ANÁLISIS FUNCIONAL
Y
OPTIMIZACIÓN

Apuntes

Curso: 2010/11

José D. Martín Gómez

Índice general

Capítulo 1

ELEMENTOS DE ANÁLISIS FUNCIONAL

"La distancia más corta entre personas es una sonrisa"
E. Zeidler

1.1. Recordatorio de Análisis Funcional

En esta sección recordamos algunos resultados que se usarán a menudo a lo largo del curso, la mayoría de los cuales corresponden a la asignatura Análisis Funcional.

A lo largo del curso, los espacios vectoriales se considerarán siempre construidos sobre el cuerpo de los números reales $\mathbb{R}$.

Definición 1.1.1 *Sea X un espacio vectorial, una norma en X es una aplicación $\|\cdot\| : X \to \mathbb{R}^+$ que verifica:*

1. $\|x\| = 0$ *si y sólo si* $x = 0$,
2. $\|\lambda x\| = |\lambda|\|x\|, \forall\, \lambda \in \mathbb{R}, \forall\, x \in X$,
3. $\|x + y\| \leq \|x\| + \|y\|$, $\forall\, x, y \in X$.

Un espacio dotado de una norma, se llama espacio normado (en adelante e.n.). Los espacios normados (en adelante también e.n.) son espacios métricos con la distancia

$$d(x, y) = \|x - y\|, \quad \forall\, x, y \in X.$$

Si el espacio resultante es completo (i.e. si toda sucesión de A. Cauchy es convergente) se dice que es un espacio de S. Banach.

Proposición 1.1.2 *Un subespacio de un espacio de Banach sigue siendo de Banach con la norma inducida si y sólo si es cerrado.*

Proposición 1.1.3 *Dado un e.n. X, la aplicación norma $\|\cdot\| : X \to \mathbb{R}$ es continua, de hecho verifica*

$$|\|x\| - \|y\|| \leq \|x - y\|, \quad \forall\, x, y \in X.$$

También son continuas las aplicaciones $(x, y) \in X \times X \to x + y \in X$,$(\lambda, x) \in \mathbb{R} \times X \to \lambda x \in X$.

Observación 1.1.4 *A veces en un mismo espacio X podemos tener definidas dos normas $\|\cdot\|_a$, $\|\cdot\|_b$ distintas. En general es falso que la aplicación $\|\cdot\|_b$ sea continua considerando el espacio X con la topología definida a partir de $\|\cdot\|_a$.*

Definición 1.1.5 *Dos normas definidas en un mismo espacio X, se dicen normas equivalentes si generan la misma topología.*

Proposición 1.1.6 *Dos normas $\|\cdot\|_a, \|\cdot\|_b$ son equivalentes si y sólo si existen $C_1, C_2 > 0$ tales que*

$$C_1\|x\|_a \leq \|x\|_b \leq C_2\|x\|_a, \quad \forall\, x \in X.$$

Ejemplo 1.1.7 *Si X, Y son dos e.n., el espacio producto $X \times Y$, es un e.n. con cualquiera de las normas equivalentes*

$$\begin{cases} \|(x,y)\|_\infty = \text{máx}\{\|x\|_X, \|y\|_Y\}, \quad \forall\, x \in X, \forall\, y \in Y, \\ \|(x,y)\|_1 = \|x\|_X + \|y\|_Y, \quad \forall\, x \in X, \forall\, y \in Y, \\ \|(x,y)\|_2 = \left(\|x\|_X^2 + \|y\|_Y^2\right)^{1/2}, \quad \forall\, x \in X, \forall\, y \in Y. \\ \|(x,y)\|_p = \left(\|x\|_X^p + \|y\|_Y^p\right)^{1/p}, \, p \in (1,+\infty), \quad \forall\, x \in X, \forall\, y \in Y. \end{cases}$$

Definición 1.1.8 *Sea X un espacio vectorial, una aplicación $(\cdot|\cdot) : X \times X \to \mathbb{R}$, se dice que es un producto escalar si es bilineal, simétrica y definida positiva, es decir*

1. $(\lambda x + \mu y|z) = \lambda(x|z) + \mu(y|z) \quad \forall\lambda, \mu \in \mathbb{R}, \quad \forall x, y, z \in X,$

2. $(x|y) = (y|x), \quad \forall x, y \in X,$

3. $(x|x) > 0, \quad \forall x \in X \setminus \{0\}.$

Dado un producto escalar en X, se puede definir una norma sobre X (llamada norma inducida por el producto escalar) mediante

$$\|x\| = \sqrt{(x|x)}, \quad \forall\, x \in X.$$

Un e.n. en el que la norma se define a través de un producto escalar, se llama prehilbertiano. En los espacios prehilbertianos se verifica la desigualdad de A. Cauchy-H. Schwarz

$$|(x|y)| \leq \|x\|\|y\|, \quad \forall\, x, y \in X.$$

Si el espacio es completo, se dice que es un espacio de D. Hilbert.

Se recuerda

Definición 1.1.9 *Dado X un espacio vectorial, se dice que $C \subset X$ es convexo si para todo $x, y \in C$ se tiene que el segmento*

$$[x,y] = \big\{\lambda x + (1-\lambda)y \,/\, \lambda \in [0,1]\big\}$$

está contenido en C.

En los espacios de Hilbert se tiene el siguiente resultado fundamental:

Teorema 1.1.10 *(Teorema de la proyección) Sea H un espacio de Hilbert, K un conjunto convexo cerrado no vacío y $x \in H$, entonces existe un único elemento $\hat{x} \in K$ (llamado la proyección de x sobre K) tal que*

$$\|x - \hat{x}\| = \text{ínf}\{\|x - k\| \,/\, k \in K\}.$$

Además $\hat{x}$ está caracterizado por

$$\hat{x} \in K \quad y \quad (x - \hat{x}|k - \hat{x}) \leq 0, \quad \forall k \in K. \tag{1.1}$$

1.2. Ejemplos de espacios normados.

1.2.1. El espacio $\mathbb{R}^N$

El espacio vectorial $\mathbb{R}^N$, formado por las N-tuplas $x = (x_1, \cdots, x_N)$ de números reales, es un espacio de Hilbert con el producto escalar usual:

$$(x|y) = \sum_{i=1}^{N} x_i y_i, \forall\, x, y \in \mathbb{R}^N.$$

La norma correspondiente es la norma euclídea. También se pueden definir otras normas como por ejemplo

$$\begin{cases} \|x\|_\infty = \text{máx}\{|x_1|, \cdots, |x_N|\}, \forall\, x \in \mathbb{R}^N, \\ \|x\|_1 = \sum_{i=1}^{N} |x_i|, \forall\, x \in \mathbb{R}^N, \\ \|x\|_p = (\sum_{i=1}^{N} |x_i|^p)^{1/p},\ p \in (1, +\infty), \forall\, x \in \mathbb{R}^N. \end{cases}$$

En $\mathbb{R}^N$ y en general en todo espacio de dimensión finita, todas las normas son equivalentes.

1.2.2. Los espacios ℓ^p

Sobre el espacio vectorial de las sucesiones de números reales y para $1 \leq p < +\infty$, se define ℓ^p por

$$\ell^p = \big\{x = \{x_n\}_{n\in\mathbb{N}} \,/\, x_n \in \mathbb{R},\ \forall\, n \in \mathbb{N},\ \sum_{n=1}^{\infty} |x_n|^p < +\infty\big\}.$$

Los espacios ℓ^p son espacios de Banach dotados de la norma

$$\|x\|_p = \|x\|_{\ell^p} \overset{\text{def.}}{=} \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}}, \quad \forall\, x = \{x_n\} \in \ell^p.$$

Para $p = 2$, el espacio ℓ^2 es un espacio de Hilbert con el producto escalar

$$(x|y) = \sum_{n=1}^{\infty} x_n y_n, \quad \forall\, x = \{x_n\},\ y = \{y_n\} \in \ell^2.$$

Se define ℓ^∞ por

$$\ell^\infty = \big\{x = \{x_n\}_{n\in\mathbb{N}} \,/\, x_n \in \mathbb{R},\ \forall\, n \in \mathbb{N},\ \{x_n\} \text{ está acotada}\big\}.$$

El espacio ℓ^∞ es un espacio de Banach dotado de la norma

$$\|x\|_\infty = \|x\|_{\ell^\infty} \overset{\text{def.}}{=} \sup_{n\in\mathbb{N}} |x_n|, \quad \forall\, x = \{x_n\} \in \ell^\infty.$$

El espacio c_0 se define como el subespacio cerrado de ℓ^∞ definido por las sucesiones que convergen a cero y por tanto es también un espacio de Banach.

En los espacios ℓ^p se verifica la desigualdad de O. Hölder: Sean $p, p' \in [1, +\infty]$ dos exponentes conjugados, i.e. tales que

$$\frac{1}{p} + \frac{1}{p'} = 1 \tag{1.2}$$

(con el convenio que $p = 1 \Longrightarrow p' = +\infty$ y $p = +\infty \Longrightarrow p' = 1$). Si $\{x_n\} \in \ell^p$, $\{y_n\} \in \ell^{p'}$ entonces $\{x_n y_n\} \in \ell^1$ y se tiene

$$\|\{x_n y_n\}\|_1 \leq \|\{x_n\}\|_p \|\{y_n\}\|_{p'}.$$

Cuando $1 < p, p' < +\infty$, la desigualdad anterior se escribe

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} \left(\sum_{n=1}^{\infty} |y_n|^{p'} \right)^{1/p'}.$$

Si $p = 1$, $p' = +\infty$ se tiene

$$\sum_{n=1}^{\infty} |x_n y_n| \leq (\sup_{n \in \mathbb{N}} |y_n|) \sum_{n=1}^{\infty} |x_n|.$$

Entre los espacios ℓ^p, se tiene la siguiente relación de inclusión

Proposición 1.2.1 *Para $1 \leq p \leq q \leq +\infty$ se tiene $\ell^p \subset \ell^q$ y además:*

$$\|x\|_q \leq \|x\|_p, \quad \forall x \in \ell^p. \tag{1.3}$$

Demostración: Primero observamos que para todo $n \in \mathbb{N}$, se tiene

$$|x_n| = (|x_n|^p)^{1/p} \leq \left(\sum_{m=1}^{\infty} |x_m|^p \right)^{1/p} = \|x\|_p. \tag{1.4}$$

En el caso $q = +\infty$ la desigualdad anterior implica

$$\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n| \leq \|x\|_p,$$

lo que prueba (1.3) para el caso $q = +\infty$.

