcal{C}([-1,1],\mathbb{R})$, y a una función $\bar{f}$ en el e.v.n. $\mathcal{C}_0([-1,1],\mathbb{R})$ entonces, $f = \bar{f}$. Ahora bien, es fácil verificar que $\{f_k\}$ converge en $\mathcal{C}([-1,1],\mathbb{R})$ a la función constante $f(x) = 1$, y como esta función no pertenece a $\mathcal{C}_0([-1,1],\mathbb{R})$ concluímos que en este último e.v.n. la sucesión en cuestión no converge.

El e.v. $\mathcal{C}([-1,1],\mathbb{R})$ dotado de la norma $\|\cdot\|_1$ tampoco es un espacio de Banach, como lo muestra la sucesión $\{f_k\}$ definida por $f_k(x) = \sqrt[2k+1]{x}$, que es de Cauchy y no es convergente.

Contraejemplo 3.5.3. Los Teoremas 1.5.2 y 1.5.3 de la Sección 1.5 nos mostraban que en un e.v.n. de dimensión finita los conjuntos compactos son los cerrados y acotados, por lo tanto las bolas son siempre conjuntos compactos. Veremos aquí un ejemplo de e.v.n. cuyas bolas no son compactas. El ejemplo está dado en $\mathcal{C}([0,1],\mathbb{R})$ dotado de la norma $\|\cdot\|_\infty$. Demostremos entonces que la bola con centro en la función nula y radio 1, $B(0,1)=\{f\in\mathcal{C}([0,1],\mathbb{R}):\ \|f\|_\infty\leq 1\}$, no es compacta a pesar de ser un conjunto cerrado y acotado. Para esto basta considerar la sucesión $\{f_k\}$, definida por $f_k(x)=x^k$, que está en $B(0,1)$ (en efecto, $\|f_k\|_\infty=1$) y comprobar que no tiene ninguna subsucesión convergente. Para esto, vemos que todas las subsucesiones de $\{f_k\}$ convergen simplemente a la función

$$f(x)=\begin{cases} 0 & \text{si } x\in[0,1[\\ 1 & \text{si } x=1 \end{cases}$$

que por ser discontinua, deja en evidencia de acuerdo al Teorema 3.4.1, que $\{f_k\}$ no tiene subsucesiones uniformemente convergentes.

Es interesante notar que esta sucesión de funciones es convergente a la función nula si dotamos al espacio $\mathcal{C}([0,1],\mathbb{R})$ de la norma definida en (3.5.1), en efecto $\|f_k\|_1=1/(k+1)$. Este hecho constituye otra forma de ver que estas dos normas no son equivalentes.

En general, se puede demostrar que en un e.v.n. de dimensión infinita, las bolas no son jamás compactas.

Contraejemplo 3.5.4. El Teorema 2.7.4 nos decía que toda función lineal definida en un e.v.n. de dimensión finita, con valores en un e.v.n. $\vec{F}$ es continua. Veremos aquí una función lineal definida en un e.v.n. (de dimensión infinita) con valores en $\mathbb{R}$, que no es continua. Sea $\mathcal{P}$ el e.v. de los polinonios de una variable real dotado de la norma

$$\|p\|=\max_{x\in[0,1]}|p(x)|. \tag{3.5.3}$$

No es difícil verificar que es una norma en $\mathcal{P}$. Consideremos entonces la función lineal $l:\mathcal{P}\to\mathbb{R}$ definida por $l(p)=p(3)$ y demostremos que ella no es continua en 0 (polinomio nulo). Para esto, tomemos en $\mathcal{P}$ la sucesión $p_k(x)=(\frac{x}{2})^k$. Como $\|p_k\|=2^{-k}$, vemos que se trata de una sucesión convergente a 0. Si l fuera continua, de acuerdo al Teorema 2.3.5 debería tenerse que $\lim l(p_k)=l(0)=0$, pero esto no se tiene pues $l(p_k)=p_k(3)=(\frac{3}{2})^k$, lo que muestra que l no es continua.

3.6 Teorema de Weierstrass-Stone

Weierstrass demostró el año 1885 que toda función $f\in\mathcal{C}([a,b],\mathbb{R})$ puede aproximarse tanto como se quiera por un polinomio, esto es, dado $\varepsilon>0$ existe un polinomio p tal que $\|p-f\|_\infty\leq\varepsilon$. La demostración que daremos de este importante resultado fue hecha por Berstein el año 1912 y tiene la ventaja de ser constructiva. La demostración original de Weierstrass no es constructiva, es decir, dada la función f demuestra que existe una sucesión de polinomios $\{p_k\}$ convergente uniformemente a f, sin explicitar la forma de estos polinomios.

Definición 3.6.1. Dada una función $f \in \mathcal{C}([0,1],\mathbb{R})$ se llama polinomio de Berstein de grado k asociado a la función f, al polinomio

$$b_k(x) := \sum_{i=0}^{k} f\left(\frac{i}{k}\right) p_{k,i}(x) \tag{3.6.1}$$

donde $p_{k,i}$ es el polinomio de grado k:

$$p_{k,i}(x) := \binom{k}{i} x^i (1-x)^{k-i} \tag{3.6.2}$$

Lema 3.6.1. *Para todo $k \in \mathbb{N}$ los polinomios $p_{k,i}$ definidos por* (3.6.2) *verifican*

(i) $\sum\limits_{i=0}^{k} p_{k,i}(x) = 1$

(ii) $\sum\limits_{i=0}^{k} \frac{i}{k} p_{k,i}(x) = x$

(iii) $\sum\limits_{i=0}^{k} (\frac{i}{k} - x)^2 p_{k,i}(x) = \frac{x(1-x)}{k}$.

Demostración.

(i) $\sum\limits_{i=0}^{k} p_{k,i}(x) = [x + (1-x)]^k = 1$

(ii) $\sum\limits_{i=0}^{k} \frac{i}{k} p_{k,i}(x) = x \sum\limits_{i=1}^{k} \binom{k-1}{i-1} x^{i-1}(1-x)^{(k-1)-(i-1)}$
$= x \sum\limits_{j=0}^{k-1} \binom{k-1}{j} x^j (1-x)^{(k-1)-j} = x[x + (1-x)]^{k-1} = x$

(iii) $\sum\limits_{i=0}^{k} (\frac{i}{k} - x)^2 p_{k,i}(x) = \sum\limits_{i=0}^{k} \frac{i^2}{k^2} p_{k,i}(x) - 2x \sum\limits_{i=0}^{k} \frac{i}{k} p_{k,i}(x) + x^2 \sum\limits_{i=0}^{k} p_{k,i}(x)$
$= \sum\limits_{i=0}^{k} \frac{i(i-1)}{k^2} p_{k,i}(x) + \frac{1}{k} \sum\limits_{i=0}^{k} \frac{i}{k} p_{k,i}(x) - 2x^2 + x^2$

$$= x^2 \tfrac{k-1}{k} \sum_{i=2}^{k} \tbinom{k-2}{i-2} x^{i-2}(1-x)^{(k-2)-(i-2)} + \tfrac{x}{k} - x^2$$
$$= x^2 \tfrac{k-1}{k} \sum_{j=0}^{k-2} \tbinom{k-2}{j} x^{j}(1-x)^{(k-2)-j} + \tfrac{x}{k} - x^2 = \tfrac{x-x^2}{k}.$$

□

Teorema 3.6.1. *Dada una función $f \in \mathcal{C}([a,b],\mathbb{R})$, existe una sucesión de polinomios convergente uniformemente a f.*

Demostración. Para simplificar la demostración supondremos que $[a,b]=[0,1]$. Demostremos entonces que la sucesión de polinomios de Berstein $\{b_k\}$ asociados a la función f, converge uniformemente a f.

Dado $\varepsilon > 0$, debemos probar que existe $k_0 \in \mathbb{N}$ tal que

$$\|b_k - f\|_\infty \leq \varepsilon \text{ para todo } \ k \geq k_0.$$

Del Lema anterior (i) vemos que para cada $x \in [0,1]$ y $k \in \mathbb{N}$ se tiene

$$|b_k(x) - f(x)| = \left|\sum_{i=0}^{k} f\left(\frac{i}{k}\right) p_{k,i}(x) - f(x)\right|$$

$$= \left|\sum_{i=0}^{k} f\left(\frac{i}{k}\right) p_{k,i}(x) - \sum_{i=0}^{k} f(x)p_{k,i}(x)\right|$$

$$= \left|\sum_{i=0}^{k} p_{k,i}(x)\left(f\left(\frac{i}{k}\right) - f(x)\right)\right|$$

$$\leq \sum_{i=0}^{k} p_{k,i}(x)\left|f\left(\frac{i}{k}\right) - f(x)\right|. \qquad (*)$$

Debemos ahora acotar por ε esta última sumatoria. Para esto vamos a separar la suma en dos partes y, usando argumentos distintos, vamos a acotar por $\varepsilon/2$ cada una de ellas.

Del Teorema 2.6.1 sabemos que f es uniformemente continua en $[0,1]$, por lo tanto existe $\delta > 0$ tal que

$$|x' - x| \leq \delta \Rightarrow |f(x') - f(x)| \leq \varepsilon/2$$

y así, del lema anterior (i), obtenemos para cada $x \in [0,1]$ y $k \in \mathbb{N}$

$$\sum_{i \in I_k(x)} p_{k,i}(x) \left| f\left(\frac{i}{k}\right) - f(x) \right| \leq \varepsilon/2 \tag{**}$$

donde $I_k(x) = \{i \leq k : |\frac{i}{k} - x| \leq \delta\}$.

Consideremos ahora el complemento $I_k^c(x)$ del conjunto $I_k(x)$, es decir, el conjunto de índices $I_k^c(x) = \{i \leq k : |\frac{i}{k} - x| > \delta\}$.

Usando la desigualdad $|f(\frac{i}{k}) - f(x)| \leq 2\|f\|_\infty$ para todo $i \leq k$ y, la desigualdad $\frac{1}{\delta^2}(\frac{i}{k} - x)^2 > 1$ para todo $i \in I_k^c(x)$, vemos que

$$\begin{aligned} \sum_{i \in I_k^c(x)} p_{k,i}(x) \left| f\left(\frac{i}{k}\right) - f(x) \right| &\leq 2\|f\|_\infty \sum_{i \in I_k^c(x)} p_{k,i}(x) \\ &\leq 2\|f\|_\infty \frac{1}{\delta^2} \sum_{i \in I_k^c(x)} \left(\frac{i}{k} - x\right)^2 p_{k,i}(x) \end{aligned}$$

y usando ahora el lema anterior (iii) obtenemos

$$\sum_{i \in I_k^c(x)} p_{k,i}(x) \left| f\left(\frac{i}{k}\right) - f(x) \right| \leq \frac{2\|f\|_\infty}{\delta^2} \frac{x(1-x)}{k}$$

$$\leq \frac{2\|f\|_\infty}{\delta^2 k} \leq \frac{\varepsilon}{2} \text{ para todo } k \geq k_0 \tag{***}$$

donde k_0 es un entero que verifica $k_0 \geq \frac{4\|f\|_\infty}{\delta^2 \varepsilon}$. De las desigualdades (*), (**), (***) se obtiene la desigualdad

$$|b_k(x) - f(x)| \leq \varepsilon \text{ para todo } k \geq k_0 \text{ y todo } x \in [0,1]$$

y, tomando supremo sobre x, concluímos que

$$\|b_k - f\|_\infty \leq \varepsilon \text{ para todo } k \geq k_0.$$

$\square$

Nota 3.6.1. En el Contraejemplo 3.5.2 construímos en un e.v.n. una sucesión de Cauchy que no era convergente. El teorema de Weierstrass-Stone nos muestra que en el e.v. de los polinomios de una variable real definidos en $[0,1]$, dotado de la norma $\|\cdot\|_\infty$, es fácil construir sucesiones de Cauchy que no convergen. Basta para esto tomar una función continua f de $[0,1]$ en $\mathbb{R}$ que no sea un polinomio y considerar la sucesión de polinomios $\{b_k\}$ definida en (3.6.1). Esta sucesión será convergente en $\mathcal{C}([0,1],\mathbb{R})$ dotado de la norma $\|\cdot\|_\infty$, a la función f y, como f no es un polinomio vemos que en el e.v. de los polinomios, dotado de la norma $\|\cdot\|_\infty$, $\{b_k\}$ será de Cauchy pero no convergente.

3.7 Ejercicios

1. Estudiar la convergencia puntual, la convergencia uniforme y la convergencia para la norma $\|\cdot\|_1$, de las siguientes sucesiones de funciones en $\mathcal{C}([0,1],\mathbb{R})$ para f_k definidas de a) a e) y en $\mathcal{C}([0,2],\mathbb{R})$ para f_k definida en f).

   a) $f_k(x) = \frac{\sin(kx)}{k}$

   b) $f_k(x) = xe^{-kx}$

   c) $f_k(x) = \begin{cases} kx & \text{si } x \in [0,1/k] \\ (kx)^{-1} & \text{si } x \in [1/k,1] \end{cases}$

   d) $f_k(x) = kxe^{-kx}$

   e) $f_k(x) = \frac{kx^2}{1+kx}$

   f) $f_k(x) = \begin{cases} -1 & \text{si } x \in [0,1-1/k] \\ kx-k & \text{si } x \in [1-1/k,1+1/k] \\ 1 & \text{si } x \in [1+1/k,2] \end{cases}$

2. Demuestre que si $\{f_k\}$ es una sucesión en $\mathcal{C}([a,b],\mathbb{R})$ uniformemente convergente a una función f, entonces

$$\int_a^b f(x)dx = \lim_k \int_a^b f_k(x)dx.$$

3. Sean $\{f_n\}$ y $\{g_n\}$ dos sucesiones de funciones en $\mathcal{A}(A,\mathbb{R})$ uniformemente convergentes a f y g respectivamente. Demuestre que la sucesión $\{h_n\}$ definida por $h_n := f_n \cdot g_n$ es uniformemente convergente a la función $h = f \cdot g$.

4. Sea $\vec{F}$ un e.v.n., $\{f_n\}$ una sucesión en $\mathcal{A}(A,\vec{F})$ uniformemente convergente a una función f y $\{g_n\}$ una sucesión de funciones de A en $\vec{F}$ que verifica $\|f_n - g_n\|_\infty \to 0$. Demuestre que la sucesión $\{g_n\}$ está en $\mathcal{A}(A,\vec{F})$ y converge uniformemente a la función f.

5. Sea $\{p_n\}$ una sucesión de polinomios de una variable real definidos por la formula recurrente

$$p_1 = 0,\ p_{n+1}(x) = \frac{1}{2}(x + 2p_n(x) - p_n(x)^2).$$

   Demuestre que para todo $x \in [0,1]$ se tiene

$$0 \le \sqrt{x} - p_n(x) \le \frac{2\sqrt{x}}{2+n\sqrt{x}}$$

y que la sucesión $\{p_n\}$ en $\mathcal{A}([0,1],\mathbb{R})$ converge uniformemente a la función f definida por $f(x) = \sqrt{x}$.

6. Sea $\{f_n\}$ una sucesión de funciones monótonas definidas en un intervalo $[a,b]$ con valores en $\mathbb{R}$, convergente puntualmente a una función continua $f : [a,b] \to \mathbb{R}$. Demuestre que $\{f_n\}$ converge uniformemente a f.

CAPÍTULO 4

ESPACIOS DE HILBERT

4.1 Introducción

Los espacios de Hilbert son los espacios vectoriales normados más usados en los modelos matemáticos de la ingeniería y de la física. Este capítulo constituye una muy breve introducción a estos espacios.

Todo el capítulo gira en torno al hecho que en un espacio de Hilbert la proyección de un punto sobre un conjunto convexo cerrado siempre existe, es única y puede caracterizarse por la desigualdad (4.4.1).

Terminamos el capítulo con tres resultados importantes: el teorema de representación de Riesz, el teorema de Hanh-Banach y el lema de Farkas.

4.2 Producto interno en un espacio vectorial

Definición 4.2.1. Dado un e.v. $\vec{E}$ sobre el cuerpo $\mathbb{R}$, se llama producto interno o producto escalar en $\vec{E}$ a toda función bilineal simétrica y definida positiva $b : \vec{E} \times \vec{E} \to \mathbb{R}$, esto es, para todo $\vec{u}, \vec{v}, \vec{w} \in \vec{E}$ y $\lambda \in \mathbb{R}$ se tiene:

(i) $b(\vec{u} + \vec{v}, \vec{w}) = b(\vec{u}, \vec{w}) + b(\vec{v}, \vec{w})$ y $b(\vec{u}, \vec{v} + \vec{w}) = b(\vec{u}, \vec{v}) + b(\vec{u}, \vec{w})$;

(ii) $b(\lambda\vec{u}, \vec{v}) = \lambda b(\vec{u}, \vec{v})$ y $b(\vec{u}, \lambda\vec{v}) = \lambda b(\vec{u}, \vec{v})$;

(iii) $b(\vec{u},\vec{v}) = b(\vec{v},\vec{u})$;

(iv) $b(\vec{u},\vec{u}) > 0$ para todo $\vec{u} \neq 0$.

Nota 4.2.1. Las propiedades (i) y (ii) corresponden al hecho que b es bilineal, la propiedad (iii) a la simetría de b y la propiedad (iv) al hecho de ser definida positiva.

Nota 4.2.2. En adelante usaremos para el producto interno la notación $\langle \vec{u},\vec{v}\rangle$ en lugar de $b(\vec{u},\vec{v})$.

Teorema 4.2.1. *Todo producto interno en un e.v. $\vec{E}$ verifica la llamada desigualdad de Cauchy-Schwarz:*

$$|\langle \vec{u},\vec{v}\rangle| \leq \langle \vec{u},\vec{u}\rangle^{\frac{1}{2}}\langle \vec{v},\vec{v}\rangle^{\frac{1}{2}} \quad \textit{para todo } \vec{u},\vec{v} \in \vec{E}. \tag{4.2.1}$$

Además, la igualdad se tiene si y solo si u y v son colineales, esto es, si $\vec{u} = \lambda\vec{v}$ para algún $\lambda \in \mathbb{R}$ o si $\vec{v} = \vec{0}$.

Demostración. Dados $\vec{u},\vec{v} \in \vec{E}$ no colineales, para todo $\lambda,\mu \in \mathbb{R}\setminus\{0\}$ se tiene

$$0 < \langle \lambda\vec{u} - \mu\vec{v}, \lambda\vec{u} - \mu\vec{v}\rangle = \lambda^2\langle \vec{u},\vec{u}\rangle - 2\lambda\mu\langle \vec{u},\vec{v}\rangle + \mu^2\langle \vec{v},\vec{v}\rangle$$

y como $\langle \vec{v},\vec{v}\rangle > 0$ y $\langle \vec{u},\vec{u}\rangle > 0$, haciendo $\lambda := \langle \vec{v},\vec{v}\rangle^{\frac{1}{2}}$ y $\mu := \langle \vec{u},\vec{u}\rangle^{\frac{1}{2}}$ obtenemos

$$2\langle \vec{v},\vec{v}\rangle^{\frac{1}{2}}\langle \vec{u},\vec{u}\rangle^{\frac{1}{2}}\langle \vec{u},\vec{v}\rangle < 2\langle \vec{v},\vec{v}\rangle\langle \vec{u},\vec{u}\rangle$$

y dividiendo por $[\langle \vec{v},\vec{v}\rangle^{\frac{1}{2}}\langle \vec{u},\vec{u}\rangle^{\frac{1}{2}}]$ llegamos a la desigualdad

$$\langle \vec{u},\vec{v}\rangle < \langle \vec{u},\vec{u}\rangle^{\frac{1}{2}}\langle \vec{v},\vec{v}\rangle^{\frac{1}{2}} \text{ para todo } \vec{u},\vec{v} \in \vec{E} \text{ no colineales.}$$

Si en la desigualdad anterior hacemos $\vec{u} = -\vec{v}$ y $\vec{v} = \vec{u}$ obtenemos

$$-\langle \vec{u},\vec{u}\rangle^{\frac{1}{2}}\langle \vec{v},\vec{v}\rangle^{\frac{1}{2}} < \langle \vec{u},\vec{v}\rangle \quad \text{para todo} \quad \vec{u},\vec{v} \in \vec{E} \text{ no colineales}$$

Se obtiene así la desigualdad de Cauchy-Schwarz para $\vec{u},\vec{v}$ no colineales.

La igualdad, cuando u y v son colineales, se obtiene trivialmente. □

Nota 4.2.3. Es fácil demostrar que un producto interno b en un e.v. $\vec{E}$, define una norma en $\vec{E}$, esta es $\|\vec{u}\| = [\langle \vec{u},\vec{u}\rangle]^{1/2}$. Recíprocamente no toda norma se puede definir a partir de un producto interno.

Ejemplo 4.2.1. De las tres normas que hemos definido en $\mathbb{R}^n$ (ver Ejemplo 1.2.1) sólo la norma $\|\cdot\|_2$ se puede definir a partir de un producto interno, este es

$$\langle \vec{u}, \vec{v} \rangle := u_1 v_1 + ... + u_n v_n$$

que usualmente se denota por $\vec{u} \cdot \vec{v}$.

De las dos normas que hemos usado en $\mathcal{C}([a,b],\mathbb{R})$, la norma $\|\cdot\|_\infty$ definida por (3.2.1) y la norma $\|\cdot\|_1$ definida por (3.5.1), ninguna se puede definir a partir de un producto interno. En este mismo e.v. se puede definir el producto interno

$$\langle f, g \rangle := \int_a^b f(x)g(x)dx \tag{4.2.2}$$

que define en $\mathcal{C}([a,b],\mathbb{R})$ la norma

$$\|f\|_2 := [\int_a^b [f(x)]^2 dx]^{\frac{1}{2}} \tag{4.2.3}$$

Definición 4.2.2. Se llama espacio de Hilbert a todo espacio de Banach (ver Definición 1.4.4) cuya norma está definida por un producto interno.

Ejemplo 4.2.2. De acuerdo al Teorema 1.4.7, $\mathbb{R}^n$ con la norma $\|\cdot\|_2$ es un espacio de Hilbert. Pero, de acuerdo a lo que decíamos en el Ejemplo 4.2.1, $\mathbb{R}^n$ con la norma $\|\cdot\|_1$ o $\mathbb{R}^n$ con la norma $\|\cdot\|_\infty$ no son espacios de Hilbert.

Teorema 4.2.2. *Sea $\vec{E}$ un espacio de Hilbert. Las funciones $\langle \cdot, \cdot \rangle$ de $\vec{E} \times \vec{E}$ en $\mathbb{R}$ y $\langle \vec{b}, \cdot \rangle$ de $\vec{E}$ en $\mathbb{R}$, con $\vec{b} \in \vec{E}$ fijo, son continuas.*

Demostración. Demostremos que la función bilineal $\langle \cdot, \cdot \rangle$ es continua en el e.v.n. producto $\vec{E} \times \vec{E}$. Para esto usemos el Teorema 2.3.5. Debemos probar entonces que dado $(\vec{a}, \vec{b}) \in \vec{E} \times \vec{E}$, toda sucesión $(\vec{x}_k, \vec{y}_k) \to (\vec{a}, \vec{b})$ verifica que $\langle \vec{x}_k, \vec{y}_k \rangle \to \langle \vec{a}, \vec{b} \rangle$. Aplicando la desigualdad de Cauchy-Schwarz se tiene que

$$\begin{aligned} |\langle \vec{x}_k, \vec{y}_k \rangle - \langle \vec{a}, \vec{b} \rangle| &= |\langle \vec{x}_k - \vec{a}, \vec{y}_k \rangle + \langle \vec{a}, \vec{y}_k - \vec{b} \rangle| \\ &\leq \|\vec{x}_k - \vec{a}\|\|\vec{y}_k\| + \|\vec{a}\|\|\vec{y}_k - \vec{b}\| \end{aligned}$$

y si $(\vec{x}_k, \vec{y}_k) \to (\vec{a}, \vec{b})$, del Teorema 1.4.3 vemos que $\vec{x}_k \to \vec{a}$ e $\vec{y}_k \to \vec{b}$, lo que nos permite concluir

La continuidad de la función lineal $\langle\vec{b},\cdot\rangle$ es una consecuencia inmediata del teorema anterior que nos dice que

$$|\langle\vec{b},\vec{x}\rangle| \leq \|\vec{b}\|\|\vec{x}\| \text{ para todo } \vec{x}\in\vec{E}$$

y de la caracterización de la continuidad de una función lineal dada por el Teorema 2.7.1. □

4.3 Proyección de un punto sobre un conjunto en un espacio de Hilbert

Definición 4.3.1. Una parte C de un e.v. $\vec{E}$ se dice convexa si para todo par de elementos $\vec{a},\vec{b}\in C$ se tiene que

$$[\vec{a},\vec{b}] := \{\vec{a}+\lambda(\vec{b}-\vec{a}) : \lambda\in[0,1]\}\subset C. \tag{4.3.1}$$

Nota 4.3.1. Los conjuntos convexos en un e.v. tienen una gran importancia. En el próximo teorema veremos una de sus propiedades fundamentales. Más adelante veremos otras.

Nota 4.3.2. Recordemos que se llama proyección de un elemento $\vec{a}$ de un e.v.n. $\vec{E}$ sobre un subconjunto A de $\vec{E}$ a todo elemento $\vec{p}\in A$ que verifica la igualdad

$$d_A(\vec{a}) := \inf_{\vec{x}\in A}\|\vec{x}-\vec{a}\| = \|\vec{p}-a\| \tag{4.3.2}$$

y, que en el Teorema 2.5.4 demostramos que si A es compacto entonces $\vec{p}$ siempre existe. Después, en la Nota 2.5.3 veíamos que si E era de dimensión finita, bastaba con suponer A cerrado para asegurar la existencia de $\vec{p}$. En esta sección daremos un nuevo teorema de existencia de la proyección. Previamente necesitamos el siguiente lema técnico.

Lema 4.3.1. *Sean $\vec{a},\vec{b},\vec{c}$ tres elementos en un espacio de Hilbert $\vec{E}$ y $\vec{m} := \frac{\vec{b}+\vec{c}}{2}$ el punto medio del trazo $[b,c]$. Entonces se tiene la igualdad:*

$$\|\vec{a}-\vec{b}\|^2+\|\vec{a}-\vec{c}\|^2 = 2\|\vec{a}-\vec{m}\|^2+\frac{1}{2}\|\vec{b}-\vec{c}\|^2$$

Demostración.

$$\begin{aligned}
\|\vec{a}-\vec{b}\|^2+\|\vec{a}-\vec{c}\|^2 &= \|\vec{a}\|^2-2\langle\vec{a},\vec{b}\rangle+\|\vec{b}\|^2+\|\vec{a}\|^2-2\langle\vec{a},\vec{c}\rangle+\|\vec{c}\|^2\\
&= 2\|\vec{a}\|^2-2\langle\vec{a},\vec{b}+\vec{c}\rangle+\|\vec{b}+\vec{c}\|^2-2\langle\vec{b},\vec{c}\rangle\\
&= 2[\|\vec{a}\|^2-2\langle\vec{a},\vec{m}\rangle+\|\vec{m}\|^2]+\frac{1}{2}\|\vec{b}+\vec{c}\|^2-2\langle\vec{b},\vec{c}\rangle\\
&= 2\|\vec{a}-\vec{m}\|^2+\frac{1}{2}\|\vec{b}-\vec{c}\|^2
\end{aligned}$$

$\square$

Teorema 4.3.1. *Si $\vec{E}$ es un espacio de Hilbert y C un conjunto convexo cerrado en $\vec{E}$, entonces todo elemento $\vec{a} \in \vec{E}$, tiene una proyección sobre C. Esto significa, de acuerdo a la igualdad* (4.3.2)*, que existe un elemento $\vec{p}(\vec{a}) \in C$ que verifica*

$$d_C(\vec{a}) = \|\vec{p}(\vec{a}) - \vec{a}\|.$$

Se tiene además que esta proyección es única.

Demostración. De acuerdo a la definición de $d_C(\vec{a})$ dada por (2.2.3), debe existir una sucesión $\{\vec{x}_k\}$ en C tal que

$$d_C(\vec{a}) = \lim \|\vec{x}_k - \vec{a}\| \tag{*}$$

Demostremos primero que toda sucesión $\{x_k\}$ en C que verifica (*) es una sucesión de Cauchy. Para esto apliquemos el lema anterior al trío $\vec{a}, \vec{x}_j, \vec{x}_k$

$$\|\vec{x}_j - \vec{x}_k\|^2 = 2\|\vec{a} - \vec{x}_j\|^2 + 2\|\vec{a} - \vec{x}_k\|^2 - 4\|\vec{a} - \frac{\vec{x}_j + \vec{x}_k}{2}\|^2.$$

Como C es un convexo se tiene que

$$\frac{\vec{x}_j + \vec{x}_k}{2} = \vec{x}_k + (1 - \frac{1}{2})(\vec{x}_j - \vec{x}_k) \in C$$

lo que implica

$$-4\|\vec{a} - \frac{\vec{x}_j + \vec{x}_k}{2}\|^2 \leq -4d_C^2(\vec{a})$$

obteniendo la desigualdad

$$\|\vec{x}_j - \vec{x}_k\|^2 \leq 2\|\vec{a} - \vec{x}_j\|^2 + 2\|\vec{a} - \vec{x}_k\|^2 - 4d_C^2(\vec{a}). \tag{**}$$

Así entonces, dado $\varepsilon > 0$, concluímos de la relación (*) que existe $k_0 \in \mathbb{N}$ tal que

$$\Big|\ 2\|\vec{a} - \vec{x}_j\|^2 + 2\|\vec{a} - \vec{x}_k\|^2 - 4d_A^2(\vec{a})\ \Big| \leq \varepsilon \quad \text{para todo} \quad j, k \geq k_0$$

y de la desigualdad (**) obtenemos

$$\|\vec{x}_j - \vec{x}_k\|^2 \leq \varepsilon \quad \text{para todo} \;\; j, k \geq k_0$$

lo que muestra que $\{\vec{x}_k\}$ es una sucesión de Cauchy.