Supongamos ahora $p \leq q < +\infty$. Si $x = 0$ el resultado es evidente, en otro caso la desigualdad (1.4) prueba que

$$\frac{|x_n|}{\|x\|_p} \leq 1, \quad \forall\, n \in \mathbb{N}$$

y por tanto

$$\left(\frac{|x_n|}{\|x\|_p} \right)^q \leq \left(\frac{|x_n|}{\|x\|_p} \right)^p, \quad \forall\, n \in \mathbb{N}.$$

Sumando la desigualdad anterior para todo $n \in \mathbb{N}$, se deduce entonces (1.3). $\square$

1.2.3. Los espacios $L^p(\Omega)$

Sea $\Omega \subset \mathbb{R}^N$ un conjunto medible en el sentido de H. Lebesgue. Para una función f, integrable en el sentido de Lebesgue, se denota la integral de f por alguna de las siguientes expresiones:

$$\int_\Omega f = \int_\Omega f\,dx = \int_\Omega f(x)dx,$$

y la medida de Lebesgue de un conjunto medible Ω la denotaremos por $|\Omega|$. Para $1 \le p < +\infty$, se define

$$L^p(\Omega) = \left\{ f : \Omega \to \mathbb{R} \ \text{ medible} \,/ \ \int_\Omega |f|^p < +\infty \right\}.$$

Las funciones de $L^p(\Omega)$ se entienden definidas salvo conjuntos de medida nula, es decir, dos funciones que son iguales salvo en un conjunto de medida nula son consideradas como la misma función, luego en realidad no son funciones sino clases de funciones. Los espacios $L^p(\Omega)$ son espacios de Banach con las normas

$$\|f\|_p = \|f\|_{L^p} \overset{\text{def.}}{=} \left(\int_\Omega |f|^p \right)^{1/p}, \ \forall\, f \in L^p(\Omega).$$

Para $p = 2$, el espacio $L^2(\Omega)$ es un espacio de Hilbert con el producto escalar

$$(f|g) = \int_\Omega fg\,, \quad \forall\, f,\ g \in L^2(\Omega).$$

Se define $L^\infty(\Omega)$ como el espacio formado por las funciones $f : \Omega \to \mathbb{R}$ medibles tales que

$$\text{supes}_\Omega |f| < +\infty,$$

donde "supes" es el supremo esencial de $|f|$ en Ω, el cual está definido por

$$\text{supes}_\Omega |f| = \text{ínf}\{\sup\{|f(x)| \,/\, x \in \Omega \setminus N\} \,/\, N \subset \Omega \text{ medible, de medida nula}\}.$$

Obsérvese que este valor no depende del representante de la función f escogido. El espacio $L^\infty(\Omega)$ es un espacio de Banach con la norma

$$\|f\|_\infty = \|f\|_{L^\infty} \overset{\text{def.}}{=} \text{supes}_\Omega |f|.$$

En $L^p(\Omega)$ se verifica la desigualdad de Hölder: Sean $f \in L^p(\Omega)$, $g \in L^{p'}(\Omega)$, con $1 \le p, p' < +\infty$, verificando (1.2), entonces fg pertenece a $L^1(\Omega)$ y se verifica

$$\|fg\|_1 \le \|f\|_p \|g\|_{p'},$$

lo que en el caso $1 < p, p' < +\infty$, se escribe

$$\int_\Omega |fg| \le \left(\int_\Omega |f|^p \right)^{1/p} \left(\int_\Omega |g|^{p'} \right)^{1/p'},$$

mientras que en el caso $p = 1$, $p' = +\infty$ se tiene

$$\int_\Omega |fg| \le (\text{supes}_\Omega |g|) \int_\Omega |f|\,.$$

Cuando Ω es de medida finita, en los espacios $L^p(\Omega)$, se tiene la siguiente relación de inclusión (obsérvese que las relaciones van al contrario que en el caso ℓ^p).

Proposición 1.2.2 *Supongamos $|\Omega| < +\infty$, si $1 \leq q \leq p \leq \infty$ entonces $L^p(\Omega) \subset L^q(\Omega)$ y se tiene*

$$\|f\|_q \leq |\Omega|^{(p-q)/pq} \|f\|_p, \, \forall f \in L^p(\Omega). \tag{1.5}$$

Demostración: Supongamos primero $p < +\infty$. En este caso, basta usar la desigualdad de Hölder con los valores p/q, $p/(p-q)$ (que son mayores que 1) de la siguiente forma

$$\|f\|_q^q = \int_\Omega |f|^q \, 1 \leq \left(\int_\Omega |f|^p \right)^{q/p} \left(\int_\Omega 1^{p/(p-q)} \right)^{(p-q)/p} = \|f\|_p^q |\Omega|^{(p-q)/p},$$

lo que extrayendo raíz q-ésima prueba (1.5).

Cuando $p = +\infty$, se tiene análogamente

$$\|f\|_q^q = \int_\Omega |f|^q \leq \int_\Omega (\text{supes}\, |f|)^q = |\Omega| \|f\|_\infty^q,$$

con lo que de nuevo se obtiene el resultado extrayendo raíz q-ésima. □

1.2.4. Los espacios $C^k([a,b])$

Sea $a, b \in \mathbb{R}, a < b$ e $I = [a,b]$, para $k \in \mathbb{Z}$, $k \geq 0$ se define el espacio $C^k(I)$ como el espacio formado por las restricciones a I de las funciones de clase k en $\mathbb{R}$ (i.e. tales que son derivables hasta el orden k con derivadas continuas, y en el caso $k = 0$ son simplemente las funciones continuas en I). El espacio $C^k(I)$ es un espacio de Banach con la norma

$$\|f\|_{C^k(I)} = \sum_{i=0}^{k} \max_{t \in I} |f^{i)}(t)|.$$

1.3. Aplicaciones lineales y continuas. Espacio dual.

Proposición 1.3.1 *Sean X, Y dos e.n., una aplicación lineal $A : X \to Y$ es continua si y solo si existe $C \geq 0$ tal que*

$$\|Ax\|_Y \leq C \|x\|_X, \quad \forall \, x \in X.$$

Ejemplo 1.3.2 *Por la Proposición 1.2.1, para $1 \leq p \leq q \leq +\infty$, podemos considerar la aplicación (inyección) $i : \ell^p \to \ell^q$ por $i(x) = x$, para todo $x \in \ell^p$. Claramente esta aplicación es lineal y por la desigualdad (1.3) se tiene*

$$\|i(x)\|_q = \|x\|_q \leq \|x\|_p.$$

Por tanto i es continua. En particular esto significa que si una sucesión converge en ℓ^p, también converge en ℓ^q. Análogamente, por la Proposición 1.2.2, si $|\Omega| < +\infty$, entonces para $1 \leq q \leq p \leq +\infty$ la aplicación $i : L^p(\Omega) \to L^q(\Omega)$ definida por $i(u) = u$, para toda $u \in L^p(\Omega)$ es continua.

Definición 1.3.3 *El espacio vectorial formado por las aplicaciones lineales y continuas de X en Y se denota por $\mathcal{L}(X,Y)$ y es un e.n. con la norma*

$$\|A\|_{\mathcal{L}(X,Y)} = \inf\{C \geq 0 \, / \, \|Ax\|_Y \leq C \|x\|_X, \quad \forall \, x \in X\} =$$

$$= \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|=1} \|Ax\|_Y = \sup_{\|x\| \leq 1} \|Ax\|_Y, \quad \forall A \in \mathcal{L}(X,Y).$$

Obsérvese que se verifica

$$\|Ax\|_Y \leq \|A\|_{\mathcal{L}(X,Y)}\|x\|_X, \quad \forall\, x \in X.$$

Se tiene

Proposición 1.3.4 *Si Y es un espacio de Banach, entonces $\mathcal{L}(X,Y)$ es un espacio de Banach.*

Definición 1.3.5 *Dado un e.n. X, se denomina espacio dual de X y se denota por X' al espacio $\mathcal{L}(X,\mathbb{R})$.*

Nótese que, por la Proposición 1.3.4, X' siempre es un Banach (aunque X no lo sea). Si $x' \in X'$ y $x \in X$ se suele denotar

$$\langle x', x\rangle = x'(x)$$

y se denomina producto de dualidad de x' con x.

Cuando sea importante especificar los espacios, usaremos la notación $\langle x', x\rangle_{X',X}$.

1.4. Identificación de algunos espacios duales

1.4.1. Espacios de Hilbert

En los espacios de Hilbert X se verifica el Teorema de F. Riesz que permite identificar X con X'. Concretamente, se tiene que la aplicación $R : X \to X'$ definida por

$$\langle Rx, y\rangle = (x|y), \quad \forall\, x, y \in X,$$

es un isomorfismo isómetrico entre X y X', es decir es lineal, biyectiva y verifica

$$\|Rx\|_{X'} = \|x\|_X, \quad \forall\, x \in X,$$

en particular, tanto R como R^{-1} son continuas. Este resultado permite identificar X con X'.