Por ser $\vec{E}$ un espacio de Hilbert sabemos que $\{\vec{x}_k\}$ debe converger a un elemento $\vec{p} \in \vec{E}$ y como C es un conjunto cerrado, de acuerdo al Teorema 1.4.1, $\vec{p} \in C$. La continuidad de la función $\| \cdot - \vec{a}\|$ nos permite concluir, de la relación (*), que

$$d_C(\vec{a}) = \|\vec{p} - \vec{a}\|$$

es decir, que $\vec{p}$ es una proyección de $\vec{a}$ sobre C.

Para terminar debemos demostrar la unicidad de la proyección. Si $\vec{p}\,'$, es otra proyección de $\vec{a}$ en C, se tendrá que la sucesión $\{\vec{x}_k^{\,'}\}$, definida por $\vec{x}_k^{\,'} := \vec{x}_k$ si k es par y $\vec{x}_k^{\,'} := \vec{p}\,'$ si k es impar, también va a verificar la igualdad (*) (puesto que $\|\vec{p} - \vec{a}\| = \|\vec{p}\,' - \vec{a}\|$) y, será entonces de Cauchy y por lo tanto convergente. Del Teorema 1.4.5 vemos que $\vec{p}\,'$ y $\vec{p}$ son puntos de acumulación de $\{\vec{x}_k^{\,'}\}$ y como ella es convergente, de la Nota 1.4.3 concluímos que $\vec{p}\,' = \vec{p}$. $\square$

Nota 4.3.3. Es importante tener claro que en un espacio de Banach la proyección sobre un convexo cerrado no es necesariamente única, como lo muestra el ejemplo i) que vimos en la Nota 2.5.2. El hecho que la norma pueda definirse a partir de un producto interno, es entonces el punto clave para asegurar la unicidad.

4.4 Caracterización de la proyección sobre un conjunto convexo

Teorema 4.4.1. *Si C es una parte convexa cerrada de un espacio de Hilbert $\vec{E}$, entonces la proyección de $\vec{a} \in \vec{E}$ sobre C es el único elemento $\vec{p} \in C$ que verifica la desigualdad*

$$\langle \vec{a} - \vec{p}, \vec{x} - \vec{p}\,\rangle \leq 0 \quad \textit{para todo} \quad \vec{x} \in C. \tag{4.4.1}$$

Demostración. Demostremos que (4.4.1) implica que $\vec{p}(\vec{a}) = \vec{p}$. De la identidad

$$\begin{aligned} \|\vec{a} - \vec{x}\|^2 &= \|\vec{a} - \vec{p} - (\vec{x} - \vec{p})\|^2 \\ &= \|\vec{a} - \vec{p}\|^2 - 2\langle \vec{a} - \vec{p}, \vec{x} - \vec{p}\rangle + \|\vec{x} - \vec{p}\|^2 \end{aligned}$$

vemos que (4.4.1) implica que

$$\|\vec{a} - \vec{x}\| \geq \|\vec{a} - \vec{p}\| \quad \text{para todo } \vec{x} \in C$$

lo que equivale a decir que $\vec{p}$ es la proyección $\vec{a}$ sobre C.

Demostremos ahora que la proyección $\vec{p}(\vec{a})$ verifica la desigualdad (4.4.1).

Puesto que C es convexo se tiene que $p(\vec{a}) + t(\vec{x} - \vec{p}(\vec{a})) \in C$ para todo $\vec{x} \in C$ y todo $t \in [0,1]$. Se tendrá entonces que

$$\|\vec{a} - [\vec{p}(\vec{a}) + t(\vec{x} - \vec{p}(\vec{a}))]\|^2 \geq \|\vec{a} - \vec{p}(\vec{a})\|^2$$

y, desarrollando el cuadrado de la izquierda se obtiene para todo $\vec{x} \in C$ la desigualdad

$$-2\langle \vec{a} - \vec{p}(\vec{a}), t(\vec{x} - p(\vec{a}))\rangle + t^2\|\vec{x} - \vec{p}(\vec{a})\|^2 \geq 0 \quad \forall t \in [0,1]$$

y simplificando por t

$$2\langle \vec{a} - \vec{p}(\vec{a}), \vec{x} - \vec{p}(\vec{a})\rangle \leq t\|\vec{x} - \vec{p}(\vec{a})\|^2 \quad \forall t \in]0,1]$$

lo cual implica, haciendo tender t a cero, que $\vec{p}(\vec{a})$ verifica la desigualdad (4.4.1). $\square$

Teorema 4.4.2. *Si $\vec{S}$ es un subespacio vectorial cerrado de un espacio de Hilbert $\vec{E}$, entonces la proyección de $\vec{a} \in \vec{E}$ sobre $\vec{S}$ es el único elemento $\vec{p} \in \vec{S}$ que verifica*

$$\langle \vec{a} - \vec{p}, \vec{x}\rangle = 0 \quad \textit{para todo } \vec{x} \in \vec{S}. \tag{4.4.2}$$

Demostración. Por ser $\vec{S}$ un conjunto convexo cerrado, podemos aplicar el teorema anterior y concluir que la proyección $\vec{p}(\vec{a})$ es el único $\vec{p} \in \vec{S}$ que verifica

$$\langle \vec{a} - \vec{p}, \vec{x} - \vec{p}\rangle \leq 0 \text{ para todo } \vec{x} \in \vec{S} \tag{*}$$

Como $\vec{S}$ es un subespacio vectorial y $\vec{p} \in \vec{S}$, es claro que

$$\vec{x} \in \vec{S} \Leftrightarrow \vec{x} - \vec{p} \in \vec{S} \Leftrightarrow -(\vec{x} - \vec{p}) \in S.$$

Haciendo entonces en (*) los cambios de variable $\vec{y} = \vec{x} - \vec{p}$ y $\vec{z} = -(\vec{x} - \vec{p})$ concluímos que (*) es equivalente a " $\langle \vec{a} - \vec{p}, \vec{y}\rangle \leq 0$ para todo $\vec{y} \in S$ " y a " $\langle \vec{a} - \vec{p}, \vec{z}\rangle \geq 0$ para todo $\vec{z} \in \vec{S}$ ", respectivamente. Esas dos desigualdades nos muestran que (*) es equivalente a (4.4.2) que es lo que queríamos demostrar. $\square$

Nota 4.4.1. Se demuestra fácilmente que todo subespacio vectorial de dimensión finita de un e.v.n. $\vec{E}$ es un conjunto cerrado. En particular todo subespacio vectorial de $\mathbb{R}^n$ será cerrado. Por otra parte, en la Sección 3.6 vimos que el sub e.v. de los polinomios, en el e.v. $\mathcal{C}([0,1],\mathbb{R})$ dotado la norma $\|\cdot\|_\infty$, no es cerrado.

Definición 4.4.1. Dos elementos $\vec{a}, \vec{b}$ en un espacio de Hilbert se dirán ortogonales o perpendiculares si

$$\langle \vec{a}, \vec{b} \rangle = 0. \tag{4.4.3}$$

Un elemento $\vec{a} \in \vec{E}$ se dirá ortogonal o perpendicular a un subespacio vectorial $\vec{S}$ de $\vec{E}$ si es ortogonal a todos los elementos de $\vec{S}$, es decir si

$$\langle a, \vec{x} \rangle = 0 \quad \text{para todo } \vec{x} \in \vec{S}. \tag{4.4.4}$$

Dos subespacios vectoriales $\vec{S}, \vec{G}$ de un espacio de Hilbert $\vec{E}$ se dirán ortogonales si

$$\langle \vec{x}, \vec{y} \rangle = 0 \quad \text{para todo} \quad \vec{x} \in \vec{S}, \vec{y} \in \vec{G}.$$

Nota 4.4.2. Del teorema anterior y de la definición anterior podemos decir que la proyección de un elemento $\vec{a}$ de un espacio de Hilbert $\vec{E}$ sobre un subespacio vectorial $\vec{S}$ es el elemento $\vec{p} \in \vec{S}$ tal que $\vec{a} - \vec{p}$ es ortogonal a todos los elementos de $\vec{S}$.

Definición 4.4.2. Se llama sistema ortonormado en un espacio de Hilbert $\vec{E}$ a todo conjunto $\{\vec{E}_1, ..., \vec{E}_n\}$ de elementos de $\vec{E}$ que son ortogonales entre si y de norma uno.

Teorema 4.4.3. *Sea $\vec{S}$ el subespacio vectorial generado por un sistema ortonormado $\{\vec{E}_1, ..., \vec{E}_n\}$ en un espacio de Hilbert $\vec{E}$. Entonces la proyección $\vec{p}(\vec{a})$ de un elemento $\vec{a} \in \vec{E}$ sobre $\vec{S}$ está dada por*

$$\vec{p}(\vec{a}) = \sum_{i=1}^{n} \lambda_i \vec{E}_i \quad \textit{donde} \quad \lambda_i = \langle \vec{a}, \vec{E}_i \rangle \tag{4.4.5}$$

y se tiene

$$\|\vec{a} - \vec{p}(\vec{a})\|^2 = \|\vec{a}\|^2 - \sum_{i=1}^{n} \lambda_i^2. \tag{4.4.6}$$

Demostración. Como decíamos en la Nota 4.4.1, por ser $\vec{S}$ de dimensión finita, es cerrado en $\vec{E}$. Entonces de la relación (4.4.2), que caracteriza la proyección de $\vec{a}$ sobre $\vec{S}$, vemos que $\vec{p}(\vec{a}) = \sum \lambda_i \vec{E}_i$ debe verificar

$$\langle \vec{a} - \sum_{i=1}^{n} \lambda_i \vec{E}_i, \vec{E}_j \rangle = 0 \quad \text{para todo} \quad j = 1, ..., n$$

y, como los $\vec{E}_i$ son ortonormales, estas n igualdades se reducen a

$$\langle \vec{a}, \vec{E}_j \rangle - \lambda_j = 0 \quad \text{para todo} \quad j = 1, ..., n$$

lo que demuestra la fórmula (4.4.5).

Calculemos ahora $\|\vec{a} - \vec{p}(\vec{a})\|^2$:

$$\begin{aligned} \|\vec{a} - \vec{p}(\vec{a})\|^2 &= \langle \vec{a} - \vec{p}(\vec{a}), \vec{a} - \vec{p}(\vec{a}) \rangle = \langle \vec{a} - \vec{p}(\vec{a}), \vec{a} \rangle \\ &= \|\vec{a}\|^2 - \langle \vec{p}(\vec{a}), \vec{a} \rangle = \|\vec{a}\|^2 - \langle \sum \lambda_i \vec{E}_i, \vec{a} \rangle \\ &= \|\vec{a}\|^2 - \sum \lambda_i \langle \vec{E}_i, \vec{a} \rangle = \|\vec{a}\|^2 - \sum \lambda_i^2 \end{aligned}$$

que es lo que queríamos probar. □

Teorema 4.4.4. *La proyección $\vec{p}(\vec{a})$ de un elemento $\vec{a}$ de un espacio de Hilbert $\vec{E}$ sobre la recta generada por un elemento $\vec{b} \neq \vec{0}$, está dada por*

$$\vec{p}(\vec{a}) = \langle \vec{a}, \frac{\vec{b}}{\|\vec{b}\|} \rangle \frac{\vec{b}}{\|\vec{b}\|}. \tag{4.4.7}$$

Demostración. Se trata de proyectar $\vec{a}$ sobre el subespacio vectorial generado por el sistema ortonormado $\{\frac{\vec{b}}{\|\vec{b}\|}\}$ formado por un solo elemento. La fórmula (4.4.7) es entonces una consecuencia inmediata de la fórmula (4.4.5). □

4.5 Continuidad de la proyección sobre un conjunto convexo

Teorema 4.5.1. *La función $\vec{p}(\cdot)$ que a todo elemento de un espacio de Hilbert $\vec{E}$ le hace corresponder su proyección sobre un subespacio $\vec{S}$ de dimensión finita, es lineal y continua.*

Demostración. De acuerdo al Teorema 4.4.3 hay que demostrar que la función

$$\vec{p}(\vec{x}) = \sum_{i=1}^{n} \langle \vec{x}, \vec{E}_i \rangle \vec{E}_i$$

donde $\{\vec{E}_1, ..., \vec{E}_n\}$ es una base ortonormada de $\vec{S}$, es lineal y continua, lo que es una consecuencia inmediata de la linealidad y continuidad de cada una de las funciones $\langle \cdot, \vec{E}_i \rangle$. □

Teorema 4.5.2. *La función $\vec{p}(\cdot)$ que a todo elemento de un espacio de Hilbert $\vec{E}$ le hace corresponder su proyección sobre un convexo cerrado C de $\vec{E}$, es Lipschitziana:*

$$\|\vec{p}(\vec{x}) - \vec{p}(\vec{z})\| \leq \|\vec{x} - \vec{z}\| \text{ para todo } \vec{x}, \vec{z} \in \vec{E}. \tag{4.5.1}$$

Demostración. Si escribimos $\vec{x} - \vec{z} = \vec{p}(\vec{x}) - \vec{p}(\vec{z}) + \vec{u}$ vemos que

$$\begin{aligned} \|\vec{x} - \vec{z}\|^2 &= \|\vec{p}(\vec{x}) - \vec{p}(\vec{z})\|^2 + 2\langle \vec{p}(\vec{x}) - \vec{p}(\vec{z}), \vec{u} \rangle + \|\vec{u}\|^2 \\ &\geq \|\vec{p}(\vec{x}) - \vec{p}(\vec{z})\|^2 + 2\langle \vec{p}(\vec{x}) - \vec{p}(\vec{z}), \vec{u} \rangle. \end{aligned}$$

Por lo tanto, para obtener (4.5.1) bastará con demostrar que $\langle \vec{p}(\vec{x}) - \vec{p}(\vec{z}), \vec{u} \rangle \geq 0$. Puesto que $\vec{u} = \vec{x} - \vec{p}(\vec{x}) - (\vec{z} - \vec{p}(\vec{z}))$ se tendrá

$$\langle \vec{p}(\vec{x}) - \vec{p}(\vec{z}), \vec{u} \rangle = \langle \vec{x} - \vec{p}(\vec{x}), \vec{p}(\vec{x}) - \vec{p}(\vec{z}) \rangle + \langle \vec{z} - \vec{p}(\vec{z}), \vec{p}(\vec{z}) - \vec{p}(\vec{x}) \rangle$$

y del Teorema 4.4.1 se desprende de inmediato que cada uno de los dos sumandos de la derecha de esta igualdad son no negativos. $\square$

4.6 Espacios suplementarios y proyección

Definición 4.6.1. Dos subespacios vectoriales $\vec{S}_1, \vec{S}_2$ de un e.v. $\vec{E}$ se dirán suplementarios si $\vec{S}_1 \cap \vec{S}_2 = \{\vec{0}\}$ y para todo $\vec{x} \in \vec{E}$ existen $\vec{x}_1 \in \vec{S}_1, \vec{x}_2 \in \vec{S}_2$ tales que $\vec{x} = \vec{x}_1 + \vec{x}_2$.

Teorema 4.6.1. *Si $\vec{S}_1$ y $\vec{S}_2$ son dos subespacios vectoriales cerrados suplementarios ortogonales en un espacio de Hilbert $\vec{E}$, entonces*

$$\vec{x} = \vec{p}_1(\vec{x}) + \vec{p}_2(\vec{x}) \quad \text{para todo} \quad \vec{x} \in \vec{E}$$

donde $\vec{p}_1(\vec{x})$ y $\vec{p}_2(\vec{x})$ denotan la proyecciones de $\vec{x}$ sobre $\vec{S}_1$ y $\vec{S}_2$ respectivamente.

Demostración. Dado $\vec{x} \in \vec{E}$, sean $\vec{x}_1 \in \vec{S}_1, \vec{x}_2 \in \vec{S}_2$ tales que $\vec{x} = \vec{x}_1 + \vec{x}_2$. Como $\vec{S}_1$ y $\vec{S}_2$ son ortogonales (ver Definición 4.4.1), se tiene

$$\langle \vec{x} - \vec{x}_1, \vec{z} \rangle = 0 \quad \text{para todo } \vec{z} \in \vec{S}_1$$

y

$$\langle \vec{x} - \vec{x}_2, \vec{z} \rangle = 0 \quad \text{para todo } \vec{z} \in \vec{S}_2$$

lo que muestra, de acuerdo al Teorema 4.4.2, que $\vec{x}_1 = \vec{p}_1(\vec{x})$ y $\vec{x}_2 = \vec{p}_2(\vec{x})$. $\square$

Definición 4.6.2. Se llama hiperplano en un e.v. $\vec{E}$ a todo subespacio vectorial $\vec{H}$ en $\vec{E}$ de codimensión 1. Esto significa que existe $\vec{a} \in \vec{E} \setminus \vec{H}$ tal que el subespacio vectorial de dimensión 1 generado por $\vec{a}$ y el subespacio vectorial $\vec{H}$ son suplementarios en $\vec{E}$.

Teorema 4.6.2. *Sea l una función lineal no nula de un e.v. $\vec{E}$ en $\mathbb{R}$. Entonces el subespacio vectorial $\vec{H} = \{\vec{x} \in \vec{E} : l(\vec{x}) = 0\}$ es un hiperplano en $\vec{E}$. Si $\vec{E}$ es un e.v.n. y si l es continua, entonces $\vec{H}$ será un hiperplano cerrado.*

Demostración. Sea $\vec{e} \in \vec{E}$ tal que $l(\vec{e}) \neq 0$ y definamos $\vec{a} = \frac{\vec{e}}{l(\vec{e})}$ (de modo que $l(\vec{a}) = 1$). Demostremos entonces que todo $\vec{x} \in \vec{E}$ se puede escribir $\vec{x} = \lambda\vec{a} + \vec{y}$ con $\lambda \in \mathbb{R}, \vec{y} \in \vec{H}$. Pero esto es evidente pues basta con definir $\lambda = l(\vec{x})$ e $\vec{y} = \vec{x} - l(\vec{x})\vec{a}$.

Es inmediato verificar que H es cerrado usando el Teorema 1.4.1 que caracteriza los conjuntos cerrados mediante sucesiones y el Teorema 2.3.5 que caracteriza las funciones continuas mediante sucesiones. $\square$

4.7 Tres Teoremas importantes

Teorema 4.7.1 (Representación de Riesz)**.** *Sea $\vec{E}$ un espacio de Hilbert y $l : \vec{E} \to \mathbb{R}$ una función lineal continua. Entonces existe $\vec{\omega} \in \vec{E}$ tal que*

$$l(\vec{x}) = \langle \vec{\omega}, \vec{x} \rangle \quad \textit{para todo} \quad \vec{x} \in \vec{E}. \tag{4.7.1}$$

Demostración. Si l es la función nula, se tendrá (4.7.1) con $\vec{\omega} = \vec{0}$. Supongamos ahora que l no es la función nula. De acuerdo al Teorema 4.6.2, $\vec{H} = \{\vec{x} \in \vec{E} : l(\vec{x}) = 0\}$ es un hiperplano cerrado. Entonces, si $\vec{a} \in \vec{E} \setminus \vec{H}$ y si $\vec{p}_1(\vec{a})$ es la proyección de $\vec{a}$ en $\vec{H}$, de acuerdo al Teorema 4.4.2, el elemento $\vec{b} := \vec{a} - p_1(\vec{a})$ genera un subespacio de dimensión 1 ortogonal a $\vec{H}$. Del Teorema 4.6.1, se tiene

$$\vec{x} = \vec{p}_1(\vec{x}) + \vec{p}_2(\vec{x}) \quad \text{para todo} \quad \vec{x} \in \vec{E} \tag{*}$$

donde $\vec{p}_1(\vec{x})$ es la proyección de $\vec{x}$ en el espacio generado por $\vec{b}$ y $\vec{p}_2(\vec{x})$ la proyección de $\vec{x}$ sobre $\vec{H}$. Del Teorema 4.4.4 vemos que $\vec{p}_1(\vec{x}) = \langle \frac{\vec{b}}{\|\vec{b}\|}, \vec{x} \rangle \frac{\vec{b}}{\|\vec{b}\|}$ por lo tanto, definiendo $\vec{\omega} := l(\frac{\vec{b}}{\|\vec{b}\|})\frac{\vec{b}}{\|\vec{b}\|}$ obtenemos, a partir de (*), la fórmula (4.7.1). $\square$

Teorema 4.7.2 (Separación de Hanh-Banach). *Sea C una parte convexa cerrada de un espacio de Hilbert $\vec{E}$ que no contiene al origen. Existe entonces una función lineal continua $l : \vec{E} \to \mathbb{R}$ y $\alpha > 0$ tales que*

$$l(\vec{x}) \geq \alpha \quad \textit{para todo} \quad \vec{x} \in C. \tag{4.7.2}$$

Demostración. Sea $\vec{p} \in C$ la proyección de $\vec{0}$ sobre C, cuya existencia está garantizada por el Teorema 4.3.1. Del Teorema 4.4.1 sabemos que

$$\langle -\vec{p}, \vec{x} - \vec{p} \rangle \leq 0 \quad \text{para todo} \quad \vec{x} \in C$$

es decir

$$\langle \vec{p}, \vec{x} \rangle \geq \|\vec{p}\|^2 \quad \text{para todo} \quad \vec{x} \in C.$$

Definiendo $\alpha := \|\vec{p}\|^2 > 0$ y $l : E \to \mathbb{R}$ por $l(\vec{x}) := \langle \vec{p}, \vec{x} \rangle$, del Teorema 4.2.2 vemos que l es una función lineal continua que verifica (4.7.2). $\square$

Nota 4.7.1. Si C es un convexo cerrado de un espacio de Hilbert $\vec{E}$ y $\vec{a} \notin C$, aplicando el teorema anterior es fácil ver que existe una función lineal afín continua $h : \vec{E} \to \mathbb{R}$ y $\alpha > 0$ tales que

$$h(\vec{a}) = 0 \quad \text{y} \quad h(\vec{x}) \geq \alpha \quad \text{para todo} \quad \vec{x} \in C. \tag{4.7.3}$$

En efecto, si definimos $C' = C - \{a\} := \{\vec{x}' \in \vec{E} : \vec{x}' = \vec{x} - \vec{a}$ para algún $\vec{x} \in C\}$, puesto que $0 \notin C'$ debe existir $l \in \mathcal{L}(\vec{E}, \mathbb{R})$ y $\alpha > 0$ tales que $l(\vec{x}) \geq \alpha$ para todo $\vec{x}' \in C'$. Definiendo entonces $h(\vec{x}) := l(\vec{x}) - l(\vec{a})$, se obtiene (4.7.3).

Teorema 4.7.3 (Lema de Farkas). *Dadas $n+1$ funciones lineales continuas $\ell_0, \ell_1, ..., \ell_n$ definidas en un espacio de Hilbert $\vec{E}$ con valores en $\mathbb{R}$, la inclusión*

$$\{\vec{x} \in \vec{E} : \ell_i(\vec{x}) \leq 0 \quad \textit{para todo} \quad i = 1...n\} \subset \{\vec{x} \in \vec{E} : \ell_0(\vec{x}) \leq 0\} \tag{4.7.4}$$

implica la existencia de escalares $\lambda_1 \geq 0, ..., \lambda_n \geq 0$ tales que

$$\ell_0 = \sum_{i=1}^{n} \lambda_i \ell_i. \tag{4.7.5}$$

Demostración. Del Teorema 4.7.1 sabemos que existen $\vec{\omega}_0, \vec{\omega}_1, ..., \vec{\omega}_n \in \vec{E}$ tales que $\ell_i(\vec{x}) = \langle \vec{\omega}_i, \vec{x} \rangle$ para todo $\vec{x} \in \vec{E}$ y todo $i = 1, ..., n$. Si definimos el convexo

$$C := \{\vec{x} \in \vec{E} : \vec{x} = \sum_{i=1}^{n} \lambda_i \vec{\omega}_i \quad \text{con} \quad \lambda_i \geq 0 \quad \text{para todo} \quad i = 1, ..., n\}$$

(4.7.5) equivale a demostrar que $\vec{\omega}_0 \in C$. Como C es un convexo cerrado (ver Nota 4.7.2), de acuerdo al Teorema 4.3.1 esto equivale a demostrar que la proyección $\vec{p}(\vec{\omega}_0)$ de $\vec{\omega}_0$ sobre C es igual a $\vec{\omega}_0$. Escribamos $\vec{p}(\vec{\omega}_0) = \Sigma\bar{\lambda}_i\vec{\omega}_i$ con $\bar{\lambda}_i \geq 0$ para todo $i = 1, ..., n$. Del Teorema 4.4.1 sabemos que $\langle\vec{\omega}_0 - \vec{p}(\vec{\omega}_0), \vec{x} - \vec{p}(\vec{\omega}_0)\rangle \leq 0$ para todo $\vec{x} \in C$ y, aplicando esta desigualdad a $\vec{x}_j := p(\vec{\omega}_0) + \vec{\omega}_j \in C$ obtenemos

$$\langle\vec{\omega}_0 - \vec{p}(\vec{\omega}_0), \vec{\omega}_j\rangle \leq 0 \quad \text{para todo} \quad j = 1, ..., n$$

es decir, $\ell_j(\vec{\omega}_0 - \vec{p}(\vec{\omega}_0)) \leq 0$ para todo $j = 1, ..., n$. De (4.7.4) podemos concluir que

$$\langle\vec{\omega}_0 - \vec{p}(\vec{\omega}_0), \vec{\omega}_0\rangle \leq 0. \tag{*}$$

Por otra parte, como $\vec{0} \in C$, se tiene que $\langle\vec{\omega}_0 - \vec{p}(\vec{\omega}_0), -\vec{p}(\vec{\omega}_0)\rangle \leq 0$ y sumándole la desigualdad (*) vemos que $\|\vec{\omega}_0 - \vec{p}(\vec{\omega}_0)\|^2 \leq 0$ lo que muestra que $\vec{\omega}_0 = \vec{p}(\vec{\omega}_0)$. $\square$

Nota 4.7.2. Probar que el conjunto C de la demostración del teorema anterior, es cerrado, es fácil si los vectores $\vec{\omega}_1, ..., \vec{\omega}_n$ son linealmente independientes. Sin la hipótesis de independencia lineal, la dificultad aumenta.

4.8 Ejercicios

1. Demueste que la norma en un espacio de Hilbert verifica la igualdad

$$\|\vec{x} - \vec{y}\|^2 + \|\vec{x} + \vec{y}\|^2 = 2(\|\vec{x}\|^2 + \|\vec{y}\|^2) \quad \text{para todo} \quad \vec{x}, \vec{y}.$$

2. Demuestre que si $\vec{E}$ es un e.v.n. cuya norma verifica la igualdad del Ejercicio 1, entonces la función $b : E \times E \to E$ definida por

$$b(\vec{x}, \vec{y}) = \frac{1}{4}(\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2)$$

es un producto interno. Demuestre entonces que $\vec{E}$ es un espacio de Hilbert.

3. Dado un punto $\vec{b}$ en un espacio de Hilbert $\vec{E}$, demuestre que $\|\langle\vec{b}, \cdot\rangle\| = \|\vec{b}\|$ (la primera norma es la de $\mathcal{L}(\vec{E}, \mathbb{R})$ y la segunda la de $\vec{E}$).

4. Demuestre que la proyección en un espacio de Hilbert $\vec{E}$, de un punto $\vec{a}$, sobre la bola cerrada $B(\vec{c}, 1)$ (suponiendo que $\vec{a} \notin B(\vec{c}, 1)$), está dada por $\vec{p}(\vec{a}) = \vec{c} + \frac{\vec{a} - \vec{c}}{\|\vec{a} - \vec{c}\|}$.

5. Calcule en $\mathbb{R}^3$, dotado de la norma $\|\cdot\|_2$, la proyección del punto (1,2,3) sobre el plano de ecuación $x_1 + x_2 + x_3 = 0$. Calcule la proyección del mismo punto sobre el plano afín de ecuación $x_1 + x_2 + x_3 = 1$.

CAPÍTULO 5

DERIVADAS PARCIALES Y DIFERENCIAL DE FUNCIONES DEFINIDAS EN UN E.V.N. $\vec{E}$ CON VALORES EN UN E.V.N. $\vec{F}$

5.1 Introducción

En este capítulo abordamos el cálculo diferencial de funciones definidas en un e.v.n. con valores en otro e.v.n.

Las siguientes dos secciones definen las dos nociones fundamentales del capítulo. En la primera definimos la derivada parcial de una función f en un punto $\vec{a}$ de su dominio con respecto a un vector $\vec{v}$, que denotamos $Df(\vec{a};\vec{v})$. En la segunda se define la diferencial de una función en un punto $\vec{a}$, que denotamos $Df(\vec{a})$. La diferencial $Df(\vec{a})$ es una función lineal continua que nos da una aproximación de primer orden de la función en el punto $\vec{a}$ (ver Nota 5.3.2) y la derivada parcial $Df(\vec{a};\vec{v})$ es un vector en el espacio donde f toma sus valores, que nos permite calcular en forma efectiva la diferencial mediante la fórmula $Df(\vec{a})(\vec{v}) = Df(\vec{a};\vec{v})$ para todo $\vec{v}$ (fórmula (5.3.5)). En estas secciones se entregan también algunas fórmulas para el cálculo de la derivada parcial y de la diferencial de la suma de dos funciones, del producto de una función por un escalar, etc. El cálculo para la composición de funciones (regla de la cadena) está en la Sección 5.5.

También se presenta el teorema del valor medio que es fundamental en cálculo diferencial, este teorema solo es válido para funciones con valores en $\mathbb{R}$. Una variante de este teorema, para funciones a valores en un espacio vectorial, esta dado por el Teorema 5.3.7 llamado de los incrementos finitos.

La Sección 5.4 se introducen las funciones de clase C^1 que constituyen la clase más importante de funciones diferenciables.

Dos teoremas fundamentales se desarrollan en la Sección 5.7, el teorema de la función inversa y el de la función implícita.

Haciendo uso de la noción de derivada de orden superior que se ve en la Sección 5.8, en la siguiente se define la noción de desarrollo limitado de orden N de una función en un punto de su dominio. El resultado fundamental de esta sección está dado por el Teorema 5.9.3 donde se calcula explícitamente el desarrollo limitado de orden dos de una función de clase C^2.

5.2 Derivada parcial con respecto a un vector

Definición 5.2.1. Dada una función f definida en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, se llama derivada parcial de f en $\vec{a} \in A$, con respecto al vector $\vec{v} \in \vec{E}$, al elemento de $\vec{F}$ definido por:

$$Df(\vec{a};\vec{v}) := \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(\vec{a}+t\vec{v})-f(\vec{a})}{t} \tag{5.2.1}$$

cuando el límite existe.

Nota 5.2.1. Si definimos la función $\phi :]-\varepsilon,\varepsilon[\to \vec{F}$ por $\phi(t) = f(\vec{a}+t\vec{v})$, es fácil ver que se tiene la fórmula $Df(\vec{a};\vec{v}) = \phi'(0)$, es la derivada de ϕ en 0 según el cálculo de funciones de una variable real. Con esto vemos que cuando $\vec{F} = \mathbb{R}$ y $\|\vec{v}\| = 1$, la cantidad $Df(\vec{a};\vec{v})$ se interpreta como la pendiente de f en $\vec{a}$ en la dirección $\vec{v}$.

Definición 5.2.2. La función f de la Definición 5.2.1 se dirá parcialmente derivable en $\vec{a}$, si $Df(\vec{a};\vec{v})$ existe para todo $\vec{v} \in \vec{E}$.

Definición 5.2.3. Si en la Definición 5.2.1 consideramos que $\vec{E} = \mathbb{R}^n$ y si denotamos por $\vec{e}_1, ..., \vec{e}_n$ la base canónica de $\mathbb{R}^n$, entonces la derivada parcial $Df(\vec{a};\vec{e}_i)$, cuando existe, la denotaremos $\frac{\partial f}{\partial x_i}(\vec{a}) \in \vec{F}$ o bien $\partial_i f(\vec{a})$ y, la llamaremos derivada parcial de f en $\vec{a}$ con respecto a x_i.

Nota 5.2.2. De la fórmula (5.2.1) vemos que la derivada parcial de f en $\vec{a}$ con respecto a x_i

$$\begin{aligned}\frac{\partial f}{\partial x_i}(\vec{a}) &= \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(\vec{a}+t\vec{e}_i)-f(\vec{a})}{t}\\ &= \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a_1,...,a_i+t,...,a_n)-f(a_1,...,a_i,...,a_n)}{t}\end{aligned}$$

corresponde a la derivada de la función de una variable $f(a_1, ..., a_{i-1}, \cdot\, , a_{i+1}, ..., a_n)$ en a_i.

Nota 5.2.3. Cuando se habla de la derivada parcial de una función f con respecto a x_i, se entiende que se trata de la función $\frac{\partial f}{\partial x_i} : A \to \vec{F}$ que a cada $\vec{x} \in A$ le hace corresponder $\frac{\partial f}{\partial x_i}(\vec{x}) \in \vec{F}$.

Teorema 5.2.1. *Si $Df(\vec{a}; \vec{v})$ existe para la función f de la Definición 5.2.1, $Df(\vec{a}; \lambda\vec{v})$ también existe para todo $\lambda \in \mathbb{R}$ y se tiene la igualdad*

$$Df(\vec{a}; \lambda\vec{v}) = \lambda Df(\vec{a}; \vec{v}). \tag{5.2.2}$$

Demostración. Si $\lambda = 0$, la fórmula es evidente. Si $\lambda \neq 0$, hacemos el cambio de variable $s = \lambda t$, y obtenemos

$$Df(\vec{a}; \lambda\vec{v}) = \lim_{\substack{\frac{s}{\lambda} \to 0 \\ \frac{s}{\lambda} \neq 0}} \lambda \frac{f(\vec{a} + s\vec{v}) - f(\vec{a})}{s} = \lambda \lim_{\substack{s \to 0 \\ s \neq 0}} \frac{f(\vec{a} + s\vec{v}) - f(\vec{a})}{s} = \lambda Df(\vec{a}; \vec{v}).$$

□

Teorema 5.2.2. *Dadas dos funciones f y g definidas en un abierto A de un e.v.n. $\vec{E}$, con valores en un e.v.n. $\vec{F}$ y, dos elementos $\vec{a} \in A$ y $\vec{v} \in \vec{E}$ tales que $Df(\vec{a}; \vec{v})$ y $Dg(\vec{a}; \vec{v})$ existen, entonces*

(i) *$D[f + g](\vec{a}; \vec{v})$ existe y se tiene*

$$D[f + g](\vec{a}; \vec{v}) = Df(\vec{a}; \vec{v}) + Dg(\vec{a}; \vec{v}); \tag{5.2.3}$$

(ii) *Para $\lambda \in \mathbb{R}, D[\lambda f](\vec{a}; \vec{v})$ existe y se tiene*

$$D[\lambda f](\vec{a}; \vec{v}) = \lambda Df(\vec{a}; \vec{v}); \tag{5.2.4}$$

(iii) *Si $\vec{F} = \mathbb{R}, D[fg](\vec{a}; \vec{v})$ existe y se tiene*

$$D[fg](\vec{a}; \vec{v}) = g(\vec{a})Df(\vec{a}; \vec{v}) + f(\vec{a})Dg(\vec{a}; \vec{v}); \tag{5.2.5}$$

(iv) *Si $\vec{F} = \mathbb{R}$ y $f(\vec{a}) \neq 0, D[1/f](\vec{a}; \vec{v})$ existe y se tiene*

$$D[1/f](\vec{a}; \vec{v}) = -\frac{1}{f^2(\vec{a})} Df(\vec{a}; \vec{v}). \tag{5.2.6}$$

Demostración. Las fórmulas (5.2.3) y (5.2.4) son una consecuencia directa de la Definición 5.2.1 y del Teorema 2.3.2. Siguiendo la Nota 5.2.1, si definimos las funciones $\phi(t) = f(\vec{a}+t\vec{v})$ y $\xi(t) = g(\vec{a}+t\vec{v})$, vemos que las fórmulas (5.2.5) y (5.2.6) se escriben

$$(\phi\xi)'(0) = \phi'(0)\xi(0) + \phi(0)\xi'(0) \quad \text{y} \quad \left(\frac{1}{\phi}\right)'(0) = -\frac{\phi'(0)}{\phi(0)^2}$$

respectivamente. Estas dos fórmulas fueron demostradas en el curso de cálculo de funciones de una variable real. $\square$

Teorema 5.2.3. *Si f es una función definida en un abierto A de un e.v.n $\vec{E}$, con valores en $\mathbb{R}^m$ y, si denotamos $f_1, ..., f_m$ las funciones componentes de f, entonces para $\vec{a} \in A$ y $\vec{v} \in \vec{E}$, la derivada parcial $Df(\vec{a};\vec{v})$ existe si y sólo si las derivadas parciales $Df_i(\vec{a};\vec{v})$ existen para todo $i = 1, ..., m$ y, en ese caso se tiene que*

$$Df(\vec{a};\vec{v}) = (Df_1(\vec{a};\vec{v}), ..., Df_m(\vec{a};\vec{v})) \in \mathbb{R}^m. \tag{5.2.7}$$

Demostración. Es una consecuencia directa de la Definición 5.2.1 y del Teorema 2.4.2. $\square$

5.3 Diferencial

Definición 5.3.1. Una función f definida en un abierto A de un e.v.n. $\vec{E}$, con valores en un e.v.n. $\vec{F}$, se dirá diferenciable en $\vec{a} \in A$, si existe una función $\ell \in \mathcal{L}(\vec{E},\vec{F})$ (ℓ lineal y continua de $\vec{E}$ en $\vec{F}$) tal que para todo $\vec{\delta} \in \vec{E}$, con $\vec{a}+\vec{\delta} \in A$, se tenga

$$f(\vec{a}+\vec{\delta}) = f(\vec{a}) + \ell(\vec{\delta}) + o(\vec{\delta}) \tag{5.3.1}$$

donde $o(\cdot)$ es una función de $\vec{E}$ en $\vec{F}$ que verifica $o(\vec{0}) = \vec{0}$ y

$$\lim_{\substack{\vec{\delta}\to\vec{0} \\ \vec{\delta}\neq\vec{0}}} \frac{o(\vec{\delta})}{\|\vec{\delta}\|} = \vec{0}. \tag{5.3.2}$$

Haciendo el cambio de variable $\vec{x} = \vec{a}+\vec{\delta} \in A$, la relación (5.3.1) toma la forma

$$f(\vec{x}) = f(\vec{a}) + \ell(\vec{x}-\vec{a}) + o(\vec{x}-\vec{a}) \quad \text{para todo } \vec{x} \in A.$$

A la función lineal continua ℓ (más adelante veremos que es única) se le llama diferencial de f en $\vec{a}$ y se le denota usualmente $Df(\vec{a}) \in \mathcal{L}(\vec{E},\vec{F})$ o bien $df(\vec{a})$.

La función f se dirá diferenciable si es diferenciable en todo punto de A y llamamos diferencial de f a la función de A en $\mathcal{L}(\vec{E},\vec{F})$ que a cada $\vec{x}\in A$ le hace corresponder la diferencial $Df(\vec{x})\in\mathcal{L}(\vec{E},\vec{F})$ de f en $\vec{x}$. A la diferencial de f la denotaremos Df o df.

Nota 5.3.1. La igualdad (5.3.1) junto con (5.3.2) equivale a

$$\lim_{\substack{\vec{\delta}\to\vec{0}\\ \vec{\delta}\neq\vec{0}}} \frac{f(\vec{a}+\vec{\delta})-f(\vec{a})-\ell(\vec{\delta})}{\|\vec{\delta}\|}=\vec{0} \tag{5.3.3}$$

lo que a su vez equivale a decir que: para todo $\varepsilon>0$ existe $\eta>0$ tal que

$$\|\vec{\delta}\|\leq\eta\Rightarrow\|f(\vec{a}+\vec{\delta})-f(\vec{a})-\ell(\vec{\delta})\|\leq\varepsilon\|\vec{\delta}\|. \tag{5.3.4}$$

Ejemplo 5.3.1. Toda función $\ell\in\mathcal{L}(\vec{E},\vec{F})$ es diferenciable y se tiene $D\ell(\vec{a})=\ell$. Esto es una consecuencia inmediata de (5.3.3) y de la identidad $\ell(\vec{a}+\vec{\delta})-\ell(\vec{a})-\ell(\vec{\delta})=\vec{0}$ para todo $\vec{\delta}\in\vec{E}$.

Teorema 5.3.1. *Si f es una función definida en un abierto A de un e.v.n $\vec{E}$, con valores en un e.v.n $\vec{F}$, diferenciable en $\vec{a}\in A$, entonces existe una única función $\ell\in\mathcal{L}(\vec{E},\vec{F})$ que verifica* (5.3.1).

Demostración. Si ℓ_1 y ℓ_2 verifican (5.3.1), usando (5.3.4) que es equivalente a (5.3.1), se obtiene fácilmente que para todo $\varepsilon>0$ existe $\eta>0$ tal que

$$\|\vec{\delta}\|\leq\eta\Rightarrow\|\ell_1(\vec{\delta})-\ell_2(\vec{\delta})\|\leq\varepsilon\|\vec{\delta}\|$$

y como $\|\frac{\eta}{\|\vec{v}\|}\vec{v}\|\leq\eta$ para todo $\vec{v}\in\vec{E}$ no nulo, vemos que para todo $\vec{v}$ no nulo se tendrá

$$\|\ell_1(\frac{\eta}{\|\vec{v}\|}\vec{v})-\ell_2(\frac{\eta}{\|\vec{v}\|}\vec{v})\|\leq\varepsilon\|\frac{\eta}{\|\vec{v}\|}\vec{v}\|$$

y simplificando por $\frac{\eta}{\|\vec{v}\|}$ obtenemos

$$\|\ell_1(\vec{v})-\ell_2(\vec{v})\|\leq\varepsilon\|\vec{v}\|\quad\text{para todo }\vec{v}\in\vec{E}$$

y como esto se tiene para todo $\varepsilon>0$, haciendo tender ε a 0 concluímos que

$$\|\ell_1(\vec{v})-\ell_2(\vec{v})\|=0\quad\text{para todo }\vec{v}\in\vec{E}$$

es decir, $\ell_1(\vec{v})=\ell_2(\vec{v})$ para todo $\vec{v}\in\vec{E}$. □

Nota 5.3.2. La diferencial de f en $\vec{a}$ permite obtener la aproximación lineal afín o aproximación de primer orden de la función f en el punto $\vec{a}$, que está dada por la función $h : \vec{E} \to \vec{F}$ definida por

$$h(\vec{x}) := f(\vec{a}) + Df(\vec{a})(\vec{x} - \vec{a})$$

es decir, la única función lineal afín que verifica $h(\vec{a} + \vec{\delta}) - f(\vec{a} + \vec{\delta}) = o(\vec{\delta})$.

Dos funciones f y h definidas en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, se dicen tangentes en $\vec{a} \in A$ si $f(\vec{a} + \vec{\delta}) - h(\vec{a} + \vec{\delta}) = o(\vec{\delta})$ donde o es una función de $\vec{E}$ en $\vec{F}$ que verifica $o(\vec{0}) = \vec{0}$ y la relación (5.3.2).

De lo anterior vemos que la aproximación de primer orden de la función f en $\vec{a} \in A$ corresponde a la función lineal afín que es tangente a la función f en $\vec{a}$.

Geométricamente, este hecho se expresa diciendo que en $\vec{E} \times \vec{F}$, el subespacio afín definido por el grafo de la función lineal afín h, es tangente al grafo de f en $(\vec{a}, f(\vec{a}))$. Recordemos que el grafo de una función f de A en $\vec{F}$ es el conjunto $\{(\vec{x}, \vec{z}) \in A \times \vec{F} \ : \ \vec{z} = f(\vec{x})\}$. Cuando $\vec{F} = \mathbb{R}$ decimos que $z = h(\vec{x})$ es la ecuación del hiperplano afín tangente al grafo de f en $(\vec{a}, f(\vec{a}))$.

Nota 5.3.3. Si en la Definición 5.3.1 cambiamos las normas en $\vec{E}$ y $\vec{F}$ por otras equivalentes, de (5.3.4) se desprende fácilmente que f sigue siendo diferenciable en $\vec{a}$ y, su diferencial en ese punto es el mismo.

Teorema 5.3.2. *Si f es una función definida en un abierto A de un e.v.n $\vec{E}$, con valores en un e.v.n $\vec{F}$, diferenciable en $\vec{a} \in A$, entonces ella es continua en $\vec{a}$, parcialmente derivable en $\vec{a}$ y, se tiene la igualdad*

$$Df(\vec{a})(\vec{v}) = Df(\vec{a}; \vec{v}) \qquad \textit{para todo } \vec{v} \in \vec{E}. \tag{5.3.5}$$

Demostración. Usando la expresión (5.3.4) y haciendo el cambio de variable $\vec{x} = \vec{a} + \vec{\delta}$, obtenemos que para todo $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\begin{aligned} \|\vec{x} - \vec{a}\| \leq \eta \ &\Rightarrow \ \|f(\vec{x}) - f(\vec{a}) - Df(\vec{a})(\vec{x} - \vec{a})\| \leq \varepsilon \|\vec{x} - \vec{a}\| \\ &\Rightarrow \ \|f(\vec{x}) - f(\vec{a})\| - \|Df(\vec{a})(\vec{x} - \vec{a})\| \leq \varepsilon \|\vec{x} - \vec{a}\| \end{aligned}$$

y como $Df(\vec{a})$ es una función lineal continua, aplicando la desigualdad (2.7.1) obtenemos

$$\|f(\vec{x}) - f(\vec{a})\| \leq (\|Df(\vec{a})\| + \varepsilon)\|\vec{x} - \vec{a}\| \quad \forall \vec{x} \in B(\vec{a}, \eta)$$

lo que implica que f es continua en $\vec{a}$.

Ahora, dado $\vec{v} \in \vec{E}$ no nulo, para $\vec{\delta} = t\vec{v}$ la relación (5.3.3) implica que

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(\vec{a}+t\vec{v}) - f(\vec{a}) - Df(\vec{a})(t\vec{v})}{|t|\|\vec{v}\|} = \vec{0}$$

multiplicando por $\|\vec{v}\|$ y usando la linealidad de $Df(\vec{a})$ se obtiene

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(\vec{a}+t\vec{v}) - f(\vec{a})}{t} = Df(\vec{a})(\vec{v})$$

que prueba que f es parcialmente derivable en $\vec{a}$ y que se tiene la igualdad (5.3.5).

El resultado para $\vec{v} = 0$ es evidente. $\square$

Nota 5.3.4. La fórmula (5.3.5) nos da la relación que existe entre la diferencial de una función y su derivada parcial. Su importancia es crucial pues constituye la única forma de calcular la diferencial de una función en un punto. En el teorema también demostramos que si la diferencial de una función existe en un punto de su dominio implica que ella es continua y parcialmente derivable en ese punto. Para apreciar cuanto más fuerte es la diferenciabilidad que la derivabilidad parcial, en el próximo ejemplo veremos que esta última no implica ni siquiera la continuidad.

Ejemplo 5.3.2. La función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x_1, x_2) := \frac{x_1 x_2^3}{x_1^2 + x_2^6}$ si $(x_1, x_2) \neq (0,0)$ y $f(0,0) := 0$, no es continua en $(0,0)$ pero si es parcialmente derivable. En efecto, $f(k^{-1}, k^{-\frac{1}{3}}) = \frac{k^{-2}}{k^{-2}+k^{-2}} = \frac{1}{2} \neq f(0,0)$ para todo $k \in \mathbb{N}$, lo que muestra la discontinuidad de f en $(0,0)$. Por otra parte $Df(\vec{0}; \vec{v}) = \lim\limits_{t\to 0} \frac{f(t\vec{v}) - f(\vec{0})}{t} = \lim\limits_{t\to 0} \frac{t^4 v_1 v_2^3}{t^3 v_1^2 + t^7 v_2^6} = \lim\limits_{t\to 0} \frac{t v_1 v_2^3}{v_1^2 + t^4 v_2^6} = 0$ para todo $\vec{v} \in \vec{E}$, lo que muestra que f es parcialmente derivable en $(0,0)$.

Teorema 5.3.3. *Dadas dos funciones f y g definidas en un abierto A de un e.v.n $\vec{E}$, con valores en un e.v.n $\vec{F}$, diferenciables en $\vec{a} \in A$, entonces*

(i) *$f + g$ es diferenciable en $\vec{a}$ y se tiene*

$$D[f+g](\vec{a}) = Df(\vec{a}) + Dg(\vec{a}); \tag{5.3.6}$$

(ii) *Para $\lambda \in \mathbb{R}, \lambda f$ es diferenciable en $\vec{a}$ y se tiene*

$$D[\lambda f](\vec{a}) = \lambda Df(\vec{a}); \tag{5.3.7}$$

(iii) *Si* $\vec{F} = \mathbb{R}$, *entonces* $f \cdot g$ *es diferenciable en* $\vec{a}$ *y se tiene*

$$D[f \cdot g](\vec{a}) = g(\vec{a})Df(\vec{a}) + f(\vec{a})Dg(\vec{a}); \tag{5.3.8}$$

(iv) *Si* $\vec{F} = \mathbb{R}$ *y* $f(\vec{a}) \neq 0$, *entonces* $1/f$ *es diferenciable en* $\vec{a}$ *y se tiene*

$$D\left[\frac{1}{f}\right](\vec{a}) = -\frac{1}{f^2(\vec{a})}Df(\vec{a}). \tag{5.3.9}$$

Demostración. Una forma directa de demostrar este teorema consiste en verificar para cada uno de los cuatro casos la igualdad (5.3.3). De este modo se demuestra simultáneamente la diferenciabilidad de cada una de las cuatro funciones y su respectiva fórmula.

Los casos (i) y (ii) son los más fáciles y los dejamos como ejercicio.

(iii)

$$\lim_{\substack{\vec{\delta}\to\vec{0}\\ \vec{\delta}\neq\vec{0}}} \frac{(f\cdot g)(\vec{a}+\vec{\delta}) - (f\cdot g)(\vec{a}) - [g(\vec{a})Df(\vec{a}) + f(\vec{a})Dg(\vec{a})](\vec{\delta})}{\|\vec{\delta}\|} =$$

$$\lim \frac{f(\vec{a}+\vec{\delta})[g(\vec{a}+\vec{\delta}) - g(\vec{a}) - Dg(\vec{a})(\vec{\delta})]}{\|\vec{\delta}\|} + \frac{g(\vec{a})[f(\vec{a}+\vec{\delta}) - f(\vec{a}) - Df(\vec{a})(\vec{\delta})]}{\|\vec{\delta}\|}$$

$$+\frac{[f(\vec{a}+\vec{\delta}) - f(\vec{a})]Dg(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|} = \lim f(\vec{a}+\vec{\delta}) \lim \frac{g(\vec{a}+\vec{\delta}) - g(\vec{a}) - Dg(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|}$$

$$+g(\vec{a})\lim \frac{f(\vec{a}+\vec{\delta}) - f(\vec{a}) - Df(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|} + \lim[f(\vec{a}+\vec{\delta}) - f(\vec{a})]\frac{Dg(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|} = 0.$$

El último límite de la expresión anterior es nulo debido a que la continuidad de la función lineal $Dg(\vec{a})$ implica que el cuociente $\frac{Dg(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|}$ es acotado y, la diferenciabilidad de f en $\vec{a}$ implica su continuidad.

(iv)

$$\lim_{\substack{\vec{\delta}\to\vec{0}\\ \vec{\delta}\neq\vec{0}}} \frac{\frac{1}{f(\vec{a}+\vec{\delta})} - \frac{1}{f(\vec{a})} + \frac{Df(\vec{a})(\vec{\delta})}{f(\vec{a})^2}}{\|\vec{\delta}\|} = \lim \frac{f(\vec{a})^2 - f(\vec{a}+\vec{\delta})f(\vec{a}) + f(\vec{a}+\vec{\delta})Df(\vec{a})(\vec{\delta})}{f(\vec{a}+\vec{\delta})f(\vec{a})^2\|\vec{\delta}\|} =$$

$$\lim \frac{1}{f(\vec{a}+\vec{\delta})f(\vec{a})^2} \cdot \frac{Df(\vec{a})(\vec{\delta})[f(\vec{a}+\vec{\delta}) - f(\vec{a})] - f(\vec{a})[f(\vec{a}+\vec{\delta}) - f(\vec{a}) - Df(\vec{a})(\vec{\delta})]}{\|\vec{\delta}\|} =$$

$$\frac{1}{f(\vec{a})^3}\,\text{lím}[f(\vec{a}+\vec{\delta})-f(\vec{a})]\frac{Df(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|}-\frac{1}{f(\vec{a})^2}\,\text{lím}\,\frac{f(\vec{a}+\vec{\delta})-f(\vec{a})-Df(\vec{a})(\vec{\delta})}{\|\vec{\delta}\|}=0$$

□

Teorema 5.3.4. *Si f es una función definida en un abierto A de un e.v.n $\vec{E}$, con valores en $\mathbb{R}^m$ y, si denotamos $f_1, ..., f_m$ las funciones componentes de f, entonces f es diferenciable en $\vec{a} \in A$ si y solo si cada una de las m funciones f_i es diferenciable en $\vec{a}$ y, en ese caso se tendrá*

$$Df(\vec{a}) = (Df_1(\vec{a}), ..., Df_m(\vec{a})). \tag{5.3.10}$$

Demostración. La equivalencia de la diferenciabilidad de f con la de sus funciones componentes y la fórmula (5.3.10) se obtienen en forma directa del Teorema 2.4.2 usando la igualdad (5.3.3). □

Nota 5.3.5. Las cinco fórmulas que se dan en los dos teoremas anteriores, se pueden obtener fácilmente a partir de la fórmula (5.3.5) y de las respectivas fórmulas de los teoremas 5.2.2 y 5.2.3. No lo hicimos así debido a que previamente había que demostrar la diferenciabilidad de cada función.

Nota 5.3.6. Dados m e.v.n. $\vec{F}_1, ..., \vec{F}_m$ el teorema anterior se generaliza fácilmente, de acuerdo a la Nota 2.4.1, al caso en que f toma sus valores en el e.v.n. $\vec{F}_1 \times ... \times \vec{F}_m$.

Teorema 5.3.5. *Si f es una función definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n $\vec{F}$, diferenciable en un punto $\vec{a} \in A$, entonces se tiene la fórmula*

$$Df(\vec{a})(\vec{v}) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(\vec{a}) \tag{5.3.11}$$

para todo $\vec{v} = \sum_{i=1}^{n} v_i \vec{e}_i \in \mathbb{R}^n$, donde $\vec{e}_1, ..., \vec{e}_n$ es la base canónica de $\mathbb{R}^n$ que usamos en la Definición 5.2.3.

Demostración. Es una consecuencia inmediata del Teorema 5.3.2. □

Definición 5.3.2. Si f es una función definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, diferenciable en un punto $\vec{a} \in A$, se llama gradiente de f en $\vec{a}$ al vector

$$\nabla f(\vec{a}) := \left(\frac{\partial f}{\partial x_1}(\vec{a}), ..., \frac{\partial f}{\partial x_n}(\vec{a})\right) \in \mathbb{R}^n. \tag{5.3.12}$$

Nota 5.3.7. Con la definición anterior, cuando $\vec{F} = \mathbb{R}$, la fórmula (5.3.11) se escribe

$$Df(\vec{a})(\vec{v}) = \langle \nabla f(\vec{a}), \vec{v} \rangle \tag{5.3.13}$$

donde $\langle \cdot, \cdot \rangle$ denota al producto interno usual en $\mathbb{R}^n$. De este modo, vemos que el gradiente de una función f diferenciable en $\vec{a}$ es el vector asociado a la función lineal $Df(\vec{a}) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, según el Teorema 4.7.1.

Con esta notación, la aproximación de primer orden de f en $\vec{a}$, definida en la Nota 5.3.2, se escribe

$$h(x) = f(\vec{a}) + \langle \nabla f(\vec{a}), \vec{x} - \vec{a} \rangle \tag{5.3.14}$$

y la ecuación del hiperplano afín tangente al grafo de f en $(\vec{a}, f(\vec{a}))$, será

$$z = \langle \nabla f(\vec{a}), \vec{x} - \vec{a} \rangle + f(\vec{a}). \tag{5.3.15}$$

Denotando ahora $\langle \cdot, \cdot \rangle$ al producto interno en $\mathbb{R}^n \times \mathbb{R}$, esta ecuación se escribe

$$\langle (\nabla f(\vec{a}), -1), (\vec{x}, z) - (\vec{a}, f(\vec{a})) \rangle = 0$$

lo que muestra que el hiperplano afín en cuestión es aquel que pasa por $(\vec{a}, f(\vec{a}))$ y es paralelo al hiperplano ortogonal al vector $(\nabla f(\vec{a}), -1)$. Por esta razón, se dice que el vector $(\nabla f(\vec{a}), -1)$, trasladado a $(\vec{a}, f(\vec{a}))$, es ortogonal al grafo de f en $(\vec{a}, f(\vec{a}))$.

Nota 5.3.8. Del Teorema 4.7.1 vemos que la fórmula (5.3.13) puede generalizarse a toda función f definida en un abierto A de un espacio de Hilbert $\vec{E}$ con valores en $\mathbb{R}$. En efecto, si f es diferenciable en $\vec{a} \in A$, como $Df(\vec{a})$ es una función lineal continua de $\vec{E}$ en $\mathbb{R}$, definimos el gradiente de f en $\vec{a}$ como el único elemento $\nabla f(\vec{a}) \in \vec{E}$ que verifica

$$Df(\vec{a})(\vec{v}) = \langle \nabla f(\vec{a}), \vec{v} \rangle \quad \text{para todo } \vec{v} \in \vec{E} \tag{5.3.16}$$

y con esta definición de gradiente, se deduce fácilmente del teorema anterior que, cuando $\vec{E} = \mathbb{R}^n$ y $\vec{F} = \mathbb{R}, \nabla f(\vec{a})$ está dado por la igualdad (5.3.12).

Definición 5.3.3. Si f es una función definida en un abierto A de $\mathbb{R}^n$, con valores en $\mathbb{R}^m$, si denotamos $f_1, ..., f_m$ sus m funciones componentes y suponemos f diferenciable en $\vec{a} \in A$, entonces llamamos Jacobiano de f en $\vec{a}$ a la matriz de $m \times n$

$$J_f(\vec{a}) = \left(\frac{\partial f_i}{\partial x_j}(\vec{a}) \right) \tag{5.3.17}$$

donde $1 \leq i \leq m$ y $1 \leq j \leq n$.