1.4.2. Espacios ℓ^p

Si $1 \leq p < +\infty$, el dual del espacio ℓ^p se identifica con $\ell^{p'}$ mediante la aplicación (es un isomorfismo isométrico) $R : \ell^{p'} \to (\ell^p)'$ definida por

$$\langle Rx, y\rangle = \sum_{n=1}^{\infty} x_n y_n, \quad \forall\, x = \{x_n\} \in \ell^{p'},\ \forall\, y = \{y_n\} \in \ell^p. \tag{1.6}$$

En el caso $p = +\infty$, la aplicación anterior es lineal e isométrica (por tanto inyectiva) pero no sobreyectiva. Si en cambio definimos $R : \ell^1 \to (c_0)'$ de forma análoga, i.e.

$$\langle Rx, y\rangle = \sum_{n=1}^{\infty} x_n y_n, \quad \forall\, x = \{x_n\} \in \ell^1,\ \forall\, y = \{y_n\} \in (c_0)',$$

entonces sí resulta un isomorfismo isométrico, i.e. ℓ^1 se identifica con $(c_0)'$.

1.4.3. Espacios $L^p(\Omega)$

Si $1 \leq p < +\infty$ y $\Omega \subset \mathbb{R}^N$ es medible entonces el dual del espacio $L^p(\Omega)$ se identifica con $L^{p'}(\Omega)$ mediante la aplicación (es un isomorfismo isométrico) $R : L^{p'}(\Omega) \to (L^p(\Omega))'$ definida por

$$\langle Rf, g\rangle = \int_\Omega f g\, dx, \quad \forall\, f \in L^{p'}(\Omega),\ \forall\, g \in L^p(\Omega).$$

Si $p = +\infty$, la aplicación anterior es lineal e isométrica (por tanto inyectiva) pero no sobreyectiva. El espacio $L^1(\Omega)$ no se identifica con el dual de ningún espacio.

1.5. Algunos teoremas importantes.

Recordamos algunos resultados fundamentales referentes a la teoría de aplicaciones lineales y continuas entre e.n..

Teorema 1.5.1 *(H. Hahn-S. Banach) Sean X un e.n., M un subespacio vectorial de X y $f \in M'$, entonces existe $\tilde{f} \in X'$ tal que*

$$\tilde{f}(x) = f(x), \quad \forall\, x \in M \tag{1.7}$$

$$\|\tilde{f}\|_{X'} = \|f\|_{M'}. \tag{1.8}$$

Observación 1.5.2 *Sean X e.n., M subespacio vectorial de X, $f \in M'$, $\tilde{f} \in X'$ tales que se verifica (1.7), entonces se tiene*

$$\|\tilde{f}\|_{X'} = \sup_{\substack{x\in X\\ \|x\|_X=1}} |\tilde{f}(x)| \geq \sup_{\substack{x\in M\\ \|x\|_X=1}} |\tilde{f}(x)| = \sup_{\substack{x\in M\\ \|x\|_X=1}} |f(x)| = \|f\|_{M'},$$

luego el Teorema de Hahn-Banach, además de decirnos que toda aplicación lineal y continua de M en $\mathbb{R}$, se puede prolongar a una aplicación lineal y continua de X en $\mathbb{R}$, nos dice que esta prolongación la podemos tomar de forma que tenga la menor norma posible.

Definición 1.5.3 *Sean X un e.n., M un subespacio de X y N un subespacio de X', se define el complemento ortogonal de M por*

$$M^\perp = \{x' \in X' \,/\ \langle x', x\rangle = 0, \quad \forall\, x \in M\}$$

y el complemento ortogonal de N por

$${}^\perp N = \{x \in X \,/\ \langle x', x\rangle = 0, \quad \forall\, x' \in N\}.$$

Los conjuntos $M^\perp$ y ${}^\perp N$ son subespacios cerrados de X' y X respectivamente. El Teorema de Hahn-Banach tiene las siguientes consecuencias importantes

Corolario 1.5.4 *Sea X un e.n.*

1. *Para todo $x \in X$ existe $x' \in X'$ tal que $\|x'\|_{X'} = 1$ y x' está alineado con x, i.e.*

$$\langle x', x\rangle = \|x'\|_{X'}\|x\|_X. \tag{1.9}$$

2. Esto implica

$$\|x\|_X = \max_{\|x'\|_{X'}=1} \langle x', x\rangle = \max_{x'\neq 0} \frac{\langle x', x\rangle}{\|x'\|_{X'}}, \quad \forall\, x \in X. \tag{1.10}$$

3. Si M es un subespacio de X, y $x \in X \setminus \bar{M}$, entonces existe $x' \in M^{\perp}$ tal que $\langle x', x\rangle \neq 0$.

4. Como consecuencia se deduce

$$\bar{M} = {}^{\perp}(M^{\perp}), \ \text{para todo subespacio vectorial } M \subset X. \tag{1.11}$$

5. En particular, un subespacio $M \subset X$ es denso si y solo si $M^{\perp} = \{0\}$.

Teorema 1.5.5 *(Teorema de separación estricta de convexos) Sean C y K dos conjuntos convexos cerrados no vacíos de un e.n. X y, por ejemplo, K es compacto. Si $C \cap K = \emptyset$ entonces*

$$\exists x' \in X' \ \text{tal que} \quad \sup_{x\in C}\langle x', x\rangle < \inf_{x\in K}\langle x', x\rangle \tag{1.12}$$

Corolario 1.5.6 *Sean C un conjunto convexo cerrado no vacío de un e.n. X y $x_0 \notin C$ entonces*

$$\exists x' \in X' \ \text{tal que} \quad \sup_{x\in C}\langle x', x\rangle < \langle x', x_0\rangle \tag{1.13}$$

Teorema 1.5.7 *(Teorema de la aplicación abierta de Banach) Sean X, Y dos espacios de Banach y $A \in \mathcal{L}(X, Y)$, sobreyectiva. Entonces A transforma subconjuntos abiertos de X en subconjuntos abiertos de Y.*

Teorema 1.5.8 *(Teorema del inverso a la derecha de Banach) Bajo las mismas hipótesis que el Teorema 1.5.7 anterior, existe $C \geq 0$ tal que para todo $y \in Y$, existe $x \in X$ con $Ax = y$, $\|x\|_X \leq C\|y\|_Y$.*

Teorema 1.5.9 *(Teorema del inverso de Banach) Bajo las mismas hipótesis que el Teorema anterior si además A es inyectiva (y por tanto biyectiva), la inversa de A es lineal y continua.*

Teorema 1.5.10 *(Teorema de S. Banach-H. Steinhaus) Sean X un espacio de Banach, Y un e.n. y $\{A_i\}$ una familia de aplicaciones lineales y continuas de X en Y tales que para todo $x \in X$, $\sup_{i\in I} \|A_i x\|_Y$ es finito. Entonces*

$$\sup_{i\in I} \|A_i\|_{\mathcal{L}(X,Y)} < +\infty.$$

1.6. Espacios reflexivos

Terminado el repaso de Análisis Funcional, comenzamos ahora con la materia propia del curso introduciendo los espacios reflexivos.

Dado un e.n. X usaremos la notación $X'' = (X')'$, $X''' = (X'')'$, etc.

Definición 1.6.1 *Dado X e.n., se define $J : X \to X''$ por*

$$\langle J(x), x'\rangle_{X'',X'} = \langle x', x\rangle_{X',X}, \quad \forall\, x \in X, \ \forall\, x' \in X', \tag{1.14}$$

y se le llama inyección canónica de X en X''.

Proposición 1.6.2 *La aplicación J está bien definida, es lineal e isométrica. Por tanto permite identificar X con el subespacio $J(X)$ de X''.*

Demostración: Para ver que está bien definida hay que probar que dado $x \in X$, $J(x)$ pertenece a X'', o lo que es lo mismo, que la aplicación

$$x' \in X' \to \langle x', x \rangle_{X',X}$$

es lineal y continua. La linealidad se prueba fácilmente, respecto a la continuidad basta usar

$$|\langle x', x \rangle_{X',X}| \leq \|x\|_X \|x'\|_{X'}, \quad \forall\, x' \in X',$$

como $J(x)$ es lineal (y por tanto, si no hay confusión, pondremos Jx), esta desigualdad prueba que Jx es continua y $\|Jx\|_{X''} \leq \|x\|_X$. Para ver que de hecho esta desigualdad es una igualdad, usando el Corolario 1.5.4 consideremos ahora $\tilde{x}' \in X'$ tal que $\|\hat{x}'\|_{X'} = 1$ y $\hat{x}'$ alineado con x. Se tiene

$$\|Jx\|_{X''} = \sup_{\|x'\|_{X'}=1} \langle Jx, x' \rangle_{X'',X'} \geq \langle Jx, \hat{x}' \rangle_{X'',X'} = \langle \hat{x}', x \rangle_{X',X'} = \|x\|_X.$$

Para terminar la demostración basta probar que J es lineal lo cual se prueba fácilmente. $\square$

Definición 1.6.3 *Un e.n. X se dice reflexivo, si la aplicación J, definida por (1.14), es sobreyectiva.*

Observación 1.6.4 *Es claro a partir de la definición anterior que todo espacio reflexivo se identifica con su bidual, sin embargo esta condición no basta para que el espacio sea reflexivo. Puede ocurrir que haya una identificación de X con X'' distinta de J y que J no sea sobreyectiva.*

Como el bidual siempre es un espacio de Banach se deduce en particular que todo espacio reflexivo es Banach.

1.7. Ejemplos de espacios reflexivos

1.7.1. Espacios de Hilbert

Vamos a probar que si X es Hilbert entonces es reflexivo. Para ello hay que probar que la aplicación J definida por (1.14) es sobreyectiva. Sea entonces $x_0'' \in X''$ y consideremos la aplicación de Riesz $R : X \to X'$ definida por

$$\langle Rx, y \rangle = (x|y), \quad \forall\, x, y \in X,$$

la cual sabemos que es un isomorfismo isométrico. En particular, la composición $x_0'' \circ R$ es lineal y continua y por tanto un elemento de X', luego existe $x_0 \in X$ tal que $Rx_0 = x_0'' \circ R$. Vamos a probar que $Jx_0 = x_0''$ para ello consideramos $x' \in X'$ y $x \in X$ tal que $Rx = x'$. Se tiene

$$\langle Jx_0, x' \rangle_{X'',X'} = \langle x', x_0 \rangle_{X',X} = \langle Rx, x_0 \rangle_{X',X} =$$

$$= (x|x_0) = \langle Rx_0, x \rangle_{X',X} = \langle x_0'', Rx \rangle_{X'',X'} = \langle x_0'', x' \rangle_{X'',X'}.$$

Como la igualdad anterior es cierta para todo $x' \in X'$, se deduce $Jx_0 = x_0''$.