Nota 5.3.9. Con la definición anterior, cuando $\vec{F} = \mathbb{R}^m$, la fórmula (5.3.11) se escribe, usando notación matricial,

$$Df(\vec{a})(\vec{v}) = J_f(\vec{a})\vec{v}. \tag{5.3.18}$$

El Jacobiano de una función f en $\vec{a}$ es la matriz asociada a la función lineal $Df(\vec{a}) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, según lo estudiado en el curso de algebra lineal. Con esta notación, la aproximación de primer orden de f en $\vec{a}$, definida en la Nota 5.3.2 se escribe

$$h(\vec{x}) = f(\vec{a}) + J_f(\vec{a})(\vec{x} - \vec{a}). \tag{5.3.19}$$

5.3.1 Teorema del Valor Medio

Teorema 5.3.6. *Sea f una función definida en un abierto A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Si f es diferenciable en todo punto de un segmento $[\vec{a}, \vec{b}] \subset A$, entonces existe $\vec{c} \in]\vec{a}, \vec{b}[$ tal que*

$$f(\vec{b}) - f(\vec{a}) = Df(\vec{c})(\vec{b} - \vec{a}). \tag{5.3.20}$$

Demostración. Si definimos la función $\phi(t) = f(\vec{a} + t(\vec{b} - \vec{a}))$ para $t \in [0, 1]$ y aplicamos el teorema del valor medio para funciones de una variable real, obtenemos que existe $\eta \in]0, 1[$ tal que $\phi(1) - \phi(0) = \phi'(\eta)$ lo que corresponde exactamente a la fórmula (5.3.20) con $\vec{c} = \vec{a} + \eta(\vec{b} - \vec{a})$. $\square$

Nota 5.3.10. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, de la fórmula (5.3.13) vemos que (5.3.20) puede escribirse

$$f(\vec{b}) - f(\vec{a}) = \langle \nabla f(\vec{c}), \vec{b} - \vec{a} \rangle.$$

Ejemplo 5.3.3. Demos un ejemplo que muestre que la fórmula (5.3.20) no es en general válida si f toma sus valores en un e.v.n. $\vec{F}$. Sea $f : \mathbb{R} \to \mathbb{R}^2$ definida por $f(t) := (cos(t), sen(t))$ y sean $a := 0$ y $b := 2\pi$. Es fácil constatar que $f(b) - f(a) = (0, 0)$ y $Df(c)(b - a) = 2\pi(-sen(c), cos(c))$, lo que muestra que no existe $c \in [0, 2\pi]$ tal que se tenga la igualdad (5.3.20).

Teorema 5.3.7. *Sea f una función definida en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Si f es diferenciable en todo punto de un segmento $[\vec{a}, \vec{b}] \subset A$ y si L es una constante que verifica $L \geq \|Df(\vec{x})\|$ para todo $\vec{x} \in [\vec{a}, \vec{b}]$, entonces se tiene la desigualdad*

$$\|f(\vec{b}) - f(\vec{a})\| \leq L\|\vec{b} - \vec{a}\|. \tag{5.3.21}$$

Demostración. Supongamos lo contrario, es decir, que existe $\delta > 0$ tal que$\|f(\vec{b}) - f(\vec{a})\| - L\|\vec{b}-\vec{a}\| = \delta$. Si definimos $\vec{m} = \frac{\vec{a}+\vec{b}}{2}$, puesto que $\|f(\vec{b}) - f(\vec{m})\| + \|f(\vec{m}) - f(\vec{a})\| \geq \|f(\vec{b}) - f(\vec{a})\|$ y $\|\vec{b}-\vec{a}\| = \|\vec{b}-\vec{m}\| + \|\vec{m}-\vec{a}\|$, se tendrá

$$\|f(\vec{b}) - f(\vec{m})\| - L\|\vec{b}-\vec{m}\| \geq \frac{\delta}{2}$$

o bien

$$\|f(\vec{m}) - f(\vec{a})\| - L\|\vec{m}-\vec{a}\| \geq \frac{\delta}{2}.$$

Si se tiene la primera de estas desigualdades definimos $\vec{b}_1 := \vec{b}$ y $\vec{a}_1 := \vec{m}$, en caso contrario definimos $\vec{b}_1 := \vec{m}$ y $\vec{a}_1 := \vec{a}$. Aplicando sucesivamente el mismo procedimiento obtenemos una sucesión de intervalos encajonados $[\vec{a}_n, \vec{b}_n]$ con $\|\vec{b}_n - \vec{a}_n\| = \frac{\|\vec{b}-\vec{a}\|}{2^n}$ y tales que

$$\|f(\vec{b}_n) - f(\vec{a}_n)\| - L\|\vec{b}_n - \vec{a}_n\| \geq \frac{\delta}{2^n}.$$

Puesto que $\|\vec{a}_n - \vec{b}_n\| \to 0$ sabemos que las sucesiones $\{\vec{a}_n\}$ y $\{\vec{b}_n\}$ deben converger a un mismo $\vec{w} \in [\vec{a}, \vec{b}]$. Por otra parte por ser f diferenciable en $\vec{w}$, de la desigualdad anterior podemos escribir

$$\begin{aligned}
\frac{\delta}{2^n} &\leq \|f(\vec{b}_n) - f(\vec{w}) - (f(\vec{a}_n) - f(\vec{w}))\| - L\|\vec{b}_n - \vec{a}_n\| \\
&= \|Df(\vec{w})(\vec{b}_n - \vec{w}) + o(\vec{b}_n - \vec{w}) - Df(\vec{w})(\vec{a}_n - \vec{w}) - o(\vec{a}_n - \vec{w})\| - L\|\vec{b}_n - \vec{a}_n\| \\
&\leq \|Df(\vec{w})\|\|\vec{b}_n - \vec{a}_n\| + \|o(\vec{b}_n - \vec{w})\| + \|o(\vec{a}_n - \vec{w})\| - L\|\vec{b}_n - \vec{a}_n\|
\end{aligned}$$

y como $\|\vec{b}_n - \vec{w}\| < \|\vec{b}_n - \vec{a}_n\|$ y $\|\vec{a}_n - \vec{w}\| < \|\vec{b}_n - \vec{a}_n\|$, dividiendo la desigualdad anterior por $\|\vec{b}_n - \vec{a}_n\|$ podemos escribir

$$\frac{\delta}{\|\vec{b}-\vec{a}\|} \leq \|Df(\vec{w})\| + \frac{\|o(\vec{w} - \vec{b}_n)\|}{\|\vec{w} - \vec{b}_n\|} + \frac{\|o(\vec{w} - \vec{a}_n\|}{\|\vec{w} - \vec{a}_n\|} - L$$

y tomando límite sobre n obtenemos una contradicción con la hipótesis $\|Df(\vec{x})\| \leq L$ para todo $\vec{x} \in [\vec{a}, \vec{b}]$. □

Nota 5.3.11. Del teorema anterior se deduce que si f es diferenciable en todo punto de un conjunto convexo $C \subset A$ y si $L \geq \|Df(\vec{x})\|$ para todo $x \in C$, entonces $\|f(\vec{x}) - f(\vec{y})\| \leq L\|\vec{x} - \vec{y}\|$ para todo $\vec{x}, \vec{y} \in C$.

Definición 5.3.4. Un conjunto C en un e.v.n. $\vec{E}$ se dirá conexo si no existen conjuntos abiertos no vacíos C_1 y C_2 en $\vec{E}$ que intersecten C, tales que $C_1 \cap C_2 = \phi$ y $C \subset C_1 \cup C_2$.

Teorema 5.3.8. *Sea f una función diferenciable definida en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Si el diferencial de f es nulo en todo punto de un conjunto abierto conexo $C \subset A$, entonces la función f será constante en C.*

Demostración. Demostremos primero que f es constante en toda bola $B(\vec{x}_0, \delta) \subset C$. Sea $\vec{z} \in B(\vec{x}_0, \delta)$, puesto que $[\vec{x}_0, \vec{z}] \subset C$ y $Df(\vec{x}) = 0$ para todo $\vec{x} \in [\vec{x}_0, \vec{z}]$, del teorema anterior concluímos que $\|f(\vec{z}) - f(\vec{x}_0)\| \leq 0$, lo que equivale a decir que $f(\vec{z}) = f(\vec{x}_0)$.

Sea ahora $\vec{a} \in C, C_1 = \{\vec{x} \in C : f(\vec{x}) = f(\vec{a})\}$ y $C_2 = \{\vec{x} \in C : f(\vec{x}) \neq f(\vec{a})\}$. Puesto que la función f es continua (ver Teorema 5.3.2) y C es un conjunto abierto, es fácil demostrar que C_2 es un conjunto abierto.

Mostremos finalmente que C_1 es también un conjunto abierto. Sea $\vec{x}_0 \in C_1$ y $B(\vec{x}_0, \delta) \subset C$, de la primera parte de esta demostración concluímos que f es constante en $B(\vec{x}_0, \delta)$, y se tendrá entonces $f(\vec{x}) = f(\vec{a})$ para todo $\vec{x} \in B(\vec{x}_0, \delta)$, lo que implica que $B(\vec{x}_0, \delta) \subset C_1$. Lo anterior muestra que C_1 es abierto. Como $C_1 \cap C_2 = \phi$, como $C \subset C_1 \cup C_2$ y como $C_1 \neq \phi$ (en efecto, $\vec{a} \in C_1$), del hecho que C es conexo concluímos que $C_2 = \phi$, es decir, $C = C_1$. □

5.4 Funciones de clase $\mathcal{C}^1$

Definición 5.4.1. Una función f definida en un abierto A de $\mathbb{R}^n$, con valores en un e.v.n $\vec{F}$ se dirá de clase $\mathcal{C}^1$ si las n derivadas parciales de f con respecto a $x_1, ..., x_n$ (ver Nota 5.2.3) existen y son continuas.

Teorema 5.4.1. *Si f es una función de clase $\mathcal{C}^1$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n $\vec{F}$, entonces ella es diferenciable.*

Demostración. Con el único objeto de simplificar la notación, haremos la demostración para el caso en que $n = 2$ y usaremos la norma del máximo.

Dado un elemento cualquiera $\vec{a} := a_1\vec{e}_1 + a_2\vec{e}_2 \in A$, vamos a demostrar que f es diferenciable en $\vec{a}$. De acuerdo al Teorema 5.3.5, debemos probar entonces que la función lineal $\ell(\vec{\delta}) := \delta_1 \frac{\partial f}{\partial x_1}(\vec{a}) + \delta_2 \frac{\partial f}{\partial x_2}(\vec{a})$ es la diferencial de f en $\vec{a}$.

Dado $\eta > 0$ tal que $B(\vec{a}, \eta) \subset A$, escribamos para $\vec{\delta} \in \mathbb{R}^2$ con $\|\vec{\delta}\| \leq \eta$ la desigualdad

$$\frac{\|f(\vec{a} + \vec{\delta}) - f(\vec{a}) - \ell(\vec{\delta})\|}{\|\vec{\delta}\|} \leq \frac{\|f(a_1 + \delta_1, a_2 + \delta_2) - f(a_1, a_2 + \delta_2) - \delta_1 \frac{\partial f}{\partial x_1}(\vec{a})\|}{\|\vec{\delta}\|}$$

$$+\frac{\|f(a_1,a_2+\delta_2)-f(a_1,a_2)-\delta_2\frac{\partial f}{\partial x_2}(\vec{a})\|}{\|\vec{\delta}\|}.$$

Para concluir debemos probar que el límite cuando $\vec{\delta}\to\vec{0}$ de cada uno de los dos sumandos de la derecha de esta desigualdad es cero.

La desigualdad

$$\frac{\|f(a_1,a_2+\delta_2)-f(a_1,a_2)-\delta_2\frac{\partial f}{\partial x_2}(\vec{a})\|}{\|\vec{\delta}\|}\leq\|\frac{f(\vec{a}+\delta_2\vec{e}_2)-f(\vec{a})}{\delta_2}-\frac{\partial f}{\partial x_2}(\vec{a})\|$$

y la definición de $\frac{\partial f}{\partial x_2}(\vec{a})$, nos muestra que el segundo de estos límites, es cero.

Definamos ahora la función $\phi(t):=f(a_1+t,a_2+\delta_2)-f(a_1,a_2+\delta_2)-t\frac{\partial f}{\partial x_1}(\vec{a})$. Como f es de clase C^1, es claro que ϕ es diferenciable en todo punto del intervalo J (donde $J=[0,\delta_1]$ si $\delta_1>0$, $=[\delta_1,0]$ si $\delta_1<0$) y $D\phi(t)(v)=[\frac{\partial f}{\partial x_1}(a_1+t,a_2+\delta_2)-\frac{\partial f}{\partial x_1}(\vec{a})]v$. Escribiendo entonces para la función ϕ la desigualdad $\|\phi(\delta_1)-\phi(0)\|\leq L|\delta_1|$ con $L=\max\limits_{t\in J}\|D\phi(t)\|$, dada por el Teorema 5.3.7, obtenemos

$$\|f(a_1+\delta_1,a_2+\delta_2)-f(a_1,a_2+\delta_2)-\delta_1\frac{\partial f}{\partial x_1}(\vec{a})\|\leq|\delta_1|\max_{t\in J}\|\frac{\partial f}{\partial x_1}(a_1+t,a_2+\delta_2)-\frac{\partial f}{\partial x_1}(\vec{a})\|.$$

Puesto que $\frac{\partial f}{\partial x_1}$ es una función continua en $\vec{a}$ y que $\frac{|\delta_1|}{\|\vec{\delta}\|}\leq 1$, dividiendo por $\|\vec{\delta}\|$ esta desigualdad, vemos que el lado derecho tiende a cero cuando $\vec{\delta}\to\vec{0}$, lo que nos permite concluir nuestra demostración. □

Teorema 5.4.2. *Si f es una función de clase $\mathcal{C}^1$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$, entonces f es continua.*

Demostración. Es una consecuencia inmediata del teorema anterior y del Teorema 5.3.2. □

Definición 5.4.2. Una función definida en un abierto A de un e.v.n $\vec{E}$, con valores en un e.v.n $\vec{F}$ se dirá continuamente diferenciable si ella es diferenciable en todo punto de A y si la función $Df:A\to\mathcal{L}(\vec{E},\vec{F})$ es continua.

Teorema 5.4.3. *Si f es una función de clase $\mathcal{C}^1$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n $\vec{F}$, entonces ella es continuamente diferenciable.*

Demostración. Del Teorema 5.4.1 sabemos que f es diferenciable en $\vec{a}$. Verifiquemos ahora la continuidad de Df en $\vec{a} \in A$. De (2.7.2) y (5.3.11)

$$\begin{aligned}
\|Df(\vec{a}) - Df(\vec{x})\| &= \sup_{\vec{v} \neq \vec{0}} \frac{\|[Df(\vec{a}) - Df(\vec{x})](\vec{v})\|}{\|\vec{v}\|} = \\
&= \sup_{\vec{v} \neq \vec{0}} \frac{\|[\frac{\partial f}{\partial x_1}(\vec{a}) - \frac{\partial f}{\partial x_1}(\vec{x})]v_1 + ... + [\frac{\partial f}{\partial x_n}(\vec{a}) - \frac{\partial f}{\partial x_n}(\vec{x})]v_n\|}{\|\vec{v}\|} \\
&\leq \sup_{\vec{v} \neq \vec{0}} \frac{\|(\frac{\partial f}{\partial x_1}(\vec{a}) - \frac{\partial f}{\partial x_1}(\vec{x}))v_1\|}{\|\vec{v}\|} + ... + \sup_{\vec{v} \neq \vec{0}} \frac{\|(\frac{\partial f}{\partial x_n}(\vec{a}) - \frac{\partial f}{\partial x_n}(\vec{x}))v_n\|}{\|\vec{v}\|} \\
&\leq \|\frac{\partial f}{\partial x_1}(\vec{a}) - \frac{\partial f}{\partial x_1}(\vec{x})\| + ... + \|\frac{\partial f}{\partial x_n}(\vec{a}) - \frac{\partial f}{\partial x_n}(\vec{x})\|
\end{aligned}$$

por lo tanto, haciendo tender $\vec{x}$ a $\vec{a}$ y puesto que las funciones $\frac{\partial f}{\partial x_i}(\cdot)$ son continuas en $\vec{a}$, obtenemos

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

lo que es equivalente a $\lim\limits_{\vec{x} \to \vec{a}} Df(\vec{x}) = Df(\vec{a})$ que de acuerdo al Teorema 2.3.1 significa que Df es continua en $\vec{a}$. $\square$

Teorema 5.4.4. *Si f es una función continuamente diferenciable, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$, entonces f es de clase $\mathcal{C}^1$.*

Demostración. Vamos a demostrar la continuidad de la función $\frac{\partial f}{\partial x_i} : A \to \vec{F}$ en un punto $\vec{a} \in A$. Del Teorema 5.3.2 deducimos que

$$\begin{aligned}
\frac{\partial f}{\partial x_i}(\vec{x}) - \frac{\partial f}{\partial x_i}(\vec{a}) &= Df(\vec{x})(\vec{e}_i) - Df(\vec{a})(\vec{e}_i) \\
&= [Df(\vec{x}) - Df(\vec{a})](\vec{e}_i)
\end{aligned}$$

y de la desigualdad (2.7.3), aplicada a la función lineal continua $[Df(\vec{x}) - Df(\vec{a})]$, vemos que

$$\|\frac{\partial f}{\partial x_i}(\vec{x}) - \frac{\partial f}{\partial x_i}(\vec{a})\| \leq \|Df(\vec{x}) - Df(\vec{a})\|\|\vec{e}_i\|. \qquad (*)$$

Como por hipótesis la función $Df : A \to \mathcal{L}(\mathbb{R}^n, F)$ es continua en $\vec{a} \in A$, del Teorema 2.3.1 sabemos que $\lim\limits_{\vec{x} \to \vec{a}} Df(\vec{x}) = Df(\vec{a})$, lo que implica a partir de (*) que

$$\lim_{\vec{x} \to \vec{a}} \frac{\partial f}{\partial x_i}(\vec{x}) = \frac{\partial f}{\partial x_i}(\vec{a})$$

y de acuerdo al mismo Teorema 2.3.1, concluímos que $\frac{\partial f}{\partial x_i}$ es una función continua en $\vec{a}$. $\square$

Nota 5.4.1. Si f es una función definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$, entonces los teoremas 5.3.2, 5.4.1, 5.4.3 y 5.4.4 se resumen en el siguiente diagrama

$$\begin{aligned} f \text{ es de clase } \mathcal{C}^1 &\Leftrightarrow f \text{ es continuamente diferenciable} \\ &\Rightarrow f \text{ es diferenciable} \\ &\Rightarrow f \text{ es parc. derivable} \\ &\Rightarrow f \text{ es continua.} \end{aligned}$$

Teorema 5.4.5. *Si f y g son dos funciones continuamente diferenciables definidas en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, entonces*

(i) *$f+g$ es continuamente diferenciable.*

(ii) *Si $\lambda \in \mathbb{R}, \lambda f$ es continuamente diferenciable.*

(iii) *Si $\vec{F} = \mathbb{R}, f \cdot g$ es continuamente diferenciable.*

(iv) *Si $\vec{F} = \mathbb{R}$ y $f(\vec{x}) \neq 0$ para todo $\vec{x} \in A, 1/f$ es continuamente diferenciable.*

Demostración. Del Teorema 5.3.3 vemos que las cuatro funciones del enunciado serán diferenciables. Por otra parte las fórmulas (5.3.6) a (5.3.9) muestran que para todo $\vec{x} \in A$

(i) $D[f+g](\vec{x}) = Df(\vec{x}) + Dg(\vec{x})$

(ii) $D[\lambda f](\vec{x}) = \lambda Df(\vec{x})$

(iii) $D[f \cdot g](\vec{x}) = g(\vec{x})Df(\vec{x}) + f(\vec{x})Dg(\vec{x})$.

(iv) $D[1/f](\vec{x}) = -\frac{1}{f(\vec{x})^2}Df(\vec{x})$.

Dado que por hipótesis las funciones f, Df, g y Dg son continuas, de las fórmulas anteriores deducimos que $D[f+g], D[\lambda f], D[f \cdot g]$ y $D[\frac{1}{f}]$ son continuas. Lo anterior nos permite concluir que $f+g, \lambda f, f \cdot g$ y $1/f$ son continuamente diferenciables. $\square$

Teorema 5.4.6. *Si f es una función definida en un abierto A de un e.v.n $\vec{E}$ con valores en $\mathbb{R}^m$ y, si denotamos $f_1, ..., f_m$ las funciones componentes de f, entonces ella será continuamente diferenciable si y solo si sus m funciones componentes son continuamente diferenciables.*

Demostración. Del Teorema 5.3.4 vemos que f es diferenciable si y solo si sus m funciones componentes lo son. Por otra parte la fórmula (5.3.10) muestra que

$$Df(\vec{x}) = (Df_1(\vec{x}), ..., Df_m(\vec{x}))$$

lo que significa, de acuerdo al Teorema 2.4.1, que Df es continua si y solo si las funciones $Df_1, ..., Df_m$ son continuas. Lo anterior nos permite concluir que f es continuamente diferenciable si y solo si las funciones $f_1, ..., f_m$ también lo son. $\square$

Nota 5.4.2. Dados m e.v.n. $\vec{F}_1, ..., \vec{F}_m$, el teorema anterior se generaliza fácilmente, de acuerdo a las notas 5.3.6 y 2.4.1, al caso en que f toma sus valores en el e.v.n. producto $\vec{F}_1 \times ... \times \vec{F}_m$.

Teorema 5.4.7. *Si f es una función continuamente diferenciable, definida en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$, entonces para todo $\vec{a} \in A$ y todo $\varepsilon > 0$ existe $\delta > 0$ tal que*

$$\vec{x}, \vec{y} \in B(\vec{a}, \delta) \Rightarrow \|f(\vec{x}) - f(\vec{y})\| \leq (\|Df(\vec{a})\| + \varepsilon)\|\vec{x} - \vec{y}\|. \tag{5.4.1}$$

La función f se dice entonces localmente Lipschitziana en A.

Demostración. Sea $\vec{a} \in A$ y $\varepsilon > 0$. Como Df es una función continua en $\vec{a}$, existirá $\delta > 0$ tal que

$$\|Df(\vec{z})\| \leq \|Df(\vec{a})\| + \varepsilon \text{ para todo } \vec{z} \in B(\vec{a}, \delta).$$

Como para todo $\vec{x}, \vec{y} \in B(\vec{a}, \delta)$ se tiene que $[\vec{x}, \vec{y}] \subset B(\vec{a}, \delta)$, del Teorema 5.3.7 obtenemos la desigualdad (5.4.1). $\square$

Teorema 5.4.8. *Sea f una función continuamente diferenciable, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$. Entonces ella es Lipschitziana en toda bola $B(\vec{a},r) \subset A$ (ver Definición 2.6.2), con constante de Lipschitz $L := \max_{\vec{z} \in B(\vec{a},r)} \|Df(\vec{z})\|$.*

Demostración. Del Teorema 1.5.4 y del Teorema 2.5.2 aplicado a la función Df en $B(\vec{a}, r)$, concluímos que L está bien definido. El Teorema 5.3.7 nos permite entonces concluir. $\square$

5.5 Composición de funciones diferenciables

Teorema 5.5.1. *Sea f una función de un abierto A de un e.v.n. $\vec{E}$ en un e.v.n. $\vec{F}$ y sea g una función de un abierto B del e.v.n. $\vec{F}$, en un e.v.n. $\vec{G}$ (suponemos que $f(A) \subset B$). Entonces si f y g son diferenciables en $\vec{a} \in A$ y $f(\vec{a}) \in B$ respectivamente, la función $h := g \ o\ f$ es diferenciable en $\vec{a}$ y se tiene la fórmula*

$$Dh(\vec{a}) = Dg(f(\vec{a})) \ \circ \ Df(\vec{a}). \tag{5.5.1}$$

Demostración. De acuerdo a la relación (5.3.4) debemos probar que dado $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\|\vec{\delta}\| \leq \eta \Rightarrow \|(g \circ f)(\vec{a} + \vec{\delta}) - (g \ \circ \ f)(\vec{a}) - [Dg(f(\vec{a})) \ \circ \ Df(\vec{a})](\vec{\delta})\| \leq \varepsilon \|\vec{\delta}\|. \qquad (*)$$

Para simplificar la notación escribamos $h := g \circ f$, $\vec{b} := f(\vec{a})$ y $\vec{v} := f(\vec{a} + \vec{\delta}) - f(\vec{a})$. Entonces

$$\begin{gathered}\|h(\vec{a} + \vec{\delta}) - h(\vec{a}) - [Dg(\vec{b}) \circ Df(\vec{a})](\vec{\delta})\| \leq \\ \|g(\vec{b} + \vec{v}) - g(\vec{b}) - Dg(\vec{b})(\vec{v})\| + \|Dg(\vec{b})(\vec{v}) - [Dg(\vec{b}) \circ Df(\vec{a})](\vec{\delta})\|\end{gathered}$$

y, como de la desigualdad (2.7.3) aplicada a la función lineal continua $Dg(\vec{b})$ se tiene

$$\begin{aligned}\|Dg(\vec{b})(\vec{v}) - [Dg(\vec{b}) \circ Df(\vec{a})](\vec{\delta})\| &= \|Dg(\vec{b})(\vec{v} - Df(\vec{a})(\vec{\delta}))\| \\ &\leq \|Dg(\vec{b})\| \ \ \|f(a + \vec{\delta}) - f(\vec{a}) - Df(\vec{a})(\vec{\delta})\|\end{aligned}$$

obtenemos la desigualdad

$$\begin{gathered}\|h(\vec{a} + \vec{\delta}) - h(\vec{a}) - [Dg(\vec{b}) \circ Df(\vec{a})](\vec{\delta})\| \leq \\ \|g(\vec{b} + \vec{v}) - g(\vec{b}) - Dg(\vec{b})(\vec{v})\| + \|Dg(\vec{b})\| \ \ \|f(\vec{a} + \vec{\delta}) - f(\vec{a}) - Df(\vec{a})(\vec{\delta})\|.\end{gathered} \tag{5.5.2}$$

Sea $\varepsilon > 0$. Como f es diferenciable en $\vec{a}$, de acuerdo al Teorema 5.3.2 (ver su demostración), ella será también Lipschitziana en $\vec{a}$. Existirá entonces $\eta_1 > 0$ tal que

$$\|\vec{\delta}\| \leq \eta_1 \Rightarrow \|f(\vec{a} + \vec{\delta}) - f(\vec{a}) - Df(\vec{a})(\vec{\delta})\| \leq \frac{\varepsilon}{2\|Dg(\vec{b})\|} \|\vec{\delta}\| \tag{5.5.3}$$

y

$$\|\vec{\delta}\| \leq \eta_1 \Rightarrow \|f(\vec{a} + \vec{\delta}) - f(\vec{a})\| \leq L\|\vec{\delta}\|. \tag{5.5.4}$$

Por otra parte, como g es diferenciable en $\vec{b}$, existirá $\eta_2 > 0$ tal que

$$\|\vec{v}\| \leq \eta_2 \Rightarrow \|g(\vec{b} + \vec{v}) - g(\vec{b}) - Dg(\vec{b})(\vec{v})\| \leq \frac{\varepsilon}{2L}\|\vec{v}\|. \tag{5.5.5}$$

De las relaciones (5.5.2), (5.5.3), (5.5.4) y (5.5.5) definiendo $\eta = \text{mín}\{\eta_1, \frac{\eta_2}{L}\}$ obtenemos directamente la desigualdad (*). $\square$

Nota 5.5.1. De acuerdo a la Definición 5.3.2 y a la Nota 5.3.9 vemos que si en el teorema anterior $\vec{E} = \mathbb{R}^n, \vec{F} = \mathbb{R}^m$ y $\vec{G} = \mathbb{R}^p$, entonces el jacobiano de la función $h = g \circ f$ en $\vec{a}$, será igual al producto de los jacobianos de las funciones g y f, esto es

$$J_h(\vec{a}) = J_g(f(\vec{a})) \cdot J_f(\vec{a}) \tag{5.5.6}$$

que corresponde a la fórmula (5.5.1) escrita matricialmente. Se verifica fácilmente que el elemento (i, j) de la matriz $J_h(\vec{a})$ (de p filas y n columnas) es $\sum_{k=1}^{m} \partial_k g_i(f(\vec{a}))\partial_j f_k(\vec{a})$ donde $\partial_j f_k(\vec{a})$ representa la derivada parcial en $\vec{a}$ de la función componente f_k con respecto a la j-ésima variable y, $\partial_k g_i(f(\vec{a}))$ representa la derivada parcial en $f(\vec{a})$ de la función componente g_i con respecto a la k-ésima variable.

De (5.5.6) y (5.3.18) vemos que para todo $\vec{v} \in \mathbb{R}^n$ se tendrá

$$Dh(\vec{a})(\vec{v}) = J_g(f(\vec{a}))J_f(\vec{a})^t\vec{v}^t \tag{5.5.7}$$

que también se escribe

$$Dh(\vec{a})(\vec{v}) = (\sum_{j=1}^{n}\sum_{k=1}^{m} v_j\partial_j f_k(\vec{a})\partial_k g_1(f(\vec{a})), \quad \ldots \quad , \sum_{j=1}^{n}\sum_{k=1}^{m} v_j\partial_j f_k(\vec{a})\partial_k g_p(f(\vec{a}))).$$

En particular, para $\vec{v} = \vec{e}_j \in \mathbb{R}^n$, se tendrá para todo $j = 1, ..., n$ e $i = 1, ..., p$ la fórmula

$$\partial_j h_i(\vec{a}) = \sum_{k=1}^{m} \partial_j f_k(\vec{a})\partial_k g_i(f(\vec{a})). \tag{5.5.8}$$

Esta fórmula se llama usualmente regla de la cadena para el cálculo de las derivadas parciales de la función $h = g \circ f$.

Teorema 5.5.2. *Si suponemos que las funciones f y g del teorema anterior son continuamente diferenciables, entonces $h := f$ o g también será continuamente diferenciable.*

Demostración. Es una consecuencia inmediata de la fórmula (5.5.1) y del Teorema 2.2.2, del que se deduce que por ser Dg, f y Df funciones continuas, entonces la función compuesta Dh también será continua. □

5.6 Diferencial Parcial

En esta sección vamos a introducir la noción de diferencial parcial de una función, que en cierto sentido generaliza la noción de derivada parcial y, daremos después un resultado que generaliza el Teorema 5.3.5.

Definición 5.6.1. Dados $n+1$ e.v.n. $\vec{E}_1, ..., \vec{E}_n, \vec{F}$, una función f definida en un abierto A del e.v.n. $\vec{E}_1 \times ... \times \vec{E}_n$ con valores en $\vec{F}$ y $\vec{a} \in A$, denotaremos por $D_j f(\vec{a})$ (para $j = 1, ..., n$) al diferencial en $\vec{a}_j \in \vec{E}_j$ ($\vec{a} = (\vec{a}_1, ..., \vec{a}_n) \in \vec{E}_1 \times ... \times \vec{E}_n$) de la función $f(\vec{a}_1, ..., \vec{a}_{j-1}, \cdot, \vec{a}_{j+1}, ..., \vec{a}_n)$, cuando existe. Se tiene entonces que $D_j f(\vec{a}) \in \mathcal{L}(\vec{E}_j, \vec{F})$.

Al diferencial $D_j f(\vec{a})$ lo llamamos diferencial parcial de f en $\vec{a}$ respecto a la variable j.

Nota 5.6.1. Con los datos de la definición anterior definimos las tres funciones $p_j : \vec{E}_1 \times ... \times \vec{E}_n \to \vec{E}_j,\ i_j : \vec{E}_j \to \vec{E}_1 \times ... \times \vec{E}_n$ e $I_j : \vec{E}_j \to \vec{E}_1 \times ... \times \vec{E}_n$ por

$$p_j(\vec{x}_1, ..., \vec{x}_n) := \vec{x}_j \tag{5.6.1}$$

$$i_j(\vec{x}_j) := (\vec{0}, ..., \vec{0}, \vec{x}_j, \vec{0}, ..., \vec{0}) \tag{5.6.2}$$

$$I_j(\vec{x}_j) := \vec{a} + i_j(\vec{x}_j - \vec{a}_j). \tag{5.6.3}$$

Se deduce entonces fácilmente que

$$D_j f(\vec{a}) = D[f \circ I_j](p_j(\vec{a})). \tag{5.6.4}$$

Teorema 5.6.1. *Dados $n+1$ e.v.n. $\vec{E}_1, ..., \vec{E}_n, \vec{F}$, un abierto A en el e.v.n. $\vec{E}_1 \times ... \times \vec{E}_n$, un elemento $\vec{a} \in A$ y una función $f : A \to \vec{F}$ diferenciable en $\vec{a} \in A$, entonces para todo $\vec{v} \in \vec{E}_1 \times ... \times \vec{E}_n$ se tiene la fórmula*

$$Df(\vec{a})(\vec{v}) = \sum_{j=1}^{n} D_j f(\vec{a})(\vec{v}_j) \tag{5.6.5}$$

donde $\vec{v}_j := p_j(\vec{v})$ (ver (5.6.1)*).*

Demostración. De la igualdad (5.6.4), usando el Teorema 5.5.1 y el Ejemplo 5.3.1 que nos muestra que $DI_j(p_j(\vec{a})) = i_j$, vemos que

$$\begin{aligned} D_j f(\vec{a}) &= D[f \circ I_j](p_j(\vec{a})) \\ &= Df(I_j(p_j(\vec{a}))) \circ DI_j(p_j(\vec{a})) \qquad (*) \\ &= Df(\vec{a}) \circ i_j. \end{aligned}$$

Si componemos la igualdad (*) con p_j y sumamos sobre j, obtenemos

$$\sum_{j=1}^{n} D_j f(\vec{a}) \circ p_j = \sum_{j=1}^{n} Df(\vec{a}) \circ i_j \circ p_j$$

y como $Df(\vec{a})$ es lineal y $\sum_{j=1}^{n} i_j \circ p_j$ es la identidad en $\vec{E}_1 \times ... \times \vec{E}_n$, concluimos que

$$\sum_{j=1}^{n} D_j f(\vec{a}) \circ p_j = Df(\vec{a})$$

que corresponde exactamente a la igualdad (5.6.5). $\square$

Nota 5.6.2. Si en la Definición 5.6.1 se tiene para todo $j \in \{1, ..., n\}$ que $\vec{E}_j := \mathbb{R}^{k_j}$, $\vec{F} = \mathbb{R}^m$ y si $f_1, ..., f_m$ son las funciones componentes de f, entonces la fórmula (5.6.5) se escribe, usando notación matricial

$$Df(\vec{a})(\vec{v}) = \sum_{j=1}^{n} J_j(\vec{a}) \vec{v}_j^{\,t} \tag{5.6.6}$$

donde $J_j(\vec{a})$ es el jacobiano asociado al diferencial $D_j f(\vec{a})$, esto es la matriz de coeficientes $\partial_\ell f_i(\vec{a})$ donde el indice $i = 1, ..., m$ indica la fila y $\ell : n_j, ..., n_{j+1} - 1$ indica la columna $(n_j = \sum_{p=1}^{j} k_p)$.

5.7 Teoremas de la función inversa y de la función implícita

Nota 5.7.1. Sea $f : \mathbb{R} \to \mathbb{R}$ una función continuamente derivable y tal que $f'(a) \neq 0$. Del curso de cálculo de funciones de una variable sabemos que existe un intervalo $I :=]a - \varepsilon, a + \varepsilon[$, tal que la restricción $\tilde{f} : I \to f(I)$ de la función f es biyectiva y su función inversa $\tilde{f}^{-1}$ es continuamente derivable. Se tiene además que $(\tilde{f}^{-1})'(f(\vec{a})) = 1/f'(\vec{a})$. En el teorema que sigue vamos a generalizar este resultado al caso en que f es una función continuamente diferenciable, definida en un abierto A de un espacio de Banach $\vec{E}$ (ver Definición 1.4.4) con valores en $\vec{E}$.

Antes de enunciar el teorema vamos a demostrar en el Lema 5.7.1 una propiedad importante de las funciones continuas.

Ejemplo 5.7.1. En este ejemplo mostraremos una función $f : \mathbb{R} \to \mathbb{R}$ que es derivable, con derivada no nula en 0 y, que sin embargo no existe un intervalo $I =]-\varepsilon, \varepsilon[$ tal que la restricción de f a I sea inyectiva. Esta función está definida por la fórmula

$$f(x) = \begin{cases} \frac{x}{2} + x^2 sen(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0. \end{cases}$$

Que no exista $\varepsilon > 0$ tal que f sea inyectiva en $]-\varepsilon, \varepsilon[$ es equivalente a decir que en todo intervalo de la forma $]-\varepsilon, \varepsilon[$ la función f no es estrictamente creciente (o decreciente) lo que equivale a decir que en todo intervalo $]-\varepsilon, \varepsilon[$ la derivada f' cambia de signo. Este hecho es evidente puesto que

$$f'(x) = \begin{cases} \frac{1}{2} + 2xsen(\frac{1}{x}) - cos(\frac{1}{x}) & \text{si } x \neq 0 \\ \frac{1}{2} & \text{si } x = 0. \end{cases}$$

Lema 5.7.1. *Sea f una función continua definida en un abierto A de un e.v.n. $\vec{E}$ con valores en un e.v.n. $\vec{F}$. Entonces, la imagen recíproca $f^{-1}[\mathcal{O}]$ de un conjunto abierto $\mathcal{O} \subset \vec{F}$ es un conjunto abierto de $\vec{E}$.*

Demostración. Debemos probar que para todo $\vec{x} \in f^{-1}[\mathcal{O}]$ existe $\delta > 0$ tal que $B(\vec{x}, \delta) \subset f^{-1}[\mathcal{O}]$. Dado $\vec{x} \in f^{-1}[\mathcal{O}]$, se tiene que $f(\vec{x}) \in \mathcal{O}$ y como $\mathcal{O}$ es abierto, debe existir $\varepsilon > 0$ tal que $B(f(\vec{x}), \varepsilon) \subset \mathcal{O}$. Por otra parte, como f es continua en $\vec{x}$, debe existir $B(\vec{x}, \delta) \subset A$ tal que $f[B(\vec{x}, \delta)] \subset B(f(\vec{x}), \varepsilon)$. Lo anterior nos muestra que $f[B(\vec{x}, \delta)] \subset \mathcal{O}$ y por lo tanto $B(\vec{x}, \delta) \subset f^{-1}[\mathcal{O}]$. $\square$

Teorema 5.7.1. *Sea f una función continuamente diferenciable, definida en un abierto A de un espacio de Banach $\vec{E}$ con valores en $\vec{E}$. Si $Df(\vec{a}) \in \mathcal{L}(\vec{E}, \vec{E})$ es biyectiva, entonces existe un abierto V en $\vec{E}$ que contiene a $\vec{a}$ y, tal que la restricción $\tilde{f} : V \to f(V)$ de la función f es biyectiva y su función inversa $\tilde{f}^{-1}$ es continuamente diferenciable. Se tendrá además para $\vec{b} := f(\vec{a})$ que*

$$D\tilde{f}^{-1}(\vec{b}) = [Df(\vec{a})]^{-1}. \tag{5.7.1}$$

Demostración. Haremos la demostración en cuatro etapas. En la primera veremos que sin perder generalidad podemos suponer que $\vec{a} = \vec{0}, f(\vec{a}) = \vec{0}$ y $Df(\vec{a}) = I$ (función identidad en $\vec{E}$), con lo cual la demostración se simplifica notablemente. Luego, en la segunda parte, encontraremos $\delta > 0$ tal que la restricción $\tilde{f}$, de la función f al conjunto abierto $V := f^{-1}[B'(\vec{0}, \frac{\delta}{2})] \cap B'(0, \delta)$, (ver lema anterior), es una biyección de ese conjunto en $f(V)$. En la tercera etapa mostraremos que $\tilde{f}^{-1}$ es Lipschitziana con constante de Lipschitz 2. Finalmente probaremos que $\tilde{f}^{-1}$ es continuamente diferenciable.

1a. Parte. Por razones pedagógicas la desarrollaremos al final.

2a. Parte. De acuerdo a la primera parte, podemos suponer $\vec{a} = \vec{0}, f(\vec{0}) = \vec{0}$ y $Df(\vec{0}) = I$ (función identidad en $\vec{E}$). Definamos entonces la función g de A en $\vec{E}$ por $g(\vec{x}) = \vec{x} - f(\vec{x})$. Como g es continuamente diferenciable y $Dg(\vec{0}) = 0$ (función constante en $\vec{E}$ igual a $\vec{0}$), del Teorema 5.4.7 vemos que existe $\delta > 0$ tal que

$$\vec{x}, \vec{y} \in B(\vec{0}, \delta) \Rightarrow \|g(\vec{x}) - g(\vec{y})\| \leq \frac{1}{3}\|\vec{x} - \vec{y}\| \tag{*}$$

y como $g(\vec{0}) = \vec{0}$, si hacemos $\vec{y} = \vec{0}$ en (*) vemos que $g(B(\vec{0}, \delta)) \subset B(\vec{0}, \frac{\delta}{3}) \subset B'(\vec{0}, \frac{\delta}{2})$. Vamos ahora a demostrar que para todo $\vec{u} \in B'(\vec{0}, \frac{\delta}{2})$ existe un único $\vec{x} \in B(0, \delta)$ tal que $\vec{u} = f(\vec{x})$, lo que implica que la restricción $\tilde{f}$ de la función f al conjunto $V := f^{-1}(B'(\vec{0}, \frac{\delta}{2})) \cap B'(\vec{0}, \delta)$ es biyectiva de V en $f(V)$. Para esto definamos la función ℓ de $B(\vec{0}, \delta)$ en $\vec{E}$ por $\ell(\vec{y}) := g(\vec{y}) + \vec{u}$. Como ℓ también verifica la desigualdad (*), vemos por una parte que $\|\vec{y}\| \leq \delta \Rightarrow \|\ell(\vec{y})\| \leq \frac{1}{3}\|\vec{y}\| + \|\vec{u}\| \leq \delta$ lo que significa que ℓ es una función de $B(0, \delta)$ en $B(0, \delta)$ y por otra parte (*) nos muestra que ℓ es contractante. Estos dos hechos nos permiten concluir de acuerdo al teorema del punto fijo (Teorema 2.8.1), que existe un único $\vec{x} \in B(0, \delta)$ tal que $\ell(\vec{x}) = \vec{x}$, lo que equivale a decir que $\vec{u} = f(\vec{x})$. El Lema 5.7.1 garantiza que el conjunto V es abierto.

3a. Parte. Mostremos que

$$\vec{u}, \vec{v} \in f(V) \Rightarrow \|\tilde{f}^{-1}(\vec{u}) - \tilde{f}^{-1}(\vec{v})\| \leq 2\|\vec{u} - \vec{v}\|.$$

Dados $\vec{u}, \vec{v} \in f(V)$, sean $\vec{x}, \vec{y} \in f(V)$ tales que $\vec{x} = \tilde{f}^{-1}(\vec{u})$ y $\vec{y} = \tilde{f}^{-1}(\vec{v})$, puesto que de la segunda parte sabemos que $f + g = i$ y que g verifica (*), concluimos que

$$\begin{aligned} \|\vec{x} - \vec{y}\| &= \|(f+g)(\vec{x}) - (f+g)(\vec{y})\| \\ &\leq \|g(\vec{x}) - g(\vec{y})\| + \|f(\vec{x}) - f(\vec{y})\| \\ &\leq \frac{1}{2}\|\vec{x} - \vec{y}\| + \|f(\vec{x}) - f(\vec{y})\| \end{aligned}$$

lo que muestra $\|\vec{x} - \vec{y}\| \leq 2\|f(\vec{x}) - f(\vec{y})\|$ que equivale a escribir

$$\|\tilde{f}^{-1}(\vec{u}) - \tilde{f}^{-1}(\vec{v})\| \leq 2\|\vec{u} - \vec{v}\|.$$

4a. Parte. Demostremos ahora que $\tilde{f}^{-1}$ es diferenciable. Dado $\vec{b} \in B'(\vec{0}, \frac{\delta}{2})$ probemos que $D\tilde{f}^{-1}(\vec{b}) = [Df(\vec{a})]^{-1}$, donde $\vec{a} = \tilde{f}^{-1}(\vec{b})$. Dados $\vec{b}$ y $\vec{b} + \vec{k} \in B'(\vec{0}, \frac{\delta}{2})$, definiendo $\vec{h} := \tilde{f}^{-1}(\vec{b} + \vec{k}) - \tilde{f}^{-1}(\vec{b})$ vemos que, de acuerdo a la tercera parte de la demostración

$$\|\vec{h}\| = \|\tilde{f}^{-1}(\vec{b} + \vec{k}) - \tilde{f}^{-1}(\vec{b})\| \leq 2\|\vec{k}\|$$

por lo tanto

$$\begin{aligned} \frac{\|\tilde{f}^{-1}(\vec{b} + \vec{k}) - \tilde{f}^{-1}(\vec{b}) - Df(\vec{a})^{-1}(\vec{k})\|}{\|\vec{k}\|} &\leq \frac{2\|\vec{h} - Df(\vec{a})^{-1}(f(\vec{a} + \vec{h}) - f(\vec{a}))\|}{\|\vec{h}\|} \\ &\leq 2\frac{\|Df(\vec{a})^{-1}[Df(\vec{a})(\vec{h}) - (f(\vec{a} + \vec{h}) - f(\vec{a}))]\|}{\|\vec{h}\|} \\ &\leq 2\|Df(\vec{a})^{-1}\|\frac{\|f(\vec{a} + \vec{h}) - f(\vec{a}) - Df(\vec{a})(\vec{h})\|}{\|\vec{h}\|}. \end{aligned}$$

Como f es diferenciable y como cuando $\vec{k} \to \vec{0}$ se tiene que $\vec{h} \to \vec{0}$, obtenemos

$$\lim_{\substack{\vec{k}\to\vec{0} \\ \vec{k}\neq\vec{0}}} \frac{\tilde{f}^{-1}(\vec{b}+\vec{k}) - \tilde{f}^{-1}(\vec{b}) - Df(\vec{a})^{-1}(\vec{k})}{\|\vec{k}\|} = \vec{0}$$

que muestra que $\tilde{f}^{-1}$ es diferenciable y que $D\tilde{f}^{-1}(\vec{b}) = Df(\vec{a})^{-1}$.

1a. Parte. Si definimos la función $h : A - \vec{a} \to \vec{E}$ por $h(\vec{x}) := Df(\vec{a})^{-1}[f(\vec{a}+\vec{x}) - f(\vec{a})]$ vemos que $h(\vec{0}) = \vec{0}$ y que h es continuamente diferenciable si y solo si f es continuamente diferenciable, en efecto $Dh(\vec{x}) = Df(\vec{a})^{-1} \circ Df(\vec{a}+\vec{x})$ y en particular $Dh(\vec{0}) = I$. Vemos también que si existen abiertos $B_1, B_2 \subset \vec{E}$ que contienen a $\vec{0}$ tales que la restricción $\tilde{h}$, de h a B_1 sea una biyección de B_1 en B_2, entonces la restricción $\tilde{f}$, de f a $A_1 := B_1+\vec{a}$, será una biyección de A_1 en $A_2 := Df(\vec{a})(B_2)+f(\vec{a})$, en efecto $\tilde{f}^{-1}(\vec{u}) = \tilde{h}^{-1}(Df(\vec{a})^{-1}(\vec{u}-f(\vec{a})))+\vec{a}$. Finalmente, de la última igualdad queda claro que si $\tilde{h}^{-1}$ es continuamente diferenciable, entonces $\tilde{f}^{-1}$ también lo será y se tiene

$$D\tilde{f}(\vec{u}) = D\tilde{h}^{-1}(Df(\vec{a})^{-1}(\vec{u} - f(\vec{a}))) \circ Df(\vec{a})^{-1}.$$

$\square$

Nota 5.7.2. Dada una función $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ nos preguntamos si es posible despejar la variable y (en función de la variable x) en la ecuación

$$f(x, y) = 0 \tag{5.7.2}$$

es decir, ¿existe una función $\phi : \mathbb{R} \to \mathbb{R}$ tal que "$f(x, \phi(x)) = 0$ para todo $x \in \mathbb{R}$", o lo que es equivalente, tal que "$f(x, y) = 0 \Leftrightarrow y = \phi(x)$"?

Es fácil ver que "globalmente" esto no es posible, en general no existe ϕ que verifique $f(x, \phi(x)) = 0$ para todo $x \in \mathbb{R}$.

Sin embargo, cuando f es continuamente diferenciable, dado $(a, b) \in \mathbb{R}\times\mathbb{R}$ que verifica la ecuación (5.7.2), es decir $f(a, b) = 0$, si $\partial_2 f(a, b) \neq 0$, entonces existe $\varepsilon > 0$ y una función continuamente derivable $\phi : [a-\varepsilon, a+\varepsilon] \to \mathbb{R}$ tal que

$$b = \phi(a) \;\text{ y }\; f(x, \phi(x)) = 0 \;\text{ para todo }\; x \in [a-\varepsilon, a+\varepsilon]. \tag{5.7.3}$$

Aplicando la regla de la cadena para derivar la función nula $x \in [a-\varepsilon, a+\varepsilon] \to f(x, \phi(x))$, se tendrá

$$\phi'(a) = -\frac{\partial_1 f(a, b)}{\partial_2 f(a, b)}. \tag{5.7.4}$$

Desarrollemos esta idea para la ecuación $f(x, y) = 0$ definida por la función de clase $\mathcal{C}^1$ $f(x, y) := (x-1)^2 + (y-1)^2 - 1$. Consideremos entonces $(a, b) := (\frac{1}{2}, 1 - \frac{\sqrt{3}}{2})$, que evidentemente verifica $f(a,b)=0$ y además $\partial_2 f(a,b)=-\sqrt{3}\neq 0$. De acuerdo a lo que decíamos, debe entonces existir una función $\phi:[a-\varepsilon,a+\varepsilon]\to\mathbb{R}$ continuamente derivable tal que $\phi(a)=b$ y $f(x,\phi(x))=0$ para todo $x\in[a-\varepsilon,a+\varepsilon]$. No es difícil deducir que la función $\phi:[\frac{1}{4},\frac{3}{4}]\to\mathbb{R}$ definida por $\phi(x)=1-\sqrt{1-(x-1)^2}$ cumple con este requerimiento. Se tiene además $\phi'(\frac{1}{2})=\frac{1}{\sqrt{3}}=-\frac{\partial_1 f(\frac{1}{2},1-\sqrt{\frac{3}{2}})}{\partial_2 f(\frac{1}{2},1-\frac{\sqrt{3}}{2})}$.

Este mismo razonamiento lo podemos hacer para cualquier punto $(a,b)\in\mathbb{R}\times\mathbb{R}$ de la forma $(a,1\pm\sqrt{1-(a-1)^2})$ con $a\in]0,2[$, pues todos esos puntos verifican la ecuación $f(x,y)=0$ y además la derivada parcial de f con respecto a la segunda variable, en esos puntos, no se anula. Esto no sería posible sin embargo, en $(0,1)$ y $(2,1)$ donde $\partial_2 f$ se anula, hecho que en este ejemplo es fácil de evaluar directamente.

El teorema de la función implícita que daremos a continuación generaliza este problema a la ecuación $f(\vec{x},\vec{y})=\vec{0}$, definida por una función continuamente diferenciable $f:A\to\vec{E}_2$, donde A es un abierto en el e.v.n. $\vec{E}_1\times\vec{E}_2$, siendo $\vec{E}_1$ y $\vec{E}_2$ dos espacios de Banach.

Teorema 5.7.2. *Dados dos espacios de Banach $\vec{E}_1$ y $\vec{E}_2$, un abierto $A\subset\vec{E}_1\times\vec{E}_2$ y una función continuamente diferenciable $f:A\to\vec{E}_2$, definimos la ecuación*

$$f(\vec{x}_1,\vec{x}_2)=\vec{0}. \tag{5.7.5}$$

Entonces, si $\vec{a}=(\vec{a}_1,\vec{a}_2)\in A$ verifica la ecuación (5.7.5) *y si $D_2 f(\vec{a})$ (ver Definición 5.6.1) es biyectiva en $\vec{E}_2$, se tiene que existe una bola $B(\vec{a}_1,\varepsilon)\subset\vec{E}_1$ y una función continuamente diferenciable $\phi:B(\vec{a}_1,\varepsilon)\to\vec{E}_2$ tales que*

$$\vec{a}_2=\phi(\vec{a}_1) \quad y \quad f(\vec{x},\phi(\vec{x}))=\vec{0} \quad \textit{para todo} \quad \vec{x}\in B(\vec{a}_1,\varepsilon). \tag{5.7.6}$$

Se tiene además la fórmula

$$D\phi(\vec{a}_1)=D_2 f(\vec{a})^{-1}\circ D_1 f(\vec{a}). \tag{5.7.7}$$

Demostración. Definamos la función $\varphi:A\subset\vec{E}_1\times\vec{E}_2\to\vec{E}_1\times\vec{E}_2$ por $\varphi(\vec{x}):=(\vec{x}_1,f(\vec{x}))$, es decir, $\varphi:=(p_1,f)$ (ver (5.6.1) para la definición de p_1). Entonces, de la Nota 5.4.2, concluimos que φ es continuamente diferenciable y $D\varphi(\vec{a})=(p_1,Df(\vec{a}))$. Por otra parte, del Teorema 5.6.1 sabemos que $Df(\vec{a})(\vec{v})=D_1 f(\vec{a})(\vec{v}_1)+D_2 f(\vec{a})(\vec{v}_2)$ para todo $\vec{v}\in\vec{E}_1\times\vec{E}_2$, lo que muestra que $D\varphi(\vec{a})$ es biyectiva y

$$D\varphi(\vec{a})^{-1}(\vec{u}_1,\vec{u}_2)=(\vec{u}_1,[-D_2 f(\vec{a})^{-1}\circ D_1 f(\vec{a})](\vec{u}_1)+D_2 f(\vec{a})^{-1}(\vec{u}_2)).$$

Concluimos entonces, del Teorema 5.7.1 que existe un abierto $V\subset\vec{E}_1\times\vec{E}_2$ que contiene a $\vec{a}$ tal que la restricción $\tilde{\varphi}:V\to\varphi(V)$ de la función φ, es biyectiva y $\tilde{\varphi}^{-1}:\varphi(V)\to V$ es continuamente diferenciable. Además se tiene que

$$D\tilde{\varphi}^{-1}(\vec{b})=D\varphi(\vec{a})^{-1}$$

donde $\vec{b} := \varphi(\vec{a})$. Es fácil entonces verificar que la función $\phi : B(\vec{a}_1, \varepsilon) \to \vec{E}_2$ definida por $\phi(\vec{u}_1) := \tilde{\varphi}_2^{-1}(\vec{u}_1, \vec{0})$, donde $\tilde{\varphi}_2^{-1}$ es la función componente de $\tilde{\varphi}^{-1}$ en $\vec{E}_2$ (es decir $\tilde{\varphi}^{-1} = (\tilde{\varphi}_1^{-1}, \tilde{\varphi}_2^{-1})$) y donde $\varepsilon > 0$ es tal que $B(\vec{a}_1, \varepsilon) \times \{\vec{0}\} \subset \varphi(V)$, verifica (5.7.6) y (5.7.7). $\square$

5.8 Derivadas parciales de orden superior

Definición 5.8.1. Una función f definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$, se dirá de clase $\mathcal{C}^2$ si ella y sus derivadas parciales $\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}$ son de clase $\mathcal{C}^1$. La derivada parcial de la función $\frac{\partial f}{\partial x_i}$ con respecto a la variable x_j la denotaremos $\frac{\partial^2 f}{\partial x_j \partial x_i}$ o bien $\partial^2_{j,i} f$. Estas funciones se llaman derivadas parciales de segundo orden de la función f. Por recurrencia sobre k, diremos que f es de clase $\mathcal{C}^k$ si ella es de clase $\mathcal{C}^{k-1}$ y sus derivadas parciales de orden $k-1$, que denotaremos por

$$\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} ... \partial x_{i_1}} \quad \text{para todo} \;\; i_1, ..., i_{k-1} \in \{1, ..., n\}, \tag{5.8.1}$$

son de clase $\mathcal{C}^1$. Las derivadas parciales de orden k de f las denotaremos $\frac{\partial^k f}{\partial x_{i_k} \partial x_{i_{k-1}} ... \partial x_{i_1}}$ o bien $\partial^k_{i_k, i_{k-1}, ..., i_1} f$.

Nota 5.8.1. Es una consecuencia casi inmediata de la definición anterior que si f y g son dos funciones de clase $\mathcal{C}^k$, definidas en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$, entonces :

(i) las funciones $f + g$ y λf también son de clase $\mathcal{C}^k$ y se tiene

$$\partial^k_{i_k, ..., i_1}(f + g)(\vec{a}) = \partial^k_{i_k, ..., i_1} f(\vec{a}) + \partial^k_{i_k, ..., i_1} g(\vec{a}) \quad \text{para todo } \vec{a} \in A \tag{5.8.2}$$

$$\partial^k_{i_k, ..., i_1}(\lambda f)(\vec{a}) = \lambda \partial^k_{i_k, ..., i_1} f(\vec{a}) \quad \text{para todo } \vec{a} \in A. \tag{5.8.3}$$

(ii) Cuando $\vec{F} = \mathbb{R}$ la función $f \cdot g$ también es de clase $\mathcal{C}^k$.

Cuando $\vec{F} = \mathbb{R}^m$ se tendrá que f es de clase $\mathcal{C}^k$ si y solo si cada una de sus m funciones componentes es de clase $\mathcal{C}^k$, y en ese caso

$$\partial^k_{i_k, ..., i_1} f(\vec{a}) = (\partial^k_{i_k, ..., i_1} f_1(\vec{a}), ..., \partial^k_{i_k, ..., i_1} f_m(\vec{a})) \tag{5.8.4}$$

Teorema 5.8.1. *Sea f una función de clase $\mathcal{C}^2$, definida en un abierto A de $\mathbb{R}^2$ con valores en un e.v.n. $\vec{F}$. Entonces se tendrá para todo $\vec{a} \in A$, la fórmula*

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a}) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}). \tag{5.8.5}$$

Demostración. Sean $\vec{a} \in A$ y $\varepsilon > 0$. Por continuidad de las derivadas parciales de orden dos, existirá $\eta > 0$ tal que

$$\|\vec{x}\|_\infty \leq \eta \Rightarrow \|\partial^2_{2,1} f(\vec{a}+\vec{x}) - \partial^2_{2,1} f(\vec{a})\| \leq \varepsilon \ \text{ y } \ \|\partial^2_{1,2} f(\vec{a}+\vec{x}) - \partial^2_{1,2} f(\vec{a})\| \leq \varepsilon. \tag{5.8.6}$$

Si $\vec{e}_1$ y $\vec{e}_2$ son los vectores de la base canónica en $\mathbb{R}^2$ y $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$, denotemos

$$\Delta^2 f(\vec{x}) := f(\vec{a} + x_1\vec{e}_1 + x_2\vec{e}_2) - f(\vec{a} + x_1\vec{e}_1) - f(\vec{a} + x_2\vec{e}_2) + f(\vec{a})$$

y demostremos que

$$\|\vec{x}\|_\infty \leq \eta \Rightarrow \|\Delta^2 f(\vec{x}) - x_1 x_2 \partial^2_{2,1} f(\vec{a})\| \leq \varepsilon |x_1||x_2|. \tag{5.8.7}$$

Si definimos $\varphi(t) := f(a + t\vec{e}_1 + x_2\vec{e}_2) - f(\vec{a} + t\vec{e}_1) - t x_2 \partial^2_{2,1} f(\vec{a})$, aplicando el teorema del valor medio (t.v.m.) a esta función, vemos que existe $\bar{t} \in [0, x_1]$ tal que

$$\begin{aligned} \Delta^2 f(\vec{x}) - x_1 x_2 \partial^2_{2,1} f(\vec{a}) &= \varphi(x_1) - \varphi(0) = \varphi'(\bar{t}) x_1 \\ &= [\partial_1 f(\vec{a} + \bar{t} e_1 + x_2\vec{e}_2) - \partial_1 f(\vec{a} + \bar{t}\vec{e}_1) - \partial^2_{2,1} f(\vec{a}) x_2] x_1 \end{aligned}$$

y si definimos $\theta(\tau) := \partial_1 f(\vec{a} + \bar{t}\vec{e}_1 + \tau\vec{e}_2) - \tau \partial^2_{2,1} f(\vec{a})$ aplicando el t.v.m. a esta función, vemos que existe $\bar{\tau} \in [0, x_2]$ tal que

$$\begin{aligned} \Delta^2 f(\vec{x}) - x_1 x_2 \partial^2_{2,1} f(\vec{a}) &= [\theta(x_2) - \theta(0)] x_1 = \theta'(\bar{\tau}) x_2 x_1 \\ &= [\partial^2_{2,1} f(\vec{a} + \bar{t}\vec{e}_1 + \bar{\tau}\vec{e}_2) - \partial^2_{2,1} f(\vec{a})] x_2 x_1. \end{aligned}$$

De esta última igualdad, usando la primera desigualdad en (5.8.6), obtenemos fácilmente la implicación (5.8.7).