1.7.2. Espacios ℓ^p

Se cumple que los espacios ℓ^p con $1 < p < +\infty$ son reflexivos. Para probarlo, sea $X = \ell^p$ con $1 < p < +\infty$ y definamos J por (1.14). Como antes hay que ver que J es sobreyectiva. La demostración es muy similar a la anterior, sea $x'' \in (\ell^p)''$ y consideremos la aplicaciones $R_p : \ell^p \to (\ell^{p'})'$, $R_{p'} : \ell^{p'} \to (\ell^p)'$ definidas por

$$\langle R_p x, y\rangle_{(\ell^{p'})',\ell^{p'}} = \sum_{n=1}^{\infty} x_n y_n, \quad \forall\, x = \{x_n\} \in \ell^p,\ \forall\, y = \{y_n\} \in \ell^{p'}$$

$$\langle R_{p'} y, x\rangle_{(\ell^{p})',\ell^{p}} = \sum_{n=1}^{\infty} x_n y_n, \quad \forall\, x = \{x_n\} \in \ell^p,\ \forall\, y = \{y_n\} \in \ell^{p'},$$

las cuales sabemos que son isomorfismos isométricos. Sabemos que la composición $x'' \circ R_{p'}$ es una aplicación lineal y continua de $\ell^{p'}$ en $\mathbb{R}$ y por tanto un elemento de $(\ell^{p'})'$, luego existe $x = \{x_n\} \in \ell^p$ tal que $R_p x = x'' \circ R_{p'}$. Dado $y' \in (\ell^p)'$ e $y \in \ell^{p'}$ tal que $R_{p'} y = y'$ se tiene

$$\langle Jx, y'\rangle_{(\ell^p)'',(\ell^p)'} = \langle y', x\rangle_{(\ell^p)',(\ell^p)} = \langle R_{p'} y, x\rangle_{(\ell^p)',(\ell^p)}$$
$$= \sum_{n=1}^{\infty} x_n y_n = \langle R_p x, y\rangle_{(\ell^{p'})',\ell^{p'}} = \langle x'', R_{p'} y\rangle_{X'',X'} = \langle x'', y'\rangle_{(\ell^p)'',(\ell^p)'}.$$

Como esta igualdad es cierta para todo $y' \in (\ell^p)'$ se deduce $Jx = x''$.

1.7.3. Espacios $L^p(\Omega)$

Análogamente al resultado anterior, se tiene que los espacios $L^p(\Omega)$ con $\Omega \subset \mathbb{R}^N$ medible, $1 < p < +\infty$ son reflexivos. La demostración es prácticamente idéntica a la anterior.

1.7.4. Propiedades de los espacios reflexivos

Veamos a continuación algunas propiedades de los espacios reflexivos

Proposición 1.7.1 *Sea X un espacio de Banach, entonces X es reflexivo si y sólo si X' es reflexivo.*

Demostración: Supongamos primero X reflexivo, para probar que X' es reflexivo hay que probar que la aplicación $J_{X'} \in \mathcal{L}(X', X''')$ definida por

$$\langle J_{X'} x', x''\rangle_{X''',X''} = \langle x'', x'\rangle_{X'',X'}, \quad \forall\, x' \in X',$$

es sobreyectiva. Sea entonces $x_0''' \in X'''$ y denotemos J_X a la aplicación J definida por (1.14). Como J_X pertenece a $\mathcal{L}(X, X'')$, la composición $x_0''' \circ J_X$ es lineal y continua de X a $\mathbb{R}$ y por tanto un elemento de X'. Tomemos pues $x_0' = x_0''' \circ J_X$ y vamos a probar que $J_{X'} x_0' = x_0'''$, lo que como x_0''' es arbitrario, probará que $J_{X'}$ es sobreyectiva. Para $x'' \in X''$ y $x \in X$ tal que $J_X x = x''$, el cual siempre existe ya que J_X es sobreyectiva por hipótesis, se tiene

$$\langle J_{X'} x_0', x''\rangle_{X''',X''} = \langle x'', x_0'\rangle_{X'',X'} = \langle J_X x, x_0'\rangle_{X'',X'} =$$
$$= \langle x_0', x\rangle_{X',X} = \langle x_0''', J_X x\rangle_{X''',X''} = \langle x_0''', x''\rangle_{X''',X''}.$$

Como la igualdad anterior es cierta para todo $x' \in X'$, se tiene entonces la igualdad $J_{X'} x_0' = x_0'''$.

Supongamos ahora que X' es reflexivo y definamos $J_X, J_{X'}$ como anteriormente, para probar que X es reflexivo tenemos que ver que J_X es sobreyectiva o lo que es lo mismo $J_X(X) = X''$. Ahora bien como J_X es un isomorfismo isométrico de X sobre $J_X(X)$ y X es un espacio de Banach se tiene que $J_X(X)$ es un espacio de Banach, lo que por la Proposición 1.1.2 significa que $J_X(X)$ es cerrado. Por tanto, por el Corolario 1.5.4, $J_X(X) = X''$ si y sólo si $J_X(X)^\perp = \{0\}$. Para probar esta última igualdad, consideremos $x_0''' \in J_X(X)^\perp$, lo que significa

$$\langle x_0''', J_X x\rangle_{X''',X''} = 0, \quad x \in X.$$

Como X' es reflexivo, existe $x_0' \in X'$ tal que $J_{X'}x_0' = x_0'''$ y por tanto la igualdad anterior nos da

$$0 = \langle J_{X'}x_0', J_X x\rangle_{X''',X''} = \langle J_X x, x_0'\rangle_{X'',X'} = \langle x_0', x\rangle_{X',X},$$

para todo $x \in X$ y por tanto $x_0' = 0$, con lo que al ser $J_{X'}$ lineal se deduce $x_0''' = 0$ y por tanto $J_X(X)^\perp = \{0\}$. □

Proposición 1.7.2 *Todo subespacio cerrado de un espacio reflexivo es reflexivo.*

Demostración: Sea X reflexivo y $M \subset X$ un subespacio cerrado, queremos probar que la aplicación $J_M : M \to M''$ definida por

$$\langle J_M m, m'\rangle_{M'',M'} = \langle m', m\rangle_{M',M}$$

es sobreyectiva. Consideremos entonces $m_0'' \in M''$ y definamos $x_0'' \in X''$ por

$$\langle x_0'', x'\rangle_{X'',X'} = \langle m_0'', x'_{|M}\rangle_{M'',M'}, \quad \forall\, x' \in X'.$$

Como X es reflexivo, existe $x_0 \in X$ tal que $J_X x_0 = x_0''$, donde J_X es la aplicación J definida por (1.14). Se tiene entonces

$$\langle x', x_0\rangle_{X',X} = \langle J_X x_0, x'\rangle_{X'',X'} = \langle m_0'', x'_{|M}\rangle_{M'',M'}, \quad \forall\, x' \in X'. \tag{1.15}$$

Supongamos que x_0 no pertenece a M, como M es cerrado, por la Corolario 1.5.4, existe $x' \in X'$ tal que $\langle x', x_0\rangle_{X',X} \neq 0$ y $x'_{|M} = 0$. Sin embargo esto es imposible por (1.15), luego x_0 pertenece a M y por tanto podemos considerar $J_M x_0$. Sea ahora $m' \in M'$, por el Teorema de Hahn-Banach existe $x' \in X'$ tal que $x'_{|M} = m'$. Usando (1.15) y $x_0 \in M$, deducimos

$$\langle J_M x_0, m'\rangle_{M'',M'} = \langle m', x_0\rangle_{M',M} = \langle x'_{|M}, x_0\rangle_{M',M} =$$

$$= \langle x', x_0\rangle_{X',X} = \langle m_0'', x'_{|M}\rangle_{M'',M'} = \langle m_0'', m'\rangle_{M'',M'}.$$

Como m' es arbitrario, se tiene entonces que $J_M x_0 = m_0''$. □

1.8. Convergencia débil y $*$-débil

Definimos a continuación las nociones de convergencia débil y convergencia $*$-débil que serán de gran utilidad para probar resultados de compacidad en e.n..

Definición 1.8.1 *Sea X e.n., una sucesión $\{x_n\} \subset X$ se dice que converge débilmente a un punto $x \in X$ y se escribe $x_n \rightharpoonup x$ si y sólo si*

$$\langle x', x_n\rangle \to \langle x', x\rangle, \ \forall\, x' \in X'. \tag{1.16}$$

Definición 1.8.2 *Sea X e.n., una sucesión $\{x'_n\} \subset X'$ se dice que converge $*$-débilmente a un punto $x' \in X'$ y se escribe $x'_n \stackrel{*}{\rightharpoonup} x'$ si y sólo si*

$$\langle x'_n, x\rangle \to \langle x', x\rangle, \ \forall\, x \in X. \tag{1.17}$$

Observación 1.8.3 *La noción de convergencia débil procede de una topología, denotada por $\sigma(X, X')$, que es la topología menos fina sobre X que hace continua todas las aplicaciones $\{\varphi_{x'}\}_{x'\in X'}$ definidas por:*

$$\varphi_{x'} : x \in X \longmapsto \langle x', x\rangle \in \mathbb{R}.$$

Siendo $\sigma(X, X')$ la topología menos fina que la topología definida por la norma entonces tiene menos abiertos y menos cerrados luego, en particular, si un conjunto es cerrado para esta topología entonces es cerrado para la topología inicial de X y en general la recíproca es falsa.

Observación 1.8.4 *Obsérvese que la convergencia $*$-débil no es otra cosa que la convergencia puntual.*

Observación 1.8.5 *Los conceptos de límite débil y $*$-débil son lineales, esto es el límite de la suma es la suma de los límites y el límite de un escalar por una sucesión es el escalar por el límite de la sucesión.*

Observación 1.8.6 *Está claro que si una sucesión $\{x_n\}$ converge en norma (se dice que converge fuerte) hacia un elemento $x \in X$, entonces gracias a la continuidad de los elementos de X' se tiene que $\langle x', x_n\rangle$ converge a $\langle x', x\rangle$ para todo $x \in X$ y por tanto $\{x_n\}$ converge débil a $x \in X$. Es decir, la convergencia fuerte implica la convergencia débil.*

Cuando tenemos una sucesión $\{x'_n\}$ definida en el dual X' de un espacio X podemos considerar la convergencia fuerte que sería la convergencia definida a través de la norma en X', la convergencia $*$-débil definida por la Definición 1.8.2 y la convergencia débil definida por la Definición 1.8.1, donde decir que $\{x'_n\}$ converge débilmente a x' en X' es equivalente a decir

$$\langle x'', x'_n\rangle_{X'',X'} \to \langle x'', x'\rangle_{X'',X'}, \quad \forall\, x'' \in X''. \tag{1.18}$$

Por otra parte considerando la inyección canónica J definida por (1.14), entonces decir que $\{x'_n\}$ converge $*$-débilmente a x' es equivalente a

$$\langle Jx, x'_n\rangle_{X'',X'} \to \langle Jx, x'\rangle_{X'',X'}, \quad \forall\, x \in X,$$

o lo que es lo mismo

$$\langle x'', x'_n\rangle_{X'',X'} \to \langle x'', x'\rangle_{X'',X'}, \quad \forall\, x'' \in J(X). \tag{1.19}$$

Comparando (1.18) con (1.19) vemos que en general la convergencia débil implica la convergencia $*$-débil, siendo equivalentes los dos conceptos en el caso de espacios reflexivos.

En los espacios de dimensión finita todas las convergencias anteriores son equivalentes.

1.8.1. Ejemplos

Teniendo en cuenta las identificaciones de espacios duales que vimos anteriormente se tiene las siguientes convergencias débiles.

Espacios de Hilbert

Sea X un espacio de Hilbert, una sucesión $\{x_n\}$ en X converge débilmente a $x \in X$ si y sólo si

$$(x_n|y) \to (x|y), \quad \forall\, y \in X. \tag{1.20}$$

Al estar en espacios reflexivos, la convergencia $*$-débil es equivalente a la convergencia débil.

Espacios ℓ^p

Si $1 \leq p < +\infty$, $\{x_n\} = \{x_{n,m}\} \subset \ell^p$ converge débilmente a $x = \{x_m\}$ en ℓ^p si y sólo si

$$\sum_{m=1}^{\infty} x_{n,m} y_m \to \sum_{m=1}^{\infty} x_m y_m, \quad \text{cuando } n \to \infty, \quad \forall\, y = \{y_m\} \in \ell^{p'}.$$

Si $p = +\infty$ se tiene que $\{x_n\} = \{x_{n,m}\} \subset \ell^\infty$ converge $*$-débil a $x = \{x_m\}$ en ℓ^∞ si y sólo si

$$\sum_{m=1}^{\infty} x_{n,m} y_m \to \sum_{m=1}^{\infty} x_m y_m, \quad \text{cuando } n \to \infty, \quad \forall\, y = \{y_m\} \in \ell^1.$$

Como ℓ^1 se identifica con el dual de c_0 se tiene también que $\{x_n\} = \{x_{n,m}\} \subset \ell^1$ converge $*$-débil a $x = \{x_m\}$ en ℓ^1 si y sólo si

$$\sum_{m=1}^{\infty} x_{n,m} y_m \to \sum_{m=1}^{\infty} x_m y_m, \quad \text{cuando } n \to \infty, \quad \forall\, y = \{y_m\} \in c_0.$$

Obsérvese la diferencia con la convergencia débil, que sería decir que se tiene la convergencia anterior pero para todo y en ℓ^∞ que es un espacio más grande que c_0.

Espacios $L^p(\Omega)$

Si $1 \leq p < +\infty$, se tiene que una sucesión $\{f_n\} \subset L^p(\Omega)$ converge débilmente en $L^p(\Omega)$ a $f \in L^p(\Omega)$ si y sólo si

$$\int_\Omega f_n g\, dx \to \int_\Omega f g\, dx, \quad \text{cuando } n \to \infty, \quad \forall\, g \in L^{p'}(\Omega).$$

En el caso $L^\infty(\Omega)$ se tiene que $\{f_n\} \subset L^p(\Omega)$ converge $*$-débil en $L^\infty(\Omega)$ a $f \in L^\infty(\Omega)$ si y sólo si

$$\int_\Omega f_n g\, dx \to \int_\Omega f g\, dx, \quad \text{cuando } n \to \infty, \quad \forall\, g \in L^1(\Omega).$$

Convergencia débil y no fuerte

Consideremos para todo $n \in \mathbb{N}$, $\{e_n\} = \{e_{n,m}\}$ la sucesión de sucesiones de números reales definida por $e_{n,n} = 1$, $e_{n,m} = 0$ si $m \neq n$. Esta sucesión pertenece a ℓ^1 y por tanto a ℓ^p para todo $p \in [1, +\infty]$. Si $y = \{y_n\} \in c_0$, se tiene

$$\sum_{m=1}^{\infty} e_{n,m} y_m = y_n \to 0, \tag{1.21}$$

con lo cual $\{e_n\}$ converge $*$-débil a 0 en ℓ^1. Como ℓ^p con $1 \leq p < +\infty$ está contenido en c_0, (1.21) implica también que $\{e_n\}$ converge débil a 0 en ℓ^p para todo $p \in (1, +\infty)$ y $*$-débil en ℓ^∞. Sin embargo si definimos $y = \{y_m\} \in \ell^\infty$ por $y_m = 1$, para todo $m \in \mathbb{N}$, se tiene

$$\sum_{m=1}^{\infty} e_{n,m} y_m = 1 \not\to 0, \quad \forall\, n \in \mathbb{N},$$

y por tanto $\{e_n\}$ no converge a 0 débilmente en ℓ^1. En particular, $\{e_n\}$ no converge fuertemente a cero en ℓ^1. De hecho, como

$$\|\{e_n\}\|_\infty = \sup_{n\in\mathbb{N}} |e_{n,m}| = 1,$$

y como $\|\{e_n\}\|_{\ell^\infty} \leq \|\{e_n\}\|_{\ell^p}$ para todo $p \in [1, +\infty]$ se tiene que $\{e_n\}$ no converge fuertemente a 0 en ningún espacio ℓ^p con $1 \leq p \leq +\infty$.

Si X es un espacio de Hilbert separable infinito-dimensional, es conocido que existe una base ortonormal $\{z_n\}$ de X (si X es finito-dimensional el resultado también es cierto pero entonces $\{z_n\}$ es finita), es decir $\|z_n\| = 1$, para todo $n \in \mathbb{N}$, $(z_n|z_m) = 0$ si $n \neq m$ y

$$x = \sum_{n=1}^{\infty} (x|z_n) z_n, \quad \forall\, x \in X$$

donde la suma infinita se entiende en la topología de X, i.e.

$$\lim_{m\to\infty} \|x - \sum_{n=1}^{m} (x|z_n) z_n\| = 0.$$

Además se sabe que se verifica que la sucesión $\{(x|z_n)\}$ pertenece a ℓ^2 y se tiene (identidad de Parseval)

$$\|\{(x|z_n)\}\|_{\ell^2}^2 = \sum_{n=1}^{\infty} (x|z_n)^2 = \|x\|_X^2,$$

esto permite identificar $x \in X$ con la sucesión $\{(x|z_n)\} \in \ell^2$ y por tanto identificar todo espacio de Hilbert separable con ℓ^2. El ejemplo presentado anteriormente para los espacios ℓ^p permite por tanto dar un ejemplo en todo espacio de Hilbert separable infinito-dimensional de una sucesión que converge débil pero no fuerte, de hecho como e_n se identifica con z_n y se tiene que $\{z_n\}$ converge débil a 0 pero no fuerte.

1.8.2. Propiedades

Veamos ahora algunas propiedades de la convergencia débil y $*$-débil.

Proposición 1.8.7 *Sean X, Y e.n. y $A \in \mathcal{L}(X, Y)$. Si $\{x_n\}$ converge débilmente a x en X, entonces Ax_n converge débilmente a Ax en Y.*

Demostración: Sea $\{x_n\}$ que converge débilmente a x en X. Para todo $y' \in Y'$, se tiene que $y' \circ A$ pertenece a X' y por tanto, por definición de convergencia débil se tiene

$$\langle y', Ax_n\rangle_{Y',Y} = \langle y' \circ A, x_n\rangle_{X',X} \to \langle y' \circ A, x\rangle_{X',X} = \langle y', Ax\rangle_{Y',Y},$$

para todo $y' \in Y'$. Por tanto Ax_n converge débilmente a Ax. $\square$

Observación 1.8.8 *Si A no es lineal (aun cuando sea continua), el resultado anterior no es cierto, así hemos visto antes ejemplos de sucesiones $\{x_n\}$ que convergen débilmente a cero pero no fuertemente, por tanto la sucesión $\|x_n\|$ no converge a cero en $\mathbb{R}$ (en $\mathbb{R}$ la convergencia fuerte y débil son lo mismo) aunque la aplicación $\|\cdot\| : X \to \mathbb{R}$ es continua.*

Proposición 1.8.9 *Sea X e.n. Si $\{x_n\}$ converge débilmente a $x \in X$, entonces está acotada y se verifica*

$$\|x\|_X \leq \liminf_{n\to\infty} \|x_n\|_X. \tag{1.22}$$

Análogamente, si X es un espacio de Banach y $\{x'_n\}$ una sucesión que converge $*$-débilmente a $x' \in X'$, entonces está acotada y se verifica

$$\|x'\|_{X'} \leq \liminf_{n\to\infty} \|x_n\|_{X'}. \tag{1.23}$$

Demostración: Supongamos primero $\{x_n\}$ que converge débilmente a $x \in X$. Consideramos entonces la familia de aplicaciones $Jx_n \in X'' = \mathcal{L}(X', \mathbb{R})$ (con J la inyección canónica, definida por (1.14)), la cual, por definición de convergencia débil, verifica que para todo $x' \in X'$ la sucesión $\langle Jx_n, x'\rangle_{X'',X'} = \langle x', x_n\rangle_{X',X}$ es convergente y por tanto acotada. Como X' y $\mathbb{R}$ son espacios de Banach, el Teorema (1.5.10) de Banach-Steinhaus nos dice entonces que la sucesión $\|Jx_n\|_{X''} = \|x_n\|_X$ está acotada.