Del mismo modo se puede obtener la implicación

$$\|\vec{x}\|_\infty \leq \eta \Rightarrow \|\Delta^2 f(\vec{x}) - x_1 x_2 \partial^2_{1,2} f(\vec{a})\| \leq \varepsilon |x_1||x_2|. \tag{5.8.8}$$

De (5.8.7) y (5.8.8), sumando se obtiene que

$$\|\partial^2_{2,1} f(\vec{a}) - \partial^2_{1,2} f(\vec{a})\| \leq 2\varepsilon$$

y como ε es cualquiera, concluimos que $\partial^2_{2,1} f(\vec{a}) = \partial^2_{1,2} f(\vec{a})$. □

Teorema 5.8.2. *Sea f una función de clase $\mathcal{C}^k$, definida en un abierto A de $\mathbb{R}^n$ con valores en un e.v.n. $\vec{F}$. Entonces para toda k-tupla $(i_1, ..., i_k) \in \{1, ..., n\}^k$ y toda biyección σ en $\{1, ..., k\}$ se tendrá para todo $\vec{a} \in A$ la fórmula*

$$\frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(\vec{a}) = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} ... \partial x_{i_{\sigma(k)}}}(\vec{a}).$$

Demostración. Si $k = 2$ es una consecuencia inmediata del teorema anterior. Si $k \geq 3$, se deduce del teorema anterior por inducción sobre k. □

5.9 Desarrollos limitados

Definición 5.9.1. Se llama polinomio de grado $N \in \mathbb{N}_*$ (conjunto de los naturales incluyendo el 0) en $\mathbb{R}^n$ con coeficientes en $\mathbb{R}$ a toda función $P : \mathbb{R}^n \to \mathbb{R}$ de la forma

$$P(\delta_1, ..., \delta_n) = \sum_{|\vec{i}| \leq N} \alpha_{\vec{i}}\ \delta_1^{i_1}\ \delta_2^{i_2} ...\ \delta_n^{i_n} \tag{5.9.1}$$

donde $\vec{i} := (i_1, ..., i_n) \in \mathbb{N}^n_*$ es una n-tupla de componentes naturales, $|\vec{i}| = \sum_{j=1}^{n} i_j$ y $\alpha_{\vec{i}} \in \mathbb{R}$.

Definición 5.9.2. Sea A un conjunto abierto en $\mathbb{R}^n$ y f una función de A en $\mathbb{R}$. Se llama desarrollo limitado de orden N de la función f en un punto $\vec{a} \in A$ a todo polinomio $P : \mathbb{R}^n \to \mathbb{R}$ de grado N tal que para todo $\vec{\delta} \in \mathbb{R}^n$ con $\vec{a} + \vec{\delta} \in A$ se tiene

$$f(\vec{a} + \vec{\delta}) - P(\vec{\delta}) = o^N(\vec{\delta}) \tag{5.9.2}$$

donde o^N es una función de $\mathbb{R}^n$ en $\mathbb{R}$ que verifica $o^N(\vec{0}) = 0$ y

$$\lim_{\substack{\vec{\delta} \to \vec{0} \\ \vec{\delta} \neq 0}} \frac{o^N(\vec{\delta})}{\|\vec{\delta}\|^N} = 0. \tag{5.9.3}$$

Nota 5.9.1. De la definición anterior es evidente que en (5.9.3) se puede usar cualquier norma en $\mathbb{R}^n$ y por lo tanto el desarrollo limitado de una función f no depende de la norma. Además es fácil ver que P es un desarrollo limitado de orden N de f en $\vec{a}$, entonces $P(\vec{0}) = f(\vec{a})$ y, se deduce fácilmente del Teorema 2.3.1, que f es continua en $\vec{a}$.

Nota 5.9.2. De la Definición 5.3.1 se desprende de inmediato que f tiene un desarrollo limitado de orden 1 en un punto $\vec{a} \in A$ si y solo si es diferenciable en $\vec{a}$. En este caso se tendrá

$$P(\vec{\delta}) = f(\vec{a}) + Df(\vec{a})(\vec{\delta}) \tag{5.9.4}$$

que de acuerdo a la fórmula (5.3.13), podemos escribir

$$P(\vec{\delta}) = f(\vec{a}) + \langle \nabla f(\vec{a}), \vec{\delta} \rangle \tag{5.9.5}$$

que efectivamente es un polinomio de grado 1 en $\mathbb{R}^n$.

Nota 5.9.3. De la Nota 5.3.2 vemos que el desarrollo limitado de orden 1 en $\vec{a}$ de una función f, dado por (5.9.5), define la aproximación de primer orden de f en $\vec{a}$, mediante la fórmula $h(\vec{x}) = P(\vec{x} - \vec{a})$. Del Teorema 5.3.1 concluimos que el desarrollo limitado de orden 1 de una función en un punto (que está dado por (5.9.4)) es siempre único.

Nota 5.9.4. Dada una función f definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, su desarrollo limitado de orden N en $\vec{a} \in A$ define la llamada aproximación de orden N de f en $\vec{a}$ que está dada por la función $h(\vec{x}) := P_N(\vec{x} - \vec{a})$.

Dos funciones f y h definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, se dicen tangentes de orden N en un punto $\vec{a} \in A$, si $f(\vec{a} + \vec{\delta}) - h(\vec{a} + \vec{\delta}) = o^N(\vec{\delta})$ donde o^N es una función de $\mathbb{R}^n$ en $\mathbb{R}$ que verifica $o^N(\vec{0}) = 0$ y la relación (5.9.3).

De lo anterior vemos que la aproximación de orden N de la función f en $\vec{a} \in A$, corresponde al polinomio de grado N que es tangente de orden N a la función f en $\vec{a}$.

Teorema 5.9.1. *Sea f una función definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Si el polinomio*

$$P_N(\vec{\delta}) := \sum_{|\vec{i}| \leq N} \alpha_{\vec{i}} \; \delta_1^{i_1} \; \delta_2^{i_2} \ldots \; \delta_n^{i_n}$$

es un desarrollo limitado de orden N de la función f en $\vec{a} \in A$, entonces para todo $K < N$, el polinomio

$$P_K(\vec{\delta}) = \sum_{|\vec{i}| \leq K} \alpha_{\vec{i}} \; \delta_1^{i_1} \; \delta_2^{i_2} \ldots \; \delta_n^{i_n}$$

es un desarrollo limitado de orden K de la función f en $\vec{a}$.

Demostración. Puesto que

$$\lim_{\vec{\delta}\to\vec{0}} \frac{\frac{f(\vec{a}+\vec{\delta})-\sum\limits_{|\vec{i}|\le K} \alpha_{\vec{i}}\ \delta_1^{i_1}\dots\ \delta_n^{i_n}}{\|\vec{\delta}\|^K} - \frac{\sum\limits_{K+1\le|\vec{i}|\le N} \alpha_{\vec{i}}\ \delta_1^{i_1}\dots\ \delta_n^{i_n}}{\|\vec{\delta}\|^K}}{\|\vec{\delta}\|^{N-K}} = 0$$

y también (usando la norma infinito)

$$\lim_{\vec{\delta}\to\vec{0}} \frac{\sum\limits_{K+1\le|\vec{i}|\le N} \alpha_{\vec{i}}\ \delta_1^{i_1}\dots\ \delta_n^{i_n}}{\|\vec{\delta}\|^K} = 0$$

deducimos que

$$\lim_{\vec{\delta}\to\vec{0}} \frac{f(\vec{a}+\vec{\delta}) - \sum\limits_{|\vec{i}|\le K} \alpha_{\vec{i}}\ \delta_1^{i_1}\dots\ \delta_n^{i_n}}{\|\vec{\delta}\|^K} = 0$$

lo que equivale a decir que P_K es un desarrollo limitado de orden K de la función f en $\vec{a}$. □

Teorema 5.9.2. *Si una función f definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, admite un desarrollo limitado P_N de orden N en $\vec{a} \in A$, entonces éste es único.*

Demostración. Hagamos la demostración por inducción. En la Nota 5.9.3 vimos que la unicidad se tiene cuando $N = 1$. Supongamos ahora que f admite un desarrollo limitado P_N de orden N y que su desarrollo limitado P_{N-1} de orden $N-1$ es único. Demostremos entonces la unicidad de P_N.

Si denotamos P_N y P'_N dos desarrollos limitados de orden N de f en $\vec{a}$, del Teorema 5.9.1 y del hecho que el desarrollo limitado de orden $N-1$ es único, podemos escribir

$$\begin{aligned} P_N(\vec{\delta}) &= P_{N-1}(\vec{\delta}) + \sum_{|\vec{i}|=N} \alpha_{\vec{i}}\ \vec{\delta}_1^{i_1}\dots\ \delta_n^{i_n} \\ P'_N(\vec{\delta}) &= P_{N-1}(\vec{\delta}) + \sum_{|\vec{i}|=N} \alpha'_{\vec{i}}\ \vec{\delta}_1^{i_1}\dots\ \delta_n^{i_n} \end{aligned}$$

y aplicando a P_N y P'_N la igualdad (5.9.2) y luego restando, se obtiene

$$\phi(\vec{\delta}) := \sum_{|\vec{i}|=N} (\alpha_{\vec{i}} - \alpha'_{\vec{i}})\ \delta_1^{i_1}\dots\ \delta_n^{i_n} = o^N(\vec{\delta}). \qquad (*)$$

De este modo para todo $\vec{\delta} \in \mathbb{R}^n$ y todo $\lambda \in \mathbb{R}$ se tiene $\phi(\lambda\vec{\delta}) = \lambda^N \phi(\vec{\delta})$, es decir $\phi(\vec{\delta}) = \frac{\phi(\lambda\vec{\delta})}{\lambda^N}$. De la igualdad (*) se obtiene entonces

$$\lim_{\lambda\to 0} \frac{\phi(\lambda\vec{\delta})}{\lambda^N} = \|\vec{\delta}\|^N \lim_{\lambda\to 0} \frac{\phi(\lambda\vec{\delta})}{\|\lambda\vec{\delta}\|^N} = 0.$$

Esto implica que $\phi(\vec{\delta}) = 0$ para todo $\vec{\delta} \in \mathbb{R}^n$ y por lo tanto todos los coeficientes del polinomio ϕ deben ser nulos, esto es $\alpha_{\vec{i}} - \alpha'_{\vec{i}} = 0$ para todo $\vec{i} \in \mathbb{N}^n_*$ con $|\vec{i}| = N$. Todo esto nos permite concluir que $P_N = P'_N$. $\square$

Teorema 5.9.3. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Entonces para todo $\vec{a} \in A, f$ tiene un desarrollo limitado P_2 de orden dos en $\vec{a}$. Se tiene además que*

$$P_2(\vec{\delta}) = f(\vec{a}) + \langle \nabla f(\vec{a}), \vec{\delta} \rangle + \frac{1}{2}\vec{\delta}^t H(\vec{a})^t \vec{\delta} \tag{5.9.6}$$

donde $H(\vec{a})$ es la matriz de $n \times n$, llamada matriz Hessiana de f en $\vec{a}$, cuyos coeficientes son las derivadas parciales de orden dos de f en $\vec{a}$, esto es

$$H(\vec{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a}) ... & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}) & \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) \quad ... & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\vec{a}) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\vec{a}) ... & \frac{\partial^2 f}{\partial x_n^2}(\vec{a}) \end{pmatrix} \tag{5.9.7}$$

Demostración. Sea $r > 0$ tal que $B(\vec{a}; r) \subset A$. Sea $\vec{\delta} \in \mathbb{R}^n$ tal que $\|\vec{\delta}\| \leq r$. Definamos entonces la función $\varphi : [0, 1] \to \mathbb{R}$ por $\varphi(t) := f(\vec{a} + t\vec{\delta})$. Esta función es de clase $\mathcal{C}^1$ y de acuerdo al teorema fundamental del calculo se tiene

$$\varphi(1) = \varphi(0) + \int_0^1 \varphi'(t) dt \tag{*}$$

donde $\varphi'(t) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a} + t\vec{\delta})\delta_i$. Como las funciones $\frac{\partial f}{\partial x_i}$ son de clase $\mathcal{C}^1$, ellas son diferenciables y por lo tanto, para cada $i \in \{1, ..., n\}$ se tiene

$$\frac{\partial f}{\partial x_i}(\vec{a} + t\vec{\delta}) = \frac{\partial f}{\partial x_i}(\vec{a}) + \sum_{j=1}^{n} \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(\vec{a}) t\delta_j + o_i(t\vec{\delta})$$

y reemplazando en (*) obtenemos

$$\begin{aligned} f(\vec{a} + \vec{\delta}) = f(\vec{a}) \quad &+ \quad \int_0^1 \sum_{i=1}^{n} \left[\frac{\partial f}{\partial x_i}(\vec{a}) + \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\vec{a})\delta_j t + o_i(t\vec{\delta}) \right] \delta_i dt \\ &= \quad f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})\delta_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\vec{a})\delta_j \delta_i + \int_0^1 \sum_{i=1}^{n} o_i(t\vec{\delta})\delta_i dt \\ &= \quad f(\vec{a}) + \langle \nabla f(\vec{a}), \vec{\delta} \rangle + \frac{1}{2}\vec{\delta}^t H(\vec{a})^t \vec{\delta} + \int_0^1 \sum_{i=1}^{n} o_i(t\vec{\delta})\delta_i dt. \end{aligned}$$

De acuerdo a la Definición 5.9.2. Para terminar la demostración debemos probar que la función

$$\sigma(\vec{\delta}) := \int_0^1 \sum_{i=1}^n o_i(t\vec{\delta})\delta_i dt$$

verifica $\lim\limits_{\substack{\vec{\delta}\to\vec{0}\\ \vec{\delta}\neq\vec{0}}} \frac{\sigma(\vec{\delta})}{\|\vec{\delta}\|^2} = 0$.

Por definición de $o_i(\cdot)$, para todo $\varepsilon > 0$ existe $\eta > 0$ tal que

$$\|\vec{\delta}\| < \eta \Rightarrow |o_i(\vec{\delta})| \leq \varepsilon\|\vec{\delta}\| \quad \text{para todo} \quad i \in \{1, ..., n\}$$

entonces, con $\tilde{\eta} := \min\{\eta, r\}$ obtenemos, usando como ha sido habitual la norma ∞ en $\mathbb{R}^n$,

$$\|\vec{\delta}\| \leq \tilde{\eta} \Rightarrow |\sigma(\vec{\delta})| \leq \int_0^1 \sum_{i=1}^n |\delta_i| \; \varepsilon \; \|\vec{\delta}\| \; t \; dt \leq \varepsilon \; n\|\vec{\delta}\|^2 \frac{1}{2}$$

lo que demuestra que $\lim\limits_{\substack{\vec{\delta}\to\vec{0}\\ \vec{\delta}\neq\vec{0}}} \frac{\sigma(\vec{\delta})}{\|\vec{\delta}\|^2} \leq \varepsilon\frac{n}{2}$ y, como ε es cualquiera, concluimos que el límite es 0. $\square$

Nota 5.9.5. Del teorema anterior vemos que si f es una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, su aproximación de orden 2 o aproximación cuadrática en $\vec{a} \in A$, definida en la Nota 5.9.4, está dada por la fórmula

$$h(\vec{x}) = f(\vec{a}) + \langle \nabla f(\vec{a}), \vec{x} - \vec{a} \rangle + \frac{1}{2}(\vec{x} - \vec{a})^t H(\vec{x})(\vec{x} - \vec{a}). \tag{5.9.8}$$

Nota 5.9.6. Es fácil verificar que la aproximación cuadrática de una función cuadrática definida en $\mathbb{R}^n$ con valores $\mathbb{R}$ es, en todo punto de $\mathbb{R}^n$, ella misma.

Teorema 5.9.4. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Entonces la función ∇f definida en A con valores en $\mathbb{R}^n$ es de clase $\mathcal{C}^1$ y*

$$D[\nabla f](\vec{x})(\vec{v}) = H(\vec{x})\vec{v}^t. \tag{5.9.9}$$

Demostración. Si f es de clase $\mathcal{C}^2$ se tendrá que las funciones $\partial_1 f, ..., \partial_n f$ son de clase $\mathcal{C}^1$, lo que es equivalente de acuerdo a los teoremas 5.4.1 y 5.4.3, a decir que ellas son continuamente diferenciables. Aplicando entonces el Teorema 5.4.6 a la función $\nabla f := (\partial_1 f, ..., \partial_n f)$ concluimos que ella es continuamente diferenciable y de la fórmula (5.3.10) vemos que

$$D[\nabla f](\vec{x}) = (D[\partial_1 f](\vec{x}), ..., D[\partial_n f](\vec{x})).$$

Aplicando ahora el Teorema 5.3.5 a las funciones $\partial_i f$, obtenemos para cada $i = 1, ..., n$

$$\begin{aligned} D[\partial_i f](\vec{x})(\vec{v}) &= \sum_{j=1}^{n} \partial_{j,i} f(\vec{x}) v_j \\ &= (\partial_{1,i} f(\vec{x}), ..., \partial_{n,i} f(\vec{x}))\vec{v}^t \end{aligned}$$

que no es otra que la fórmula (5.9.9) escrita componente a componente. □

Nota 5.9.7. Terminaremos esta sección enunciando el teorema que generaliza la fórmula (5.9.6) al caso en que la función f es de clase $\mathcal{C}^N$.

Teorema 5.9.5. *Sea f una función de clase $\mathcal{C}^N$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Entonces para todo $\vec{a} \in A, f$ tiene un desarrollo limitado de orden N en $\vec{a}$, dado por el polinomio*

$$P_N(\vec{\delta}) = \sum_{|\vec{i}| \leq N} \alpha_{\vec{i}} \;\; \delta_1^{i_1} \delta_2^{i_2} .... \delta_n^{i_n} \tag{5.9.10}$$

donde

$$\alpha_{\vec{i}} = \frac{1}{i_1!...i_n!} \quad \frac{\partial^{|\vec{i}|} f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(\vec{a}), \quad \alpha_{\vec{0}} = f(\vec{a}) \tag{5.9.11}$$

con la convención $0! = 1$.

Nota 5.9.8. Del Teorema anterior vemos que si f es una función de clase $\mathcal{C}^N$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$, su aproximación de orden N en $\vec{a} \in A$, definida en la Nota 5.9.4, está dada por la fórmula

$$h(\vec{x}) = \sum_{|\vec{i}| \leq N} \alpha_{\vec{i}} (x_1 - a_1)^{i_1} (x_2 - a_2)^{i_2} ... (x_n - a_n)^{i_n} \tag{5.9.12}$$

donde los coeficientes $\alpha_{\vec{i}}$ están definidos por (5.9.11).

El polinomio (5.9.12) con los coeficientes definidos por (5.9.11) se llama también desarrollo de Taylor de orden N de la función f en $\vec{a}$.

CAPÍTULO 6

CONVEXIDAD Y EXTREMOS DE FUNCIONES DIFERENCIABLES

6.1 Introducción

En el siguiente capítulo se introduce una clase muy importante de funciones definidas sobre un e.v.n. a valores en $\mathbb{R}$. Aquella de las funciones convexas (y concavas). Esta clase, que es más general que la de las funciones lineales a valores en $\mathbb{R}$, aparece notablemente en numerosos modelos de las matemáticas aplicadas.

En primer lugar, se definen las funciones convexas, y se dan diversas caracterizaciones de aquellas que son diferenciable. También se muestran caracterizaciones de funciones convexas de aquellas funciones que son de clase $\mathcal{C}^2$. Luego, se presentan una serie de condiciones necesarias y suficientes para que un elemento del e.v.n. sea mínimo de una función convexa sobre un conjunto determinado. Estos mismos resultados son extendibles a la clase de funciones concavas. Finalmente, cerramos el capítulo con el resultado mas importante de esta parte. El Teorema de Kuhn-Tucker, el cual nos entrega una condición necesaria y suficiente para los mínimos de una función sobre un conjunto dado por desigualdades de funciones convexas.

6.2 Funciones Convexas

Definición 6.2.1. Una función f definida en una parte convexa (ver Definición 4.3.1) Q de un e.v. $\vec{E}$ con valores en $\mathbb{R}$, se dirá convexa en Q si para todo $\vec{x}, \vec{y} \in Q$ se tiene la desigualdad

$$f(\lambda\vec{x} + (1-\lambda)\vec{y}) \leq \lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) \quad \text{para todo} \quad \lambda \in [0,1]. \tag{6.2.1}$$

Si la desigualdad anterior es estricta cuando $\vec{x} \neq \vec{y}$ y $\lambda \in]0,1[$, diremos que la función f es estrictamente convexa.

Definición 6.2.2. Se llama epigrafo de una función f definida en una parte A de un e.v. $\vec{E}$ con valores en $\mathbb{R}$, al conjunto

$$epi(f) = \{(\vec{x}, z) \in A \times \mathbb{R} : f(\vec{x}) \leq z\}. \tag{6.2.2}$$

Geométricamente, esto se interpreta como el conjunto de los puntos $(\vec{x}, z) \in A \times \mathbb{R}$ que están "sobre" el grafo de f (recordemos que el grafo de una función f es el conjunto $\{(\vec{x}, z) \in A \times \mathbb{R} : f(\vec{x}) = z\}$)

Nota 6.2.1. De acuerdo a la definición anterior, la desigualdad (6.2.1) es equivalente a

$$(\lambda\vec{x} + (1-\lambda)\vec{y}\ ,\ \lambda f(\vec{x}) + (1-\lambda)f(\vec{y})) \in epi(f)$$

que escrito de otro modo toma la forma

$$\lambda(\vec{x}, f(\vec{x})) + (1-\lambda)(\vec{y}, f(\vec{y})) \in epi(f).$$

Vemos entonces que f será convexa si y solo si para todo $\vec{x}, \vec{y} \in Q$ el trazo que une los puntos $(\vec{x}, f(\vec{x}))$ e $(\vec{y}, f(\vec{y}))$ en el grafo de f, pertenece al epigrafo de f. No es difícil entonces ver que la convexidad de una función f es equivalente a la convexidad del conjunto $epi(f)$ en $\vec{E} \times \mathbb{R}$.

Teorema 6.2.1. *Dadas dos funciones convexas (resp. estrictamente convexas) f, g definidas en un conjunto convexo Q de un e.v. $\vec{E}$ con valores en $\mathbb{R}$ y dado $t \in \mathbb{R}_+$, entonces*

(i) *la función $f + g$ es convexa (resp. estrictamente convexa).*

(ii) *la función tf es convexa (resp. estrictamente convexa).*

Demostración. Dados $\vec{x}, \vec{y} \in Q$ y $\lambda \in [0,1]$, se tiene (i)

$$\begin{aligned}
[f+g](\lambda\vec{x} + (1-\lambda)\vec{y}) &= f(\lambda\vec{x} + (1-\lambda)\vec{y}) + g(\lambda\vec{x} + (1-\lambda)\vec{y}) \\
&\leq \lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) + \lambda g(\vec{x}) + (1-\lambda)g(\vec{y}) \\
&= \lambda[f+g](\vec{x}) + (1-\lambda)[f+g](\vec{y})
\end{aligned}$$

(ii)

$$\begin{aligned}[tf](\lambda\vec{x} + (1-\lambda)\vec{y}) &= t(f(\lambda\vec{x} + (1-\lambda)\vec{y})) \\ &\leq t(\lambda f(\vec{x}) + (1-\lambda) f(\vec{y})) \\ &= \lambda[tf](\vec{x}) + (1-\lambda)[tf](\vec{y}).\end{aligned}$$

□

Teorema 6.2.2. *Dada una familia $\{g_t\}_{t\in T}$ de funciones convexas definidas en un conjunto convexo Q de un e.v. $\vec{E}$ con valores en $\mathbb{R}$, si para cada $\vec{x} \in Q$ el conjunto $\{g_t(\vec{x}) : t \in T\}$ es acotado superiormente, entonces la función $f : Q \to \mathbb{R}$ definida por*

$$f(\vec{x}) = \sup\{g_t(\vec{x}) : t \in T\}$$

es convexa.

Demostración. De acuerdo con la Nota 6.2.1 basta con demostrar que $epi(f)$ es un conjunto convexo en $\vec{E} \times \mathbb{R}$. Vamos a mostrar entonces que $epi(f) = \bigcap_{t\in T} epi(g_t)$ y, como la intersección de conjuntos convexos es siempre convexa, habremos concluido que $epi(f)$ es convexo. La igualdad en cuestion es una consecuencia inmediata de la cadena de implicaciones

$$\begin{aligned}(\vec{x}, z) \in epi(f) \Leftrightarrow f(\vec{x}) \leq z &\Leftrightarrow g_t(\vec{x}) \leq z \quad \forall t \in T \Leftrightarrow (\vec{x}, z) \in epi(g_t) \quad \forall t \in T \\ &\Leftrightarrow (\vec{x}, z) \in \bigcap_{t\in T} epi(g_t)\end{aligned}$$

□

6.3 Caracterización de funciones convexas diferenciables

Nota 6.3.1. En los dos teoremas que siguen daremos una caracterización importante de la convexidad (resp. estricta convexidad) para funciones diferenciables.

Teorema 6.3.1. *Sea f una función diferenciable definida en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que*

$$f(\vec{x}) \geq f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y}) \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \tag{6.3.1}$$

$$\text{(resp. } f(\vec{x}) > f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y}) \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y}\text{)}$$

Demostración. Supongamos primero que f es convexa. Entonces para todo $\vec{x}, \vec{y} \in Q$ y todo $\lambda \in]0,1]$ se tiene de acuerdo a la desigualdad (6.2.1).

$$f(\vec{y} + \lambda(\vec{x} - \vec{y})) - f(\vec{y}) \leq \lambda[f(\vec{x}) - f(\vec{y})].$$

Dividiendo por λ y tomando límite cuando $\lambda \to 0$ con $\lambda > 0$, obtenemos la desigualdad $Df(\vec{y}; \vec{x} - \vec{y}) \leq f(\vec{x}) - f(\vec{y})$, que de acuerdo a la fórmula (5.3.5) es equivalente a (6.3.1).

Supongamos ahora que f verifica (6.3.1). Sean $\vec{x}, \vec{y} \in Q$ y $\vec{z} := \vec{y} + \lambda(\vec{x} - \vec{y})$ con $\lambda \in [0,1]$. Entonces

$$\begin{aligned} f(\vec{x}) &\geq f(\vec{z}) + Df(\vec{z})(\vec{x} - \vec{z}) \\ f(\vec{y}) &\geq f(\vec{z}) + Df(\vec{z})(\vec{y} - \vec{z}) \end{aligned}$$

y multiplicando estas desigualdades por λ y $(1 - \lambda)$ respectivamente y sumándola, se obtiene gracias a la linealidad de la función $Df(\vec{z})$

$$\begin{aligned} \lambda f(\vec{x}) + (1-\lambda) f(\vec{y}) &\geq f(\vec{z}) + Df(\vec{z})(\lambda(\vec{x} - \vec{z}) + (1-\lambda)(\vec{y} - \vec{z})) \\ &= f(\vec{z}) + Df(\vec{z})(\vec{0}) \\ &= f(\lambda\vec{x} + (1-\lambda)\vec{y}) \end{aligned}$$

que corresponde exactamente a la convexidad de f. $\square$

Nota 6.3.2. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, la fórmula (6.3.1) se escribe

$$f(\vec{x}) \geq f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \tag{6.3.2}$$

$$(\text{ resp. } f(\vec{x}) > f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \text{ con } \vec{x} \neq \vec{y}).$$

Nota 6.3.3. En la Nota 5.3.2 veíamos que la ecuación $z = h(\vec{x})$, donde $h : \vec{E} \to \vec{F}$ es la función definida por $h(\vec{x}) := f(\vec{y}) + Df(\vec{y})(\vec{x} - \vec{y})$ (con $\vec{y} \in Q$ fijo), representa el hiperplano afín tangente al grafo de una función diferenciable $f : Q \to \mathbb{R}$ en $(\vec{y}, f(\vec{y}))$. Esto nos muestra, de acuerdo al teorema anterior, que una función diferenciable f es convexa si y solo si para todo $\vec{y} \in Q$ el hiperplano afín tangente al grafo de f en $(\vec{y}, f(\vec{y}))$ está por debajo del grafo de f, es decir, para todo $\vec{y} \in Q$ la función lineal afín h minora a la función f.

Teorema 6.3.2. *Sea f una función diferenciable definida en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Una condición necesaria y suficiente para que f sea convexa (resp. estrictamente convexa) es que*

$$[Df(\vec{x}) - Df(\vec{y})](\vec{x} - \vec{y}) \geq 0 \quad \textit{para todo } \vec{x}, \vec{y} \in Q \tag{6.3.3}$$

$$(\textit{ resp. } [Df(\vec{x}) - Df(\vec{y})](\vec{x} - \vec{y}) > 0 \quad \textit{para todo} \quad \vec{x}, \vec{y} \in Q \quad \textit{con} \quad \vec{x} \neq \vec{y}).$$

Demostración. Supongamos que f es convexa. Aplicando dos veces el teorema anterior podemos escribir para todo $\vec{x}, \vec{y} \in Q$

$$\begin{array}{rcl} f(\vec{x}) & - & Df(\vec{y})(\vec{x}-\vec{y}) \geq f(\vec{y}) \\ f(\vec{y}) & + & Df(\vec{x})(\vec{x}-\vec{y}) \geq f(\vec{x}) \end{array}$$

sumando y teniendo en cuenta la linealidad de $Df(\vec{x})$ y $Df(\vec{y})$ se obtiene la desigualdad (6.3.3).