Por otra parte, para todo $x' \in X'$, con $\|x'\|_X = 1$, se tiene

$$|\langle x', x\rangle| = \lim_{n\to\infty} |\langle x', x_n\rangle| \leq \|x'\|_{X'} \liminf_{n\to\infty} \|x_n\|_X = \liminf_{n\to\infty} \|x_n\|_X$$

y por tanto gracias a (1.10), del Corolario 1.5.4 de Hahn-Banach, se deduce (1.22).

En el caso de una sucesión $\{x'_n\}$ en X', basta con aplicar el Teorema 1.5.10 de Banach-Steinhaus a la familia $\{x'_n\}$ contenida en $X' = \mathcal{L}(X, \mathbb{R})$. □

1.9. Subconjuntos débilmente cerrados

Es importante en las aplicaciones saber si el límite débil de una sucesión contenida en un determinado conjunto sigue estando en dicho conjunto. Ello conduce a la siguiente definición.

Definición 1.9.1 *Sea X un e.n. y $C \subset X$, se dice que C es secuencialmente débilmente cerrado si para todo $x \in X$ tal que existe $\{x_n\} \subset C$ que converge débilmente a x se tiene que $x \in C$.*

Usando que la convergencia fuerte implica la convergencia débil es inmediato comprobar que todo subconjunto secuencialmente débilmente cerrado de un e.n. es cerrado. El recíproco sin embargo no es cierto en general aunque sí para los subconjuntos convexos.

Teorema 1.9.2 *(S. Mazur) Todo subconjunto convexo y cerrado de un e.n. es secuencialmente cerrado.*

Demostración: Se basa en el Teorema de separación de convexos, más concretamente en el Corolario 1.5.6. Sea $\{x_n\} \subset C$ que converge débilmente a x_0. Si x_0 no pertenece a C, existe entonces $x' \in X'$ tal que

$$\exists x' \in X' \text{ tal que } \quad \sup_{x\in C}\langle x', x\rangle < \langle x', x_0\rangle,$$

por tanto

$$\langle x', x_0\rangle = \lim_{n\to\infty} \langle x', x_n\rangle \leq \sup_{x\in C}\langle x', x\rangle < \langle x', x_0\rangle.$$

Esta contradicción prueba que $x_0 \in C$ y por tanto C es secuencialmente débilmente cerrado. □

1.10. Teoremas de compacidad

Una de las dificultades principales de los espacios de dimensión infinita es que deja de ser cierto el hecho de que todo subconjunto cerrado y acotado sea compacto, tal y como sucedía en los espacios finito-dimensionales. De hecho se puede probar (Teorema de Riesz, ver por ejemplo el libro de H. Brézis) que la bola unidad cerrada de un e.n. es compacta si y sólo si el espacio es finito-dimensional. Los teoremas siguientes nos muestran que este resultado sigue siendo cierto en espacios infinito-dimensionales si nos conformamos con la convergencia débil y $*$-débil.

Teorema 1.10.1 *Sea X un e.n. separable y $\{x'_n\} \subset X'$ una sucesión acotada, entonces existe una subsucesión de $\{x'_n\}$ que converge $*$-débil.*

Demostración: Sea $\{y_n\} \subset X$ denso numerable. Se tiene

$$|\langle x'_n, y_1\rangle| \leq \|x'_n\|_{X'}\|y_1\|_X,$$

con lo que el hecho de que $\{x'_n\}$ está acotada implica que la sucesión $\langle x'_n, y_1\rangle$ está acotada en $\mathbb{R}$ y por el Teorema de B. Bolzano-K. Weierstrass existe una subsucesión $\{x'_{n_{1,j}}\}$ de $\{x'_n\}$ tal que $\langle x'_{n_{1,j}}, y_1\rangle$ converge. De la misma forma, se tiene ahora que la sucesión $\langle x'_{n_{1,j}}, y_2\rangle$ está acotada y por tanto existe una subsucesión $\{x'_{n_{2,j}}\}$ de $\{x'_{n_{1,j}}\}$ tal que $\langle x'_{n_{2,j}}, y_2\rangle$ converge. También será convergente la sucesión $\langle x'_{n_{2,j}}, y_1\rangle$ al ser una subsucesión de $\langle x'_{n_{1,j}}, y_1\rangle$. Siguiendo este proceso se construyen para todo $k \in \mathbb{N}$ sucesiones $\{x'_{n_{k,j}}\}$ tales que $\{x'_{n_{k+1,j}}\}$ es una subsucesión de $\{x'_{n_{k,j}}\}$, para todo $k \in \mathbb{N}$ y $\{x'_{n_{k,j}}\}$ es una subsucesión de $\{x'_n\}$. Además para todo $k \in \mathbb{N}$ y para todo $l \in \{1, \cdots, k\}$ existe el límite de $\langle x'_{n_{k,j}}, y_l\rangle$. Se define ahora la sucesión diagonal $x'_{n_j} = x'_{n_{j,j}}$, la cual es tal que la sucesión $\{x'_{n_j}\}_{j\geq k}$ es una subsucesión de $x'_{n_{k,j}}$ para todo $k \in \mathbb{N}$ y por tanto

$$\exists \lim_{j\to\infty} \langle x'_{n_j}, y_k\rangle, \quad \forall\, k \in \mathbb{N}.$$

Sea ahora $x \in X$ entonces para todo $k \in \mathbb{N}$, como la sucesión $\langle x'_{n_j}, y_k\rangle$ es de Cauchy, se tiene

$$\limsup_{i,j\to\infty} |\langle x'_{n_j}, x\rangle - \langle x'_{n_i}, x\rangle| = \limsup_{i,j\to\infty} |\langle x'_{n_j} - x'_{n_i}, x\rangle| \leq$$

$$\leq \limsup_{i,j\to\infty} |\langle x'_{n_j} - x'_{n_i}, x - y_k\rangle| + \limsup_{i,j\to\infty} |\langle x'_{n_j} - x'_{n_i}, y_k\rangle| =$$

$$= \limsup_{i,j\to\infty} |\langle x'_{n_j} - x'_{n_i}, x - y_k\rangle| \leq \limsup_{i,j\to\infty} \|x'_{n_j} - x'_{n_i}\|_{X'}\|x - y_k\|_X.$$

Como la sucesión $\{x'_n\}$ y por tanto la sucesión $\{x'_{n_j}\}$ está acotada y como la sucesión $\{y_k\}$ es densa, se deduce que el miembro derecho de la desigualdad anterior se puede tomar tan pequeño como se quiera. Es decir

$$\limsup_{i,j\to\infty} |\langle x'_{n_j}, x\rangle - \langle x'_{n_i}, x\rangle| = 0$$

lo que significa que la sucesión $\langle x'_{n_j}, x\rangle$ es de Cauchy y por tanto convergente, es decir se ha probado

$$\exists \lim_{j\to\infty} \langle x'_{n_j}, x\rangle, \quad \forall\, x \in X.$$

Se define entonces $x' : X \to \mathbb{R}$ por

$$x'(x) = \lim_{j\to\infty} \langle x'_{n_j}, x\rangle, \quad \forall\, x \in X. \tag{1.24}$$

Fácilmente se ve que x' es lineal, además se tiene

$$|x'(x)| = \lim_{j\to\infty} |\langle x'_{n_j}, x\rangle| \leq \limsup_{j\to\infty} \|x'_{n_j}\|_{X'}\|x\|_X \leq (\sup_{n\in\mathbb{N}} \|x'_n\|_{X'})\|x\|_X,$$

para todo $x \in X$, lo que junto con la linealidad de x' prueba que x' se encuentra en X'. La propia definición (1.24) prueba entonces que $\{x'_{n_j}\}$ converge $*$-débil a x'. □

Un resultado relacionado con el anterior es cierto para espacios reflexivos. Antes, necesitamos el siguiente resultado.

Teorema 1.10.2 *Sea X un e.n. tal que X' es separable, entonces X es separable.*

Demostración: Sea $\{x'_n\}$ un subconjunto denso y numerable de X'. Por definición de norma en el espacio dual, existe $\{x_n\} \subset X$ tal que para todo $n \in \mathbb{N}$, $\|x_n\|_X = 1$ y

$$\langle x'_n, x_n\rangle \geq \frac{1}{2}\|x'_n\|_X. \tag{1.25}$$

Sea D el espacio vectorial sobre $\mathbb{Q}$ generado por los vectores x_n, $n \in \mathbb{N}$. Teniendo en cuenta

$$D = \bigcup_{n=1}^{\infty} D_n$$

con

$$D_n = \left\{ \sum_{j=1}^{n} \alpha_j x_j : \ \alpha_j \in \mathbb{Q} \right\}$$

y que el cardinal de D_n es menor o igual que el de $\mathbb{Q}^n$ se deduce que D_n es numerable y por tanto D como unión numerable de conjuntos numerables también lo es. Si probamos que D es denso en X, habremos entonces terminado la demostración. Pero por otra parte está claro que D es denso en el espacio vectorial sobre $\mathbb{R}$ generado por los vectores x_n, el cual denotamos por L. Por tanto para probar que D es denso en X, basta probar que L es denso, para lo que por el Corolario 1.5.4 basta ver que $L^\perp$ se reduce al cero de X'. Sea pues $x' \in L^\perp$, i.e. x' verifica $\langle x', x\rangle = 0$ para todo $x \in L$.

Como $\{x'_n\}$ es denso en X', para todo $\varepsilon > 0$ existe x'_n tal que $\|x' - x'_n\|_{X'} \leq \varepsilon$. Entonces usando (1.25) se verifica

$$\|x'_n\|_{X'} \leq 2\langle x'_n, x_n\rangle = 2\langle x'_n - x', x_n\rangle \leq 2\varepsilon.$$

Esto implica $\|x'\|_{X'} \leq 3\varepsilon$ para todo $\varepsilon > 0$, lo que prueba $x' = 0$ y por tanto $L^\perp = \{0\}$. □

Corolario 1.10.3 *Un espacio reflexivo es separable si y sólo si X' es separable.*

Teorema 1.10.4 *Sea X reflexivo. Entonces toda sucesión acotada en X admite una subsucesión débilmente convergente.*

Demostración: Sea $\{x_n\}$ una sucesión acotada en X. Denotamos por M la clausura del espacio generado por los elementos de la sucesión. Entonces (razonando como con el espacio L que aparecía en la demostración del Teorema 1.10.2) M es separable. Por ser un subespacio cerrado de un espacio reflexivo, la Proposición 1.7.2 prueba que M es reflexivo y por tanto equivalente a M''. Por otra parte, el Teorema 1.10.2 prueba que M' es separable con lo cual $\{x_n\}$ se puede considerar como una sucesión acotada en el espacio M'' dual del espacio separable M' y por tanto, por el Teorema 1.10.1, existen una subsucesión $\{x_{n_j}\}$ de $\{x_n\}$ y existe $x \in M$ tales que $\{x_{n_j}\}$ converge débilmente a x en M (recuérdese que en espacios reflexivos la convergencia débil y la $*$-débil es lo mismo). Es decir, para todo $m' \in M'$ se tiene

$$\langle m', x_{n_j}\rangle_{M',M} \to \langle m', x\rangle_{M',M}.$$

Ahora bien, si x' es un elemento de X', su restricción a M $x'_{|M}$ nos da un elemento de M' y por tanto se tiene

$$\langle x', x_{n_j}\rangle_{X',X} = \langle x'_{|M}, x_{n_j}\rangle_{M',M} \to \langle x'_{|M}, x\rangle_{M',M} = \langle x', x\rangle_{X',X}.$$

Como esto es cierto para todo $x' \in X'$, se deduce que $\{x_{n_j}\}$ converge débilmente a x en X. $\square$

Los resultados anteriores son consecuencias de teoremas más generales que son válidos cuando se trabaja con la topología débil. Enunciamos ahora algunos:

Teorema 1.10.5 *(Alaoglu-Bourbaki) En todo B-espacio X, la bola unidad cerrada de X', $B_{X'} = \{x' \in X' \,/\, \|x'\| \leq 1\}$, es compacta para la topología $*$-débil.*

Teorema 1.10.6 *(Bourbaki-Kakutani-Šmulian) Todo B-espacio X es reflexivo si y sólo si la bola unidad cerrada de X, $B_X = \{x \in X \,/\, \|x\| \leq 1\}$, es compacta para la topología débil.*

Teorema 1.10.7 *Todo B-espacio X es reflexivo si y sólo si la bola B_X es secuencialmente débilmente compacta.*

Teorema 1.10.8 *En todo B-espacio X separable la bola unidad cerrada $B_{X'}$ de X' es secuencialmente $*$-débilmente compacta.*

Capítulo 2

MINIMIZACIÓN DE FUNCIONALES

2.1. Descripción del problema

En el presente tema vamos a estudiar la existencia de solución a problemas de Optimización de la forma general abstracta

$$\inf_{x\in\mathcal{U}} f(x), \tag{2.1}$$

donde $\mathcal{U}$ es un conjunto dado, generalmente un subconjunto de un e.n. X, y f es una aplicación dada (un funcional) definida sobre $\mathcal{U}$ con valores en $\mathbb{R}$. A $\mathcal{U}$ se le denomina el conjunto de las restricciones o conjunto admisible y a f el funcional coste o funcional objetivo.

Obsérvese que un problema de la forma

$$\sup_{x\in\mathcal{U}} f(x)$$

se convierte en uno de la forma (2.1) sin más que cambiar f por $-f$, así pues sólo estudiaremos el problema (2.1).

Definición 2.1.1 *Sea $f : \mathcal{U} \longrightarrow \mathbb{R}$ con $\mathcal{U}$ un subconjunto de un e.n. X, se dice que f tiene un mínimo global o absoluto en un punto $\hat{x}$ si:*

$$\hat{x} \in \mathcal{U} \ \text{y} \ f(\hat{x}) \leq f(x), \quad \forall x \in \mathcal{U}. \tag{2.2}$$

En tal caso pondremos

$$f(\hat{x}) = \min_{x\in\mathcal{U}} f(x).$$

Se dice que f tiene un mínimo local o relativo en un punto $\hat{x}$ si:

$$\hat{x} \in \mathcal{U} \ \text{y} \ \exists\, \mathcal{E}(\hat{x}) \ \text{tal que} \ f(\hat{x}) \leq f(x), \quad \forall x \in \mathcal{U} \cap \mathcal{E}(\hat{x}), \tag{2.3}$$

donde $\mathcal{E}(\hat{x})$ denota un entorno del punto $\hat{x}$.

Análogamente se definen los máximos (globales, locales) y se dice que una función tiene un extremo (global, local) si tiene un mínimo o un máximo (global, local).

En este tema abordaremos el estudio de mínimos globales o absolutos. Obviamente todo mínimo global o absoluto es un mínimo local o relativo.

Observación 2.1.2 *El resultado conocido más importante referente a la existencia de mínimos globales es el Teorema de K. Weierstrass que dice: "si $\mathcal{U}$ es compacto, no vacío, y $f : \mathcal{U} \to \mathbb{R}$ es continua, entonces f tiene un máximo y un mínimo absolutos en $\mathcal{U}$", o su Corolario, que en $\mathbb{R}^N$, dice: "si $f : \mathbb{R}^N \to \mathbb{R}$ es continua y verifica $\lim_{\|x\|\to+\infty} f(x) = +\infty$, entonces f tiene un mínimo absoluto en cualquier conjunto cerrado $\mathcal{U} \subset \mathbb{R}^N$".*

Observación 2.1.3 *El método directo del Cálculo de Variaciones, para resolver el problema (2.1), consiste en elegir una sucesión minimizante $\{x_n\} \subset \mathcal{U}$, i.e. tal que*

$$\lim_{n\to\infty} f(x_n) = \inf_{x\in\mathcal{U}} f(x),$$

y probar que alguna subsucesión de $\{x_n\}$ converge a un mínimo global $\hat{x}$ de f en $\mathcal{U}$. Para que este método sea exitoso se han de tener varias condiciones:

a) *Para que la sucesión minimizante tenga una subsucesión convergente, se ha de tener alguna condición de compacidad sobre $\mathcal{U}$. Para esto debemos elegir una adecuada topología en $\mathcal{U}$.*

b) *El límite de la subsucesión minimizante ha de estar en $\mathcal{U}$. Para esto debemos tener que $\mathcal{U}$ sea cerrado en dicha topología.*

c) *Sobre el funcional f debemos tener alguna condición de continuidad, de hecho, si se busca un mínimo, basta tener la siguiente condición de semicontinuidad inferior*

$$\text{“si } x_n \text{ converge en dicha topología a } x \text{ entonces } f(x) \le \liminf_{n\to\infty} f(x_n)\text{”},$$

para que el límite de la subsucesión minimizante sea un mínimo.

La condición de semicontinuidad inferior se facilita más cuanto más restrictiva sea la topología en $\mathcal{U}$, i.e. cuanto más fina sea la topología (cuanto más abiertos tenga) ya que una función f es continua en un punto x_0 si la imagen inversa de un entorno de $f(x_0)$ es un entorno de del punto. Sin embargo elegir una topología fuerte está en contradicción con la compacidad ya que en una topología fuerte hay "pocos" compactos. Por tanto debemos elegir una topología que haga un balance entre estas condiciones.

2.2. Semicontinuidad

Para no dificultar la exposición, nos limitaremos al caso de espacios métricos.

Definición 2.2.1 *Sea U un espacio métrico, $f : \mathcal{U} \to \mathbb{R}$, diremos que f es semicontinua inferiormente, y se escribirá f s.c.i., en un punto $x \in \mathcal{U}$ si para toda sucesión $\{x_n\} \subset \mathcal{U}$ que converge a x, se tiene*

$$f(x) \le \liminf_{n\to\infty} f(x_n) \tag{2.4}$$

Diremos que f es semicontinua inferiormente (en $\mathcal{U}$) si lo es en todo punto x de $\mathcal{U}$.

Claramente, si f es continua en $\mathcal{U}$ se tiene que es semicontinua inferiormente en $\mathcal{U}$. El recíproco sin embargo no es cierto, basta considerar la función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} 1 & \text{si} \quad x \neq 0 \\ 0 & \text{si} \quad x = 0. \end{cases}$$

la cual es semicontinua en 0 pero no continua.

Las siguientes propiedades son inmediatas:

Lema 2.2.2 *Sea $\mathcal{U}$ un espacio métrico.*

1. *Si f es s.c.i. en $\mathcal{U}$ y $\lambda \geq 0$ entonces λf es s.c.i. en $\mathcal{U}$*

2. *Si f y g son s.c.i. en $\mathcal{U}$ entonces $f+g$ también.*

3. *Si f y g son s.c.i. en $\mathcal{U}$ entonces $\inf\{f, g\}$ también.*

4. *Si $\{f_i\}_{i\in I}$ es una familia cualquiera de funciones s.c.i. en $\mathcal{U}$ entonces $\sup\{f_i \,/\, i \in I\}$ es una función s.c.i. en $\mathcal{U}$.*

Se tiene la siguiente caracterización de semicontinuidad inferior.