Supongamos ahora que f verifica (6.3.3). Dados $\vec{x}, \vec{y} \in Q$ definamos la función $\phi : [0,1] \to \mathbb{R}$ por $\phi(\lambda) := f(\vec{x} + \lambda(\vec{y} - \vec{x}))$. Del Teorema 5.5.1 vemos que ϕ es derivable y $\phi'(\lambda) = Df(\vec{x} + \lambda(\vec{y} - \vec{x}))(\vec{y} - \vec{x})$. Aplicando entonces (6.3.3) deducimos que para todo $\lambda \in [0,1]$ se tiene

$$\begin{array}{rcl} \phi'(\lambda) - \phi'(0) & = & [Df(\vec{x} + \lambda(\vec{y} - \vec{x})) - Df(\vec{x})](\vec{y} - \vec{x}) \\ & = & \dfrac{1}{\lambda}[Df(\vec{x} + \lambda(\vec{y} - \vec{x})) - Df(\vec{x})](\vec{x} + \lambda(\vec{y} - \vec{x}) - \vec{x}) \\ & \geq & 0. \end{array}$$

Aplicando ahora el teorema del valor medio a la función ϕ en $[0,1]$, vemos que existe $\bar{\lambda} \in]0,1[$ tal que $\phi(1) - \phi(0) = \phi'(\bar{\lambda})$ y de la desigualdad anterior concluimos que

$$\phi(1) - \phi(0) \geq \phi'(0)$$

que corresponde exactamente a la desigualdad (6.3.1). La función f será entonces convexa. □

Nota 6.3.4. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, la fórmula (6.3.3) se escribe

$$\langle \nabla f(\vec{x}) - \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle \geq 0 \quad \text{para todo} \quad \vec{x}, \vec{y} \in Q \tag{6.3.4}$$

$$(\text{ resp. } \langle \nabla f(\vec{x}) - \nabla f(\vec{y}), \vec{x} - \vec{y} \rangle > 0 \quad \text{para todo } \vec{x}, \vec{y} \in Q \quad \text{con} \quad \vec{x} \neq \vec{y}).$$

En el próximo teorema daremos una caracterización importante de la convexidad (y de la estricta convexidad) para funciones de clase $\mathcal{C}^2$. Como los vectores en $\mathbb{R}^n$ los denotamos por una columna, el producto de una matriz H de $n \times n$ con un vector $\vec{\delta}$ de $\mathbb{R}^n$ lo escribimos $H\vec{\delta}$ y el producto escalar del vector $H\vec{\delta}$ con un vector $\vec{u}$ en $\mathbb{R}^n$ lo escribiremos indistintamente $\vec{u}^t H\vec{\delta}$ o bien $\langle H\vec{\delta}, \vec{u} \rangle$, donde $\vec{u}^t$ representa el traspuesto del vector $\vec{u}$, es decir el vector $\vec{u}$ escrito como fila.

Teorema 6.3.3. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto convexo Q de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Una condición necesaria y suficiente para que f sea convexa (resp.*

estrictamente convexa) es que para todo $\vec{x} \in Q$ su Hessiano $H(\vec{x})$ sea definido no negativo (resp. definido positivo), es decir

$$\vec{\delta}^t H(\vec{x})\vec{\delta} \geq 0 \quad \textit{para todo} \quad \vec{\delta} \in \mathbb{R}^n \tag{6.3.5}$$

$$(\textit{ resp. } \vec{\delta}^t H(\vec{x})\vec{\delta} > 0 \quad \textit{para todo} \quad \vec{\delta} \in \mathbb{R}^n \quad \textit{con} \quad \vec{\delta} \neq \vec{0}).$$

Demostración. Supongamos que f es convexa. Dados $\vec{x} \in Q$ y $\vec{\delta} \in \mathbb{R}^n$ $(\vec{\delta} \neq \vec{0})$, definamos la función $\varphi : [0, \varepsilon] \to \mathbb{R}$, donde $\varepsilon > 0$ es tal que $\vec{x} + \varepsilon\vec{\delta} \in Q$, por $\varphi(\lambda) := \langle \nabla f(\vec{x} + \lambda\vec{\delta}), \vec{\delta} \rangle$. Como f es de clase $\mathcal{C}^2$, usando los teoremas 5.9.4 y 5.5.1 se deduce que φ es derivable y $\varphi'(\lambda) = \langle H(\vec{x} + \lambda\vec{\delta})\vec{\delta}, \vec{\delta} \rangle$. La desigualdad (6.3.5) equivale entonces a demostrar que $\varphi'(0) \geq 0$. Puesto que

$$\varphi'(0) = \lim_{\lambda \to 0} \frac{\varphi(\lambda) - \varphi(0)}{\lambda}$$

y como del teorema anterior se tiene

$$\begin{aligned} \varphi(\lambda) - \varphi(0) &= \langle \nabla f(\vec{x} + \lambda\vec{\delta}) - \nabla f(\vec{x}) \ , \ \vec{\delta} \rangle \\ &= \frac{1}{\lambda} \langle \nabla f(\vec{x} + \lambda\vec{\delta}) - \nabla f(\vec{x}), \vec{x} + \lambda\vec{\delta} - \vec{x} \rangle \\ &\geq 0 \end{aligned}$$

concluimos que $\varphi'(0) \geq 0$.

Supongamos ahora que para todo $\vec{x} \in Q$ la matriz $H(\vec{x})$ es definida nonegativa. Dados $\vec{x}, \vec{y} \in Q$, si definimos $\vec{\delta} := \vec{y} - \vec{x}$ y φ como en la parte anterior con $\varepsilon = 1$, vemos que aplicando el teorema del valor medio a φ en $[0, 1]$, existirá $\bar{\lambda} \in]0, 1[$ tal que

$$\varphi(1) - \varphi(0) = \varphi'(\bar{\lambda}).$$

Puesto que $\varphi'(\lambda) = \langle H(\vec{x} + \lambda\vec{\delta})\vec{\delta}, \vec{\delta} \rangle \geq 0$ para todo $\lambda \in]0, 1[$, concluimos que $\varphi(1) - \varphi(0) \geq 0$ lo que equivale exactamente a la desigualdad (6.3.3) del teorema anterior. La función f será entonces convexa. □

Nota 6.3.5. Del teorema anterior vemos que una forma cuadrática q definida en $\mathbb{R}^n$ con valores en $\mathbb{R}$, esto es $q(\vec{x}) := \alpha + \langle \vec{c}, \vec{x} \rangle + \vec{x}^t H \vec{x}$, es convexa si y solo si la matriz H es definida nonegativa.

Nota 6.3.6. Del teorema anterior y de la Nota 5.9.5 vemos que una función f de clase $\mathcal{C}^2$ definida en un abierto convexo Q de $\mathbb{R}^n$ con valores en $\mathbb{R}$, será convexa si y solo si en todo punto de Q la aproximación cuadrática de f es convexa.

6.4 Funciones Concavas

Definición 6.4.1. Una función f definida en una parte convexa Q de un e.v.$\vec{E}$ con valores en $\mathbb{R}$, se dirá cóncava (resp. estrictamente cóncava) en Q si la función $-f$ es convexa (resp. estrictamente convexa) en Q.

Nota 6.4.1. Todo lo dicho en las dos secciones anteriores para las funciones convexas (resp. estrictamente convexas) es válido para las funciones cóncavas (resp. estrictamente cóncavas) cambiando f por $-f$. Así entonces, en el Teorema 6.2.1 habrá que cambiar el término convexas por cóncavas; en el Teorema 6.2.2 habrá que cambiar el término convexas por cóncavas y el supremo se cambia por un ínfimo; en el Teorema 6.3.1 habrá que cambiar el término convexa por cóncava y el sentido de la desigualdad (6.3.1); en el Teorema 6.3.2 habrá que cambiar el término convexa por cóncava y el sentido de la desigualdad (6.3.3) y finalmente; en el Teorema 6.3.3 habrá que cambiar el término convexa por cóncava, $H(\vec{x})$ definido no negativo por $H(\vec{x})$ definido no positivo y el sentido de la desigualdad (6.3.5).

6.5 Mínimos y máximos de una función

Definición 6.5.1. Dada una función f definida en un conjunto A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$, un punto $\vec{x}^* \in A$ se dirá mínimo local (resp. mínimo local estricto) de f en una parte D de A si $\vec{x}^* \in D$ y existe $B(\vec{x}^*, \varepsilon)$ tal que

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D \tag{6.5.1}$$

$$(\text{ resp.} \quad f(\vec{x}^*) < f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D \quad \text{con} \quad \vec{x} \neq \vec{x}^*).$$

Definición 6.5.2. Si la desigualdad (6.5.1) de la definición anterior se verifica en sentido contrario, el punto $\vec{x}^* \in D$ se dirá máximo local (resp. máximo local estricto) de f en D.

Nota 6.5.1. De las definiciones anteriores se deduce fácilmente que un punto $\vec{x}^* \in D$ es máximo local (resp. máximo local estricto) de f en D si y solo si es mínimo local (resp. mínimo local estricto) de la función $-f$.

Definición 6.5.3. Dada una función f definida en un conjunto A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$, un punto $\vec{x}^* \in A$ se dirá mínimo (resp. mínimo estricto) de f en una parte D de A si $\vec{x}^* \in D$ y

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in D \tag{6.5.2}$$

$$(\text{ resp.} \quad f(\vec{x}^*) < f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in D \quad \text{con} \quad \vec{x} \neq \vec{x}^*).$$

Definición 6.5.4. Si la desigualdad (6.5.2) de la definición anterior se verifica en sentido contrario, el punto $\vec{x}^* \in D$ se dirá máximo (resp. máximo estricto) de f en D.

Nota 6.5.2. De las dos últimas definiciones se deduce fácilmente que un punto $\vec{x}^* \in D$ es un máximo (resp. máximo estricto) de f en D si y solo si es un mínimo (resp. mínimo estricto) de la función $-f$.

Nota 6.5.3. De las definiciones anteriores se deduce fácilmente que todo mínimo (resp. mínimo estricto) es también mínimo local (resp. mínimo local estricto) y que todo máximo (resp. máximo estricto) es también máximo local (resp. máximo local estricto).

Teorema 6.5.1. *Sea f una función convexa definida en una parte convexa Q de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Entonces, todo mínimo local (resp. mínimo local estricto) de f en una parte convexa D de Q será un mínimo (resp. mínimo estricto) de f en D.*

Demostración. Sea $\vec{x}^*$ un mínimo local de f en D. Existe entonces $B(\vec{x}^*, \varepsilon)$ tal que

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon) \cap D. \tag{*}$$

Para concluir debemos demostrar que la desigualdad anterior también se tiene para todo $\vec{x} \in D \setminus B(x^*, \varepsilon)$. Sea entonces $\vec{x} \in D$ con $\vec{x} \notin B(\vec{x}^*, \varepsilon)$. Como D es un conjunto convexo y $\lambda := \frac{\varepsilon}{\|\vec{x}-\vec{x}^*\|} \in [0,1]$, es fácil ver que $\lambda\vec{x} + (1-\lambda)\vec{x}^* \in D \cap B(\vec{x}^*, \varepsilon)$. Así entonces, de (*) y de la convexidad de f, se tiene

$$f(\vec{x}^*) \leq f(\lambda\vec{x} + (1-\lambda)\vec{x}^*) \leq \lambda f(\vec{x}) + (1-\lambda)f(\vec{x}^*) \leq f(\vec{x})$$

que es lo que queriamos demostrar. □

Teorema 6.5.2. *Sea f una función diferenciable definida en un abierto A de un e.v.n.$\vec{E}$ con valores en $\mathbb{R}$. Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en una parte convexa D de A, es que $\vec{x}^* \in D$ y*

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \quad \textit{para todo } \vec{x} \in D. \tag{6.5.3}$$

Demostración. Sea $\vec{x}^*$ un mínimo local de f en D. Dado $\vec{x} \in D$ definamos la función $\phi : [0,1] \to \mathbb{R}$ por $\phi(\lambda) := f(\vec{x}^* + \lambda(\vec{x} - \vec{x}^*))$. Del Teorema 5.5.1 vemos que ϕ es derivable y que $\phi'(0) = Df(\vec{x}^*)(\vec{x} - \vec{x}^*)$. Por otra parte, como $\phi(\lambda) \geq \phi(0)$ para todo $\lambda \in [0,1]$, concluimos que

$$\phi'(0) = \lim_{\substack{\lambda \to 0 \\ \lambda > 0}} \frac{\phi(\lambda) - \phi(0)}{\lambda} \geq 0$$

lo que corresponde exactamente a la desigualdad (6.5.3). □

Nota 6.5.4. De acuerdo a la interpretación que se dió en la Nota 5.2.1 a la cantidad $Df(\vec{x}^*; \vec{x} - \vec{x}^*) = Df(\vec{x}^*)(\vec{x} - \vec{x}^*)$ (ver fórmula (5.3.5)), vemos que la desigualdad (6.5.3) corresponde al hecho que las pendientes de f en $\vec{x}^*$, en todas las direcciones interiores al convexo D, son no negativas.

Nota 6.5.5. Es fácil ver que la condición (6.5.3) no es suficiente para que $\vec{x}^* \in D$ sea un mínimo local de f en D. Un ejemplo simple que muestra este hecho está dado por la función $f : \mathbb{R} \to \mathbb{R}$ definida por $f(\vec{x}) = x^3$, que verifica (6.5.3) para $\vec{x}^* = 0$.

El próximo teorema nos muestra que para una clase particular de funciones esta condición es suficiente para que $\vec{x}^* \in D$ sea mínimo.

Nota 6.5.6. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$, la desigualdad (6.5.3) se escribe

$$\langle \nabla f(\vec{x}^*), \vec{x} - \vec{x}^* \rangle \geq 0 \quad \text{para todo} \quad \vec{x} \in D. \tag{6.5.4}$$

Teorema 6.5.3. *Sea f una función convexa y diferenciable definida en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$. Una condición necesaria y suficiente para que $\vec{x}^* \in Q$ sea un mínimo de f en una parte convexa D en Q, es que $\vec{x}^* \in D$ y verifique* (6.5.3).

Demostración. Del teorema anterior y, aplicando el Teorema 6.5.1, deducimos que la condición es necesaria.

Supongamos ahora que $\vec{x}^* \in D$ verifica (6.5.3). Puesto que f es convexa, del Teorema 6.3.1 y de (6.5.3) se deduce

$$f(\vec{x}) \geq f(\vec{x}^*) + Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq f(\vec{x}^*) \quad \text{para todo} \quad \vec{x} \in D$$

que es lo que queriamos probar. □

Teorema 6.5.4. *Sea f una función diferenciable definida en un abierto A de un e.v.n.$\vec{E}$ con valores en $\mathbb{R}$. Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en A es que*

$$Df(\vec{x}^*) = 0 \tag{6.5.5}$$

donde 0 *representa la función nula en $\mathcal{L}(\vec{E}, \mathbb{R})$.*

Demostración. Sea $\vec{x}^* \in A$ un mínimo local de f en A. Puesto que A es un conjunto abierto, existe $B(\vec{x}^*, \varepsilon) \subset A$ tal que

$$f(\vec{x}^*) \leq f(\vec{x}) \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon)$$

y como $B(x^*, \varepsilon)$ es un conjunto convexo, usando el Teorema 6.5.2, concluimos que

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \quad \text{para todo} \quad \vec{x} \in B(\vec{x}^*, \varepsilon).$$

Haciendo el cambio de variable $\vec{v} = \vec{x} - \vec{x}^*$ en la relación anterior, vemos que ella es equivalente a

$$Df(\vec{x}^*)(\vec{v}) \geq 0 \quad \text{para todo} \quad \vec{v} \in B(\vec{0}, \varepsilon) \tag{*}$$

lo que solo puede ocurrir si $Df(\vec{x}^*)$ es la función nula. En efecto, si existiera $\vec{z} \in \vec{E}$ tal que $Df(\vec{x}^*)(\vec{z}) > 0$, definiendo $\vec{v} := -\varepsilon \frac{\vec{z}}{\|\vec{z}\|}$ tendríamos que $\vec{v} \in B(\vec{0}, \varepsilon)$ y

$$Df(\vec{x}^*)(\vec{v}) = -\frac{\varepsilon}{\|\vec{z}\|} Df(\vec{x}^*)(\vec{z}) < 0$$

lo que contradice la relación (*). □

Nota 6.5.7. De acuerdo a la interpretación que se dió en la Nota 5.2.1 a la cantidad $Df(\vec{x}^*; \vec{v}) = Df(\vec{x}^*)(\vec{v})$ (ver fórmula (5.3.5)), vemos que la igualdad (6.5.5) corresponde al hecho que las pendientes de f en $\vec{x}^*$, en todas las direcciones, son nulas.

De acuerdo a la Nota 5.3.2, la relación (6.5.5) equivale a decir que la aproximación lineal afín de f en $\vec{x}^*$ está dada por la función constante $h(\vec{x}) = f(\vec{x}^*)$. Geométricamente, (6.5.5) equivale a decir que en $\vec{E} \times \mathbb{R}$ el hiperplano afín tangente al grafo de f en $(\vec{x}^*, f(\vec{x}^*))$, es el hiperplano horizontal de ecuación $z = f(\vec{x}^*)$.

Nota 6.5.8. El mismo ejemplo de la Nota 6.5.5 nos muestra que la relación (6.5.5) no es suficiente para que $\vec{x}^* \in A$ sea un mínimo local de f en A.

El próximo teorema nos muestra que para una clase particular de funciones esta condición es suficiente para que $\vec{x}^*$ sea un mínimo de f en el abierto A.

Nota 6.5.9. Es importante tener claro que el teorema anterior no es válido si $\vec{x}^*$ es un mínimo local de f en una parte convexa D de A. En efecto, si definimos la función $f : \mathbb{R} \to \mathbb{R}$ por $f(x) = x^2$ y el conjunto $D = [1, 2]$, es evidente que $x^* = 1$ es un mínimo de f en D y sin embargo $Df(1)(v) = 2v$ no es la función nula en $\mathcal{L}(\mathbb{R}, \mathbb{R})$.

Nota 6.5.10. Si en el teorema anterior $\vec{E} = \mathbb{R}^n$ de acuerdo a la fórmula (5.3.13), la igualdad (6.5.5) es equivalente a

$$\nabla f(\vec{x}) = \vec{0}. \tag{6.5.6}$$

Teorema 6.5.5. *Sea f una función convexa y diferenciable definida en un abierto convexo Q de un e.v.n.$\vec{E}$ con valores en $\mathbb{R}$. Una condición necesaria y suficiente para que $\vec{x}^* \in Q$ sea mínimo de f en Q es que se tenga la relación* (6.5.5).

Demostración. Del teorema anterior, aplicando el Teorema 6.5.3 con $D = Q$, deducimos que la condición es necesaria.

Supongamos ahora que $\vec{x}^* \in Q$ verifica (6.5.5). Puesto que f es convexa, del Teorema 6.3.1 y de (6.5.5) se deduce

$$f(\vec{x}) \geq f(\vec{x}^*) + Df(\vec{x}^*)(\vec{x} - \vec{x}^*) = f(\vec{x}^*) \quad \text{para todo} \quad \vec{x} \in Q$$

que es lo que queriamos probar. $\square$

Teorema 6.5.6. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Una condición necesaria para que $\vec{x}^* \in A$ sea un mínimo local de f en A es que el Hessiano $H(\vec{x}^*)$ sea definido no negativo.*

Demostración. Sea $\vec{x}^*$ un mínimo de f en A. Como A es abierto, existe $B(\vec{x}^*, \varepsilon) \subset A$. Del Teorema 5.9.3 y la Definición 5.9.2 deducimos que para todo $\vec{\delta} \in B(\vec{0}, \varepsilon)$ se tiene

$$f(\vec{x}^* + \vec{\delta}) = f(\vec{x}^*) + \langle \nabla f(\vec{x}^*), \delta \rangle + \frac{1}{2}\vec{\delta}^t H(\vec{x}^*)\vec{\delta} + o^2(\vec{\delta})$$

y como f es convexa, aplicando la fórmula (6.3.2), obtenemos

$$\vec{\delta}^t H(\vec{x}^*)\vec{\delta} \geq -2o^2(\vec{\delta}) \quad \text{para todo} \quad \vec{\delta} \in B(\vec{0}, \varepsilon) \tag{*}$$

de lo cual se desprende fácilmente que $H(\vec{x}^*)$ es definida no negativa.

En efecto, sea $\vec{d} \in \mathbb{R}^n$ diferente de $\vec{0}$, aplicando la desigualdad (*) a $\vec{\delta} := \lambda\vec{d}$, donde $0 < |\lambda| \leq \frac{\varepsilon}{\|\vec{d}\|}$ de modo que $\vec{\delta} \in B(0, \varepsilon)$, se deduce

$$\vec{d}^t H(\vec{x}^*)\vec{d} \geq -2\|\vec{d}\|^2 \frac{o^2(\lambda\vec{d})}{\|\lambda\vec{d}\|^2}$$

y tomando límite cuando λ tiende a 0, de la igualdad (5.9.3) obtenemos la desigualdad

$$\vec{d}^t H(\vec{x}^*)\vec{d} \geq 0$$

que es lo que queriamos demostrar. $\square$

Nota 6.5.11. El mismo ejemplo de la Nota 6.5.5 nos muestra que la condición necesaria para mínimo local dada por el teorema anterior, no es una condición suficiente.

Nota 6.5.12. El mismo ejemplo de la Nota 6.5.9 nos muestra que la condición necesaria para mínimo local dada por el teorema anterior, deja de serlo si $\vec{x}^*$ es un mínimo local de f en una parte D del abierto A.

Teorema 6.5.7. *Sea f una función de clase $\mathcal{C}^2$ definida en un abierto A de $\mathbb{R}^n$ con valores en $\mathbb{R}$. Una condición suficiente para que $\vec{x}^* \in A$ sea un mínimo local estricto de f en A es que $\nabla f(\vec{x}^*) = 0$ y $H(\vec{x}^*)$ sea definido positivo.*

Demostración. Vamos a demostrar primero que si H es una matriz $n \times n$ definida positiva, entonces existe $\alpha > 0$ tal que

$$\vec{\delta}^t H \vec{\delta} \geq \alpha \|\vec{\delta}\|^2 \quad \text{para todo} \quad \vec{\delta} \in \mathbb{R}^n. \tag{*}$$

Para esto, definimos la función $\phi : \mathbb{R}^n \to \mathbb{R}$ por $\phi(\vec{\delta}) = \vec{\delta}^t H \vec{\delta}$. Puesto que ϕ es una función continua en el compacto $c = \{\vec{\delta} \in \mathbb{R}^n : \|\vec{\delta}\| = 1\}$ y además $\phi(\vec{\delta}) > 0$ para todo $\vec{\delta} \in C$, concluimos que existe $\alpha > 0$ tal que

$$\phi(\vec{\delta}) \geq \alpha \quad \text{para todo} \quad \vec{\delta} \in C$$

lo que es equivalente a (*) puesto que para todo $\vec{\delta} \in \mathbb{R}^n$ se tiene que $\frac{\vec{\delta}}{\|\vec{\delta}\|} \in C$ y $\phi(\frac{\vec{\delta}}{\|\vec{\delta}\|}) = \frac{\phi(\vec{\delta})}{\|\vec{\delta}\|^2}$.

Tal como decíamos en la demostración del teorema anterior, como A es abierto existe $B(\vec{x}^*, \varepsilon) \subset A$ y, del Teorema 5.9.3 y la Definición 5.9.2 deducimos que para todo $\vec{\delta} \in B(\vec{0}, \varepsilon)$

$$f(\vec{x}^* + \vec{\delta}) = f(\vec{x}^*) + \langle \nabla f(\vec{x}^*), \vec{\delta} \rangle + \frac{1}{2} \vec{\delta}^t H(\vec{x}^*) \vec{\delta} + o^2(\vec{\delta}).$$

Aplicando ahora a la matriz $H(\vec{x}^*)$ lo que demostramos en la parte anterior y del hecho que $\nabla f(\vec{x}^*) = \vec{0}$, vemos que existirá $\alpha > 0$ tal que

$$f(\vec{x}^* + \vec{\delta}) \geq f(\vec{x}^*) + \alpha \|\vec{\delta}\|^2 + o^2(\vec{\delta}) \quad \text{para todo} \quad \vec{\delta} \in B(\vec{0}, \varepsilon).$$

Para terminar la demostración bastará con mostrar que existe $\varepsilon' \in]0, \varepsilon]$ tal que $\alpha \|\vec{\delta}\|^2 + o^2(\vec{\delta}) > 0$ para todo $\vec{\delta} \in B(\vec{0}, \varepsilon')$ diferente de $\vec{0}$. Si no fuera así, existiría una sucesión $\vec{\delta}_n$ convergente a $\vec{0}$, con $\vec{\delta}_n \neq \vec{0}$ para todo n, tal que $\alpha \|\vec{\delta}_n\|^2 + o^2(\vec{\delta}_n) \leq 0$ para todo n; dividiendo esta desigualdad por $\|\vec{\delta}\|^2$ y tomando límite sobre n obtenemos que $\alpha \leq 0$, lo que muestra una contradicción. $\square$

Nota 6.5.13. La condición suficiente para mínimo local estricto dada por el teorema anterior, no es una condición necesaria como lo muestra la función $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x^4$ que tiene como mínimo local estricto a $\vec{x}^* = 0$ y en ese punto el Hessiano no es definido positivo, en efecto $H(0) = 0$.

Nota 6.5.14. De las notas 6.5.1 y 6.5.2 deducimos fácilmente que los siete teoremas que hemos visto en esta sección se pueden enunciar cambiando el término “mínimo” por “máximo”, el termino “función convexa” por “función cóncava” y el sentido de las desigualdades (a este último tipo de cambio corresponde el reemplazo del término “definido no negativo” por “definido no positivo” y “definido positivo” por “definido negativo” en los teoremas 6.5.6 y 6.5.7 respectivamente).

6.6 Mínimos con restricciones de tipo desigualdad. Teorema de Kuhn-Tucker

Nota 6.6.1. Los seis teoremas que en la sección anterior establecen condiciones necesarias y/o suficientes para mínimo local, se pueden dividir en dos grupos, los dos primeros (teoremas 6.5.2 y 6.5.3) y los cuatro últimos (teoremas 6.5.4 a 6.5.7). En los dos primeros, el mínimo o el mínimo local en cuestión está restringido a un subconjunto D del dominio de la función f. En los restantes, se trata de un mínimo o mínimo local no restringido, es decir en todo el dominio de la función f.

En esta sección vamos a dar una condición necesaria (y suficiente si la función es convexa) para mínimo local restringido de una función cuando el conjunto D toma una forma particular de gran importancia en el estudio de modelos matemáticos de la física y de la ingeniería.

Problema: Dadas n funciones convexas diferenciables $f_1, ..., f_n$, definidas en un abierto convexo Q de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$, vamos a estudiar los mínimos y mínimos locales de una función diferenciable $f_0 : Q \to \mathbb{R}$, en el conjunto D definido por

$$D := \{\vec{x} \in Q : f_1(\vec{x}) \leq 0, ..., f_n(\vec{x}) \leq 0\}. \tag{6.6.1}$$

Las funciones $f_1, ..., f_n$ se llaman restricciones del problema y f_0 se llama función objetivo.

Definición 6.6.1. Si D es el conjunto definido en (6.6.1), dado $\vec{x}^* \in D$ definimos en $\vec{E}$ los conjuntos

$$\mathcal{D}(\vec{x}^*) := \{\vec{v} \in \vec{E} : \exists \mu > 0 \quad \text{tal que} \quad f_i(\vec{x}^* + \mu\vec{v}) \leq 0 \quad \text{para todo} \quad i : 1, ..., n\} \tag{6.6.2}$$

$$\mathcal{T}(\vec{x}^*) := \{\vec{v} \in \vec{E} : Df_i(\vec{x}^*)(\vec{v}) \leq 0 \quad \text{para todo} \quad i \in I(\vec{x}^*)\} \tag{6.6.3}$$

donde

$$I(\vec{x}^*) = \{i : f_i(\vec{x}^*) = 0\}. \tag{6.6.4}$$

Diremos entonces que las restricciones del problema en cuestión verifica la hipótesis de calificación en $\vec{x}^* \in D$, si se tiene la igualdad

$$\overline{D(\vec{x}^*)} = \mathcal{T}(\vec{x}^*). \tag{6.6.5}$$

Nota 6.6.2. En los teoremas 6.6.2 y 6.6.3 daremos condiciones suficientes bastante simples para que nuestro problema verifique la hipótesis de calificación de las restricciones en un punto $\vec{x}^* \in D$. Para esto necesitaremos tres lemas previos que damos a continuación.

Lema 6.6.1. *El conjunto D definido por* (6.6.1) *es convexo.*

Demostración. $\vec{a}, \vec{b} \in D \Rightarrow f_i(\vec{a}) \leq 0$ y $f_i(\vec{b}) \leq 0$ para todo $i = 1, ..., n$

$\Rightarrow \lambda f_i(\vec{a}) \leq 0$ y $(1-\lambda) f_i(\vec{b}) \leq 0$ para todo $\lambda \in [0,1]$ y todo $i = 1, ..., n$

$\Rightarrow f_i(\lambda\vec{a} + (1-\lambda)\vec{b}) \leq 0$ para todo $\lambda \in [0,1]$ y todo $i = 1, ..., n$

$\Rightarrow \lambda\vec{a} + (1-\lambda\vec{b}) \in D$ para todo $\lambda \in [0,1]$. $\square$

Lema 6.6.2. *Si f es una función convexa definida en un convexo Q de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$ y si $\vec{x}^* \in Q$ y $\vec{v} \in \vec{E}$ verifican las desigualdades $f(\vec{x}^*) \leq 0$ y $f(\vec{x}^* + \mu\vec{v}) \leq 0$ para algún $\mu > 0$, entonces*

$$f(\vec{x}^* + \lambda\vec{v}) \leq 0 \quad \textit{para todo} \quad \lambda \in [0, \mu]. \tag{6.6.6}$$

Demostración. Del lema anterior sabemos que el conjunto $D = \{\vec{x} \in Q : f(\vec{x}) \leq 0\}$ es convexo. Para concluir basta ver que $\vec{x}^* + \lambda\vec{v} = \frac{\lambda}{\mu}\vec{x}^* + (1 - \frac{\lambda}{\mu})(\vec{x}^* + \mu\vec{v}) \in D$. $\square$

Lema 6.6.3. *Si f es una función diferenciable definida en un abierto A de un e.v.n. $\vec{E}$ con valores en $\mathbb{R}$, y si $\vec{x}^* \in A$ y $\vec{v} \in \vec{E}$ verifican la desigualdad $Df(\vec{x}^*)(\vec{v}) < 0$, entonces existe $\mu > 0$ tal que*

$$f(x^* + \lambda\vec{v}) < f(\vec{x}^*) \quad \textit{para todo} \quad \lambda \in]0, \mu]. \tag{6.6.7}$$

Demostración. Denotemos $r := Df(\vec{x}^*)(\vec{v}) < 0$. De la fórmula (5.3.5) y de la definición de límite vemos que existe $\mu > 0$ tal que

$$\frac{f(\vec{x}^* + \lambda\vec{v}) - f(\vec{x}^*)}{\lambda} \leq \frac{r}{2} \quad \text{para todo } \lambda \in]0, \mu].$$

Puesto que $\frac{r}{2} < 0$, multiplicando por λ obtenemos la desigualdad (6.6.7). $\square$

Teorema 6.6.1. *Los conjuntos* $\mathcal{D}(\vec{x}^*)$ *y* $\mathcal{T}(\vec{x}^*)$ *definidos por* (6.6.2) *y* (6.6.3) *respectivamente verifican la inclusión*

$$\overline{\mathcal{D}(\vec{x}^*)} \subset \mathcal{T}(x^*). \tag{6.6.8}$$

Demostración. Puesto que $\mathcal{T}(\vec{x}^*)$ es un conjunto cerrado, será suficiente demostrar que $\mathcal{D}(\vec{x}^*) \subset \mathcal{T}(\vec{x}^*)$. Sea $\vec{v} \in \mathcal{D}(\vec{x}^*)$ y $\mu > 0$ tal que $f_i(\vec{x}^* + \mu\vec{v}) \leq 0$ para todo $i \in I(\vec{x}^*)$. Del Lema 6.6.2 se tiene para cada función f_i con $i \in I(\vec{x}^*)$ la relación (6.6.6), entonces dividiendo por $\lambda \in]0,1]$ y tomando límite cuando $\lambda \to 0$, de la fórmula (5.3.5) obtenemos que $Df_i(\vec{x}^*)(\vec{v}) \leq 0$ para todo $i \in I(\vec{x}^*)$. Esto nos muestra que $\vec{v} \in \mathcal{T}(\vec{x}^*)$. $\square$

Teorema 6.6.2. *Una condición suficiente para que las n restricciones de nuestro problema verifiquen la llamada hipótesis de calificación en $\vec{x}^* \in D$, dada por la igualdad* (6.6.5), *es que exista $\vec{x}_0 \in \vec{E}$ tal que*

$$f_i(\vec{x}_0) < 0 \quad \textit{para todo} \quad i = 1, ..., n \tag{6.6.9}$$

(la relación (6.6.9) *se llama condición de Slater y, es equivalente a decir que el interior del conjunto D definido por* (6.6.1) *es no vacío).*

Demostración. Si $I(\vec{x}^*) = \phi$, el resultado es evidente. Supondremos entonces $I(\vec{x}^*) \neq \phi$. Como la inclusión (6.6.8) se tiene siempre, sólo debemos demostrar la contraria.

Dado $\vec{v} \in \mathcal{T}(\vec{x}^*)$ probaremos que $\vec{v} \in \overline{\mathcal{D}(\vec{x}^*)}$. Dividiremos la demostración en tres partes.

(i) De la relación (6.6.9), usando el Teorema 6.3.1 vemos que $0 > f_i(\vec{x}_0) \geq f_i(\vec{x}^*) + Df_i(\vec{x}^*)(\vec{x}_0 - \vec{x}^*)$ para todo $i = 1, ..., n$ y como $f_i(\vec{x}^*) = 0$ para todo $i \in I(\vec{x}^*)$, definiendo $\vec{v}_0 := \vec{x}_0 - \vec{x}^*$, concluimos que

$$Df_i(\vec{x}^*)(\vec{v}_0) < 0 \quad \text{para todo} \quad i \in I(\vec{x}^*).$$

(ii) Es fácil ver ahora que si definimos $\vec{v}_k := \vec{v} + \frac{1}{k}\vec{v}_0$ se tiene para todo $k \in \mathbb{N}$ y todo $i \in I(\vec{x}^*)$

$$\begin{aligned} Df_i(\vec{x}^*)(\vec{v}_k) &= Df(\vec{x}^*)(\vec{v} + \frac{1}{k}\vec{v}_0) \\ &= Df(\vec{x}^*)(\vec{v}) + \frac{1}{k}Df(\vec{x}^*)(\vec{v}_0) \\ &< 0 \qquad \text{para todo} \quad i \in I(\vec{x}^*) \end{aligned}$$

y, de acuerdo al Lema 6.6.3, esto implica que para cada $k \in \mathbb{N}$ y cada $i \in I(\vec{x}^*)$ existe $u_{k,i} > 0$ tal que

$$f_i(\vec{x}^* + \lambda\vec{v}_k) < 0 \text{ para todo } \lambda \in]0, u_{k,i}]. \tag{A}$$

(iii) Por otra parte, si $i \notin I(\vec{x}^*)$ se tiene que $f_i(\vec{x}^*) < 0$ y, por continuidad de las funciones f_i en $\vec{x}^*$, deducimos que para cada $k \in \mathbb{N}$ y cada $i \notin I(\vec{x}^*)$ existe $\mu_{k,i} > 0$ tal que

$$f_i(\vec{x}^* + \lambda \vec{v}_k) < 0 \quad \text{para todo} \quad \lambda \in]0, \mu_{k,i}]. \tag{B}$$

Si para cada $k \in \mathbb{N}$ definimos $\mu_k := \min_{i=1,\dots,n} \mu_{k,i} > 0$, de las relaciones (A) y (B) vemos que

$$f_i(\vec{x}^* + \mu_k \vec{v}_k) < 0$$

lo cual muestra que $\vec{v}_k \in \mathcal{D}(\vec{x}^*)$ para todo $k \in \mathbb{N}$. Como $\vec{v} = \lim \vec{v}_k$, concluimos que $\vec{v} \in \overline{\mathcal{D}(\vec{x}^*)}$.

$\square$

Teorema 6.6.3. *Cuando $\vec{E}$ es un espacio de Hilbert, una condición suficiente para que las n restricciones de nuestro problema verifiquen la hipótesis de calificación en $\vec{x}^* \in D$, dado por* (6.6.5)*, es que $Df_i(\vec{x}^*)$ con $i \in I(\vec{x}^*)$ sean linealmente independientes.*

Demostración. No es difícil demostrar que si $Df_i(\vec{x}^*)$ con $i \in I(\vec{x}^*)$ son linealmente independientes, entonces debe existir $\vec{v}_0 \in \vec{E}$ tal que $Df_i(\vec{x}^*)(\vec{v}_0) < 0$ para todo $i \in I(\vec{x}^*)$. La demostración sigue entonces como la del teorema anterior (partes (ii) y (iii)). $\square$

Teorema 6.6.4 (Karush-Kuhn-Tucker)**.** *Si las restricciones de nuestro problema verifican la hipótesis de calificación en $\vec{x}^* \in D$, dada por la igualdad* (6.6.5)*, entonces una condición necesaria para que $\vec{x}^* \in D$ sea un mínimo local de f_0 en D es que existan escalares $\lambda_i \geq 0$ para $i \in I(\vec{x}^*)$, llamados multiplicadores de Karush-Kuhn-Tucker, tales que*

$$Df_0(\vec{x}^*) + \sum_{i \in I(\vec{x}^*)} \lambda_i Df_i(\vec{x}^*) = 0. \tag{6.6.10}$$

Si f_0 es una función convexa en Q, entonces (6.6.10) *es una condición suficiente para que $\vec{x}^* \in D$ sea mínimo de f_0 en D.*

Demostración. Demostremos primero que cuando f_0 es convexa en Q, entonces (6.6.10) es una condición suficiente para que $\vec{x}^* \in D$ sea mínimo de f_0 en D. Para ésto, de acuerdo al Teorema 6.5.3, bastará demostrar que $Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0$ para todo $\vec{x} \in D$. Como todas las restricciones de nuestro problema son convexas, de acuerdo al Teorema 6.3.1, la igualdad (6.6.10) implica para todo $\vec{x} \in D$.

$$\begin{aligned} Df_0(\vec{x}^*)(\vec{x} - \vec{x}^*) &= -\sum \lambda_i Df_i(\vec{x}^*)(\vec{x} - \vec{x}^*) \\ &\geq \sum \lambda_i (f_i(\vec{x}^*) - f_i(\vec{x})) \\ &= -\sum \lambda_i f_i(\vec{x}) \geq 0 \end{aligned}$$

que es lo que queríamos probar.

ii) Supongamos ahora que $\vec{x}^* \in D$ es un mínimo local de f_0 en D, donde se verifica la hipótesis de calificación de las restricciones. Del Teorema 6.5.2 sabemos que

$$Df(\vec{x}^*)(\vec{x} - \vec{x}^*) \geq 0 \text{ para todo } \quad \vec{x} \in D \tag{A}$$

y no es difícil demostrar, usando el Lema 6.6.3, que esto equivale a decir que

$$Df(\vec{x}^*)(\vec{v}) \geq 0 \quad \text{ para todo } \quad \vec{v} \in \mathcal{D}(\vec{x}^*).$$

Usando entonces la hipótesis de calificación de las restricciones concluimos que (A) equivale a la inclusión

$$\{\vec{v} \in \vec{E} : -Df(\vec{x}^*)(\vec{v}) \leq 0\} \subset \{\vec{v} \in \vec{E} : Df_i(\vec{x}^*)(\vec{v}) \leq 0 \text{ para todo } i \in I(\vec{x}^*)\}.$$

El Teorema 4.7.3 (Lema de Farkas) nos permite concluir. $\square$

CAPÍTULO 7

INTEGRACIÓN

7.1 Introducción

Este último capítulo está dedicado a la Teoriía de integración de Lebesgue en $\mathbb{R}^n$. En la proxima sección se introduce la medida de Lebesgue definida a partir de conjuntos bastante simples como los n-rectángulos hasta llegar a definirla sobre una familia de conjuntos que llamaremos medibles. Luego, se introducen los primeros conceptos de integración donde destaca el importante y útil Teorema de Fubini, el cual a grandes rasgos permite ver una integral en varias variables como varias integrales en una sola variable. Finalmente se expondrán diversos resultados que permitirán facilitar el cálculo de integrales, como por ejemplo el teorema de cambio de variable.

7.2 Medida de Lebesgue en $\mathbb{R}^n$

Definición 7.2.1. Sean $\vec{a}$ y $\vec{b} \in \mathbb{R}^n$ tales que $a_i \leq b_i$ para todo $i = 1, ..., n$. Se define el *n-rectángulo* $[\vec{a}, \vec{b}]$ por

$$[\vec{a}, \vec{b}] = \{\vec{x} \in \mathbb{R}^n \ : \ a_i \leq x_i \leq b_i \quad i = 1, ..., n\}.$$

Su *medida* μ_0 estará dada por

$$\mu_0([\vec{a}, \vec{b}]) = (b_1 - a_1)(b_2 - a_2) \dots (b_n - a_n) = \prod_{i=1}^{n} (b_i - a_i).$$

Nota 7.2.1. De la definición anterior, se deduce que $\{\vec{a}\}$ es un n-rectángulo y que $\mu_0(\{\vec{a}\}) = 0$

Definición 7.2.2. Sean $P_1, P_2, \ldots, P_n$, n particiones, es decir, $P_i = \{x_0^i, x_1^i, \ldots, x_{m_i}^i\}$ donde $x_0^i < x_1^i < \ldots < x_{m_i}^i$. Se llama *reticulado* a un conjunto de la forma

$$\mathcal{R} = \bigcup_{j_i \in \{1,\ldots,m_i\}} [x_{j_1-1}^1, x_{j_1}^1] \times [x_{j_2-1}^2, x_{j_2}^2] \times \ldots \times [x_{j_n-1}^n, x_{j_n}^n].$$

Nota 7.2.2. Observe que el reticulado descrito en la definición anterior, contiene $\prod_{i=1}^{n} m_i$ n-rectángulos.

Definición 7.2.3. Una *figura* es un conjunto de $\mathbb{R}^n$ que proviene de escoger algunos n-rectángulos de algún reticulado.

Dada una figura $F = \bigcup_{k=1}^{m} I_k$ donde cada I_k es un n-rectángulo, se define su medida por

$$\mu_1(F) = \sum_{k=1}^{m} \mu_0(I_k).$$

Nota 7.2.3. Se demuestra facilmente que para toda figura F, F_1 y F_2 se tenien las siguientes propiedades:

1. $\mu_1(F) \geq 0$;
2. $F_1 \subset F_2 \Rightarrow \mu_1(F_1) \leq \mu_1(F_2)$;
3. $F_1 \cap F_2 = \emptyset \Rightarrow \mu_1(F_1 \cup F_2) = \mu_1(F_1) + \mu_1(F_2)$.

Definición 7.2.4. Sea $\theta \subset \mathbb{R}^n$ un conjunto abierto no vacío. Se define su medida por

$$\mu_2 = \sup\{\mu_1(F) \ : \ F \ \text{ es una figura }, \ F \subset \theta\}. \tag{7.2.1}$$

Nota 7.2.4. Dado que un conjunto abierto no vacío siempre contiene a una bola, es claro que el conjunto de las medidas de las figuras inscritas es no vacío, es decir,

$$\emptyset \neq \{\mu_1(F) \ : \ F \ \text{ es una figura }, \ F \subset \theta\} = J.$$

El conjunto J puede o no ser acotado superiormente. Si así fuese, el supremo en (7.2.1) será un número real mayor o igual que cero ($\mu_2(\theta) \in \mathbb{R}_+$). Si el conjunto J no es acotado superiormente, diremos que la medida de θ es $+\infty$, con la convención $+\infty \geq y$ para todo $y \in \mathbb{R}$. En general, se tendrá que la medida de un abierto pertenece a $\mathbb{R} \cup \{+\infty\}$.

Definición 7.2.5. Sea $A \subset \mathbb{R}^n$ un conjunto acotado. Diremos que A es medible si

$\forall\, \varepsilon > 0, \exists\, K,\ \theta \subset \mathbb{R}^n,\ K$ compacto, θ abierto, donde $K \subset A \subset \theta$ tales que $\mu_2(\theta \backslash K) \leq \varepsilon$.

En tal caso, definimos la medida de A por

$$\mu_3(A) = \text{ínf}\{\mu_2(\theta)\ :\ \theta \text{ abierto, } A \subset \theta\}. \tag{7.2.2}$$

Nota 7.2.5. Se prueba fácilmente que las propiedades indicadas en la Nota 7.2.3 para la medida μ_1 también son válidas para μ_2 y μ_3.

Teorema 7.2.1. *Todo conjunto abierto acotado $\theta \subset \mathbb{R}^n$ es medible y se tiene que $\mu_3(\theta) = \mu_2(\theta)$.*

Demostración. Sea $\varepsilon > 0$, y $F_1 \subset \theta$ una figura, la cual será compacta, tal que $\mu_1(F_1) \geq \mu_2(\theta) - \varepsilon$. Estimemos la medida μ_2 del abierto $\theta \setminus F_1$. Consideremos F una figura tal que $F \subset \theta \setminus F_1$, así $F \subset \theta$ y $F_1 \cap F = \emptyset$. De la Nota 7.2.3 tenemos que $\mu_1(F_1 \cup F) = \mu_1(F_1) + \mu_1(F)$ y debido a la Nota 7.2.5 podemos escribir $\mu_1(F) \leq \mu_2(\theta) - \mu_1(F_1) \leq \varepsilon$. Como esta última desigualdad se tiene para toda figura $F \subset \theta \setminus F_1$ concluímos que $\mu_2(\theta \setminus F_1) \leq \varepsilon$ y por lo tanto θ es medible. La igualdad $\mu_3(\theta) = \mu_2(\theta)$ es evidente de la definición (7.2.2). $\square$

La pregunta que naturalmente uno se puedo plantear es si todos los conjuntos acotados son medibles. La respuesta es no, como lo muestra el siguiente ejemplo en $\mathbb{R}$.

Ejemplo 7.2.1. Sea $\sim$ la relación de equivalencia en $[0, 1]$ definida por

$$x \sim y \Leftrightarrow x - y \text{ es un número racional.}$$

Sea $\{E_\alpha\}_{\alpha \in I}$ el conjunto con las clases de equivalencia definidas por $\sim$. Para cada $\alpha \in I$ elijamos un sólo elemento $x_\alpha \in E_\alpha$ y consideremos el conjunto $E = \{x_\alpha\ :\ \alpha \in I\}$. Sean $r_1, r_2, \ldots, r_n, \ldots$ una numeración de los racionales en [-1,1] y

$$E_n = r_n + E \quad n = 1, 2, \ldots$$

De manera directa se prueba que $E_n \cap E_m = \emptyset$ si $n \neq m$ y $E_n \subset [-1, 2]$. Si $x \in [0, 1]$ entonces $x \in E_\alpha$ para algún $\alpha \in I$ y por lo tanto $x \sim x_\alpha$, $x_\alpha \in E$. Pero $-1 \leq x - x_\alpha \leq 1$ por lo que deducimos que $x \in E_n$ para algún n, así

$$[0, 1] \subset \bigcup_{n \geq 1} E_n \subset [-1, 2]. \tag{7.2.3}$$

Si E es medibe, entonces lo será cada E_n y $\mu_3(E) = \mu_3(E_n)$. Si $\mu_3(E_n) = \mu_3(E) > 0$, por las propiedades presentadas en la Nota 7.2.5 y las inclusiones de (7.2.3) se tendría que

$$+\infty = \sum_{n\geq 1} \mu_3(E_n) = \mu_3\left(\bigcup_{n\geq 1} E_n\right) \leq 3,$$

lo que no puede suceder. Por otro lado, si $\mu_3(E) = 0$, entonces $\mu_3(E_n) = 0$ para todo n y por lo tanto

$$1 = \mu_3([0,1]) \leq \sum_{n\geq 1} \mu_3(E_n) = 0,$$

lo que es imposible. Concluímos que el error fue suponer E medible.

Definición 7.2.6. Sea $A \subset \mathbb{R}^n$, diremos que A es medible si para todo $r > 0$, $A \cap B(0,r)$ es medible. En tal caso, su medida se define por

$$\mu(A) = \lim_{r\to+\infty} \mu_3(A \cap B(0,r)). \tag{7.2.4}$$

Esta medida es la llamada *medida de Lebesgue*. El límite definido por (7.2.4) puede que sea $+\infty$.

Nota 7.2.6. Claramente, si A es un conjunto medible acotado, F una figura y R un n-rectángulo, entonces $\mu_3(A) = \mu(A)$, $\mu_1(F) = \mu_3(F) = \mu(F)$ y $\mu_0(R) = \mu_3(R) = \mu(R)$.

Nota 7.2.7. Si $A_1 \subset \mathbb{R}^{n_1}$, $A_2 \subset \mathbb{R}^{n_2}$, $\ldots A_m \subset \mathbb{R}^{n_m}$, μ_{n_i} la medida de Lebesgue en $\mathbb{R}^{n_i}$ y μ la medida de Lebesgue en $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \ldots \times \mathbb{R}^{n_m}$, se calcula sin gran dificultad que $\mu(A_1 \times A_2 \times \ldots \times A_n) = \prod\limits_{i=1}^{m} \mu_{n_i}(A_i)$.

Teorema 7.2.2. *Sea, $\{A_k\}$ una familia de conjuntos medibles. Entonces:*

1. *A_1^c es medible;*
2. *$\bigcup\limits_{k\in\mathbb{N}} A_k$ es medible;*
3. *$\bigcap\limits_{k\in\mathbb{N}} A_k$ es medible.*

Demostración. Para simplificar la demostración supondremos A_1 acotado y $\bigcup\limits_{k\in\mathbb{N}} A_k$ cerrado y acotado.

1. Sea $\varepsilon > 0$, entonces existe un compacto K y un abierto θ tales que $K \subset A_1 \subset \theta$ y $\mu(\theta \setminus K) \leq \varepsilon$. Sea $r > 0$ y $B_\infty(0, r)$ la bola de centro 0 y radio r con la norma $\|\cdot\|_\infty$, es decir, $B_\infty(0, r) = [-r, r]^n$, entonces

$$\theta^c \cap B_\infty(0, r) \subset A_1^c \cap B_\infty(0, r) \subset K^c \cap]-(r+\delta), (r+\delta)[^n, \ \forall\, \delta > 0$$

y como

$$(K^c \cap]-(r+\delta), (r+\delta)[^n) \setminus (\theta^c \cap B_\infty(0,r)) \subset (\theta \setminus K) \cup (]-(r+\delta), (r+\delta)[^n \setminus B_\infty(0,r))$$

se obtiene

$$\mu((K^c \cap]-(r+\delta), (r+\delta)[^n) \setminus (\theta^c \cap B_\infty(0,r))) \leq \mu(\theta \setminus K) + \mu(]-(r+\delta), (r+\delta)[^n \setminus B_\infty(0,r)) =$$

$$\mu(\theta \setminus K) + (2r)^n - (2r+2\delta)^n \quad \forall\, \delta > 0,$$

debido a que $\mu(A \setminus B) = \mu(A \cup B) - \mu(B)$ (ver Nota 7.2.3 propiedad 3). Asi concluímos $\mu((K^c \cap]-(r+\delta), (r+\delta)[^n) \setminus (\theta^c \cap B_\infty(0,r))) \leq \varepsilon$ y por lo tanto $A_1^c \cap B_\infty(0, r)$ es medible para todo $r > 0$.

2. Sea $\varepsilon > 0$, entonces para todo $k \in \mathbb{N}$ existe un compacto M_k y un abierto θ_k tales que $M_k \subset A_k \subset \theta_k$ y $\mu(\theta_k \setminus M_k) \leq \varepsilon/2^k$. Debido a que

$$\bigcup_{k\in\mathbb{N}} \theta_k \setminus \overline{\bigcup_{k\in\mathbb{N}} M_k} \subset \bigcup_{k\in\mathbb{N}} \theta_k \setminus \bigcup_{k\in\mathbb{N}} M_k \subset \bigcup_{k\in\mathbb{N}} (\theta_k \setminus M_k),$$

concluímos $\mu(\bigcup_{k\in\mathbb{N}} \theta_k \setminus \overline{\bigcup_{k\in\mathbb{N}} M_k}) \leq \varepsilon$ y por lo tanto $\bigcup_{k\in\mathbb{N}} A_k$ es medible.

3. Se obtiene gracias a los dos resultados anteriores.

$\square$

A continuación presentamos un resultado cuya demostración escapa a los objetivos de este capítulo.

Teorema 7.2.3. *Sea $\{A_k\}$ una familia de conjuntos medibles, tales que $A_k \cap A_p = \emptyset$ si $k \neq p$. Entonces,*

$$\mu\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mu(A_k).$$

Proposición 7.2.1. *$\mu(\emptyset) = 0$ y $\mu(A + \vec{x}) = \mu(A)$ para todo conjunto medible $A \subset \mathbb{R}^n$ y $\vec{x} \in \mathbb{R}^n$.*

Demostración. Es directo (ver Nota 7.2.1). $\square$

Teorema 7.2.4. *Sea $\{A_k\}$ una familia de conjuntos medibles creciente, es decir, $A_1 \subset A_2 \subset \dots A_k \subset A_{k+1} \subset \dots$. Entonces*

$$\mu\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \lim_{k\to+\infty} \mu(A_k).$$

Demostración. Sea $B_1 = A_1$, $B_2 = A_2 \setminus A_1 \dots B_k = A_k \setminus A_{k-1}$. Observamos que

$$\bigcup_{k\in\mathbb{N}} B_k = \bigcup_{k\in\mathbb{N}} A_k.$$

Como los conjuntos B_k son disjuntos, utilizando el Teorema 7.2.3 se obtiene

$$\mu\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \mu\left(\bigcup_{k\in\mathbb{N}} B_k\right) = \sum_{k\in\mathbb{N}} \mu(B_k)$$

$$= \lim_{N\to+\infty} \sum_{k=1}^{N} \mu(B_k) = \lim_{N\to+\infty} \mu(A_1) + \sum_{k=2}^{N} \mu(A_k) - \mu(A_{k-1}) = \lim_{N\to+\infty} \mu(A_N).$$

$\square$

Teorema 7.2.5. *Sea $\{A_k\}$ una familia de conjuntos medibles decreciente, es decir, $A_1 \supset A_2 \supset \dots A_k \supset A_{k+1} \supset \dots$. Entonces*

$$\mu\left(\bigcap_{k\in\mathbb{N}} A_k\right) = \lim_{k\to+\infty} \mu(A_k).$$

Demostración. Definamos $B_1 = \emptyset$, $B_2 = A_1 \setminus A_2 \dots, A_k = A_1 \setminus A_k$ que resultará ser una familia creciente de conjuntos medibles. Utilizando el teorema anterior, se obtiene que

$$\mu\left(\bigcup_{k\in\mathbb{N}} B_k\right) = \lim_{k\to+\infty} \mu(B_k) = \mu(A_k) - \lim_{k\to+\infty} \mu(A_1).$$

Por otro lado

$$\bigcup_{k\in\mathbb{N}} B_k = A_1 \setminus \bigcap_{k\in\mathbb{N}} A_k,$$

concluyendo así el resultado. $\square$

7.3 Integral de funciones positivas

Definición 7.3.1. Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ una función positiva, denotamos

$$[0, f] = \{(\vec{x}, y) \in \mathbb{R}^n \times \mathbb{R}_+ \;:\; 0 \leq y \leq f(\vec{x})\}.$$

Diremos que la función f es medible, si el conjunto $[0, f]$ es medible en $\mathbb{R}^n \times \mathbb{R}$. Si además, $\mu([0, f]) < +\infty$ se dirá que f es integrable y su integral en $\mathbb{R}^n$ estará definida por

$$\int f := \mu([0, f]). \tag{7.3.1}$$

Notaremos por $\mathcal{M}_+(\mathbb{R}^n)$ al conjunto de todas las funciones positivas medibles.

Ejemplo 7.3.1. Sea $A \subset \mathbb{R}^n$ un conjunto medible y $\beta \geq 0$. Definamos la función $f : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ por

$$f(\vec{x}) = \begin{cases} \beta & \text{si } \vec{x} \in A \\ 0 & \text{so } \vec{x} \notin A. \end{cases}$$

Entonces, debido a que $\mu(A^c \times \{0\}) = 0$ se tendrá

$$\int f = \mu([0, f]) = \mu(A \times [0, \beta] \cup A^c \times \{0\}) = \beta\mu(A)$$

Teorema 7.3.1. *Sea $\{f_k\}$ una sucesión de funciones positivas integrables tal que $\{f_k(\vec{x})\}$ es una sucesión decreciente y f el límite puntual, es decir, $f(\vec{x}) = \lim_{k\to+\infty} f_k(\vec{x})$. Entonces, f es integrable y*

$$\int f = \lim_{k\to+\infty} \int f_k.$$

Demostración. Se observa que

$$(\vec{x}, y) \in [0, f] \Leftrightarrow 0 \leq y \leq f_k(\vec{x}) \;\forall\, k \in \mathbb{N} \Leftrightarrow (\vec{x}, y) \in [0, f_k] \;\forall\, k \in \mathbb{N} \Leftrightarrow (\vec{x}, y) \in \bigcap_{k\in\mathbb{N}} [0, f_k],$$

por lo tanto, utilizando el Teorema 7.2.5, obtenemos

$$\mu([0, f]) = \lim_{k\to+\infty} \mu([0, f_k]) < +\infty.$$

□

Teorema 7.3.2. *Sea $\{f_k\}$ una sucesión creciente de funciones positivas medibles e integrables convergiendo puntuamente a la función positiva, medible e integrable f. Entonces,*

$$\int f = \lim_{k\to+\infty} \int f_k.$$

Demostración. Claramente $[0, f] = \bigcup_{k\geq 1} [0, f_k]$, por lo que utilizando el Teorema 7.2.4 se obtiene el resultado. □

7.4 Funciones simples

Definición 7.4.1. Una función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ se dice una *función simple*, si existe un número finito de conjuntos medibles disjuntos en $\mathbb{R}^n$, $A_1, A_2, \ldots, A_m$, y la misma cantidad de números reales $y_1, y_2, \ldots, y_m$ tales que

$$f(\vec{x}) = \begin{cases} y_i & \text{si } \vec{x} \in A_i \\ 0 & \text{si } \vec{x} \notin \bigcup_{i=1}^{m} A_i. \end{cases} \tag{7.4.1}$$

Si para un conjunto $A \subset \mathbb{R}^n$, escribimos su función indicatriz

$$\Psi(\vec{x}) = \begin{cases} 1 & \text{si } \vec{x} \in A \\ 0 & \text{si } \vec{x} \notin A, \end{cases} \tag{7.4.2}$$

la definición anterior se reduce a poder escribir la función f de la forma

$$f(\vec{x}) = \sum_{i=1}^{m} y_i \Psi_{A_i}(\vec{x}).$$

Si para todo $i = 1, \ldots, m$ se tiene $\mu(A_i) < +\infty$, diremos que f es integrable y su integral estará definida por

$$\int f := \sum_{i=1}^{m} y_i \mu(A_i).$$

Teorema 7.4.1. *Sean f y g dos funciones simples e integrables, entonces para todo $\lambda \in \mathbb{R}$ se tiene que $f + \lambda g$ es una función simple e integrable y*

$$\int (f + \lambda g) = \int f + \lambda \int g.$$

Demostración. Es directo. □