Proposición 2.2.3 *Sea $\mathcal{U}$ un espacio métrico. Un funcional $f : \mathcal{U} \to \mathbb{R}$ es s.c.i. en $\mathcal{U}$ si y sólo si para todo $\lambda \in \mathbb{R}$ el conjunto de subnivel E_λ definido por*

$$E_\lambda = \{x \in \mathcal{U} \,/\, f(x) \leq \lambda\} \tag{2.5}$$

es un subconjunto cerrado de $\mathcal{U}$.

Observación 2.2.4 *En el resultado anterior estamos suponiendo que $\mathcal{U}$ es todo el espacio. En las aplicaciones, lo usual es que $\mathcal{U}$ sea un subconjunto de otro espacio X, con lo que el resultado anterior habrá que aplicarlo a $\mathcal{U}$ dotado de la topología relativa de X. Por tanto lo que se tendrá es que f es s.c.i. si y sólo si para todo $\lambda \in \mathbb{R}$ existe un cerrado C_λ de X, tal que $E_\lambda = \mathcal{U} \cap C_\lambda$, lo que en el caso particular de que $\mathcal{U}$ sea cerrado es equivalente a pedir que E_λ sea cerrado de X.*

Demostración de la Proposición 2.2.3: Supongamos f s.c.i. en $\mathcal{U}$ y sea $\lambda \in \mathbb{R}$. Para probar que E_λ es cerrado consideramos una sucesión $\{x_n\} \subset E_\lambda$ que converge a un punto $x \in \mathcal{U}$, tenemos que probar que $x \in E_\lambda$ o equivalentemente que $f(x) \leq \lambda$, pero esto es evidente a partir de (2.4) y $f(x_n) \leq \lambda$ para todo $n \in \mathbb{N}$.

Para el recíproco supongamos ahora que E_λ es cerrado cualquiera que sea $\lambda \in \mathbb{R}$, consideremos $x \in \mathcal{U}$, y una sucesión $\{x_n\}$ que converge a x y denotemos

$$l = \liminf_{n\to\infty} f(x_n).$$

Para probar que f es s.c.i. tendremos que probar la desigualdad $f(x) \leq l$. Por definición de límite inferior, sabemos que existe una subsucesión $\{x_{n_k}\}$ de $\{x_n\}$ tal que

$$\exists \lim_{k\to\infty} f(x_{n_k}) = l.$$

Por tanto para todo $\varepsilon > 0$ existirá k_0 tal que $f(x_{n_k}) \leq l + \varepsilon$ para todo $k \geq k_0$, lo que significa que x_{n_k} pertenece a $E_{l+\varepsilon}$ para todo $k \geq k_0$. Teniendo en cuenta que $\left\{x_{n_k}\right\}_{k\geq k_0}$ converge a x y que $E_{l+\varepsilon}$ es cerrado, se tiene entonces que x pertenece a $E_{l+\varepsilon}$. Es decir, hemos probado que $f(x) \leq l + \varepsilon$ para todo $\varepsilon > 0$ y por tanto $f(x) \leq l$. $\square$

Observación 2.2.5 *Si $\mathcal{U}$ es un espacio topológico arbitrario, se puede tomar como definición de funcional $f : \mathcal{U} \to \mathbb{R}$ semicontinuo inferiormente en $\mathcal{U}$ precisamente el carácter cerrado de los conjuntos E_λ para todo $\lambda \in \mathbb{R}$. Cuando el espacio topológico es tal que los conceptos de cerrado y secuencialmente cerrado no son equivalentes, esta definición no coincide con la dada utilizando sucesiones en la definición 2.2.1.*

2.3. Teoremas de existencia

Usando el concepto de semicontinuidad inferior se tiene (ver la Observación 2.1.3) el siguiente Teorema de Weierstrass sobre la existencia de mínimo

Teorema 2.3.1 *(Teorema de Weierstrass) Supongamos que $\mathcal{U}$ es un espacio métrico, compacto y no vacío y que $f : \mathcal{U} \to \mathbb{R}$ sea s.c.i. en $\mathcal{U}$, entonces f tiene un mínimo global en $\mathcal{U}$.*

Demostración: Sea $\mu = \inf_{x\in\mathcal{U}} f(x)$. Por definición de ínfimo, existe una sucesión $\{x_n\} \subset \mathcal{U}$ minimizante, es decir, tal que

$$\lim_{n\to\infty} f(x_n) = \mu.$$

Como $\mathcal{U}$ es compacto, existen entonces una subsucesión $\{x_{n_k}\}$ de $\{x_n\}$ y un elemento $\hat{x} \in \mathcal{U}$ tales que x_{n_k} converge a $\hat{x}$ en $\mathcal{U}$. Por ser f s.c.i., se tendrá

$$f(\hat{x}) \leq \liminf_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = \mu,$$

y por tanto $f(\hat{x}) = \mu$. □

A la hora de aplicar el Teorema anterior surge sin embargo la dificultad ya mencionada en el tema anterior de que los conjuntos cerrados y acotados de e.n. infinito-dimensionales no son en general compactos. Para salvar esta dificultad se puede tratar de usar la topología débil en lugar de la fuerte. Introducimos por tanto la siguiente definición

Definición 2.3.2 *Sea X un e.n., $\mathcal{U} \subset X$, $f : \mathcal{U} \to \mathbb{R}$. Diremos que f es secuencialmente débilmente semicontinua inferiormente (y se escribe s.d.s.c.i.) en un punto $x \in \mathcal{U}$ si dada cualquier sucesión $\{x_n\} \subset \mathcal{U}$ tal que x_n converge débilmente a x, se tiene*

$$f(x) \leq \liminf_{n\to\infty} f(x_n).$$

Diremos que f es s.d.s.c.i. en $\mathcal{U}$ si lo es en todo punto de $\mathcal{U}$.

Observación 2.3.3 *Evidentemente, si f es s.d.s.c.i., entonces es semicontinua inferiormente, pero no recíprocamente. De hecho, existen funcionales continuos que no son s.d.s.c.i.; basta por ejemplo considerar la aplicación*

$$f : x \in \ell^2 \longmapsto 1 - \|x\|_2^2 \in \mathbb{R},$$

que evidentemente es continua. Si tomamos la sucesión $\{e_n\}_{n\geq 1} \subset \ell^2$ dada por $e_n = \{e_{n,m}\}_{m\geq 1}$ con $e_{n,m} = 1$ si $n = m$, $e_{n,m} = 0$ si $n \neq m$. Ya vimos en el tema anterior que $\{e_n\}$ converge débilmente en ℓ^2 a 1, sin embargo $f(0) = 1$ y $f(e_n) = 0$ para todo n, con lo cual

$$f(0) > \liminf_{n\to\infty} f(\{e_n\}).$$

De manera análoga a la Proposición 2.2.3 se tiene el siguiente resultado (la demostración es completamente similar)

Proposición 2.3.4 *Sean X e.n., $\mathcal{U} \subset X$ secuencialmente débilmente cerrado y $f : \mathcal{U} \to \mathbb{R}$. El funcional f es s.d.s.c.i. si y sólo si para cada $\lambda \in \mathbb{R}$ el conjunto E_λ definido por (2.5) es secuencialmente débilmente cerrado.*

A fin de obtener ejemplos de funciones s.d.s.c.i., recordamos el concepto de función convexa.

Definición 2.3.5 *Sean X un espacio vectorial y $\mathcal{U} \subset X$ un subconjunto convexo. Se dice que $f : \mathcal{U} \to \mathbb{R}$ es convexa si*

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \quad \forall\, x,\, y \in \mathcal{U},\, \forall\, \alpha \in [0,1].$$

La función f se dice estrictamente convexa si

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y), \quad \forall\, x,\, y \in \mathcal{U},\; x \neq y,\; \forall\, \alpha \in (0,1).$$

Es inmediato comprobar que las funciones convexas verifican la siguiente propiedad:

Proposición 2.3.6 *Sean X un espacio vectorial, $\mathcal{U} \subset X$ convexo y $f : \mathcal{U} \to \mathbb{R}$ convexa. Entonces para cada $\lambda \in \mathbb{R}$ el conjunto E_λ definido por (2.5) es un subconjunto convexo de X.*

El recíproco de la Proposición 2.3.6 no es cierto en general; es decir, la convexidad de los conjuntos E_λ no implica la convexidad de f. Por ejemplo, la función $f : x \in \mathbb{R} \longmapsto f(x) = x^3$ no es convexa y los conjuntos de subniveles son siempre convexos.

Gracias al Teorema 1.9.2 de Mazur, se tiene entonces:

Proposición 2.3.7 *Sean X e.n., $\mathcal{U} \subset X$ convexo y cerrado, $f : \mathcal{U} \to \mathbb{R}$ convexa. Entonces f es s.c.i. en $\mathcal{U}$ si y sólo si es s.d.s.c.i. en $\mathcal{U}$.*

El resultado principal que usaremos para la existencia de solución de un problema de minimización en e.n. es el siguiente resultado.

Teorema 2.3.8 *Sean X un espacio reflexivo, $\mathcal{U} \subset X$ un subconjunto secuencialmente débilmente cerrado no vacío y $f : \mathcal{U} \to \mathbb{R}$ una función s.d.s.c.i. Suponemos además que se verifica una de las siguientes condiciones:*

i) El conjunto $\mathcal{U}$ es acotado.

ii) El conjunto $\mathcal{U}$ no es acotado pero

$$\lim_{\substack{\|x\| \to \infty \\ x \in \mathcal{U}}} f(x) = +\infty. \tag{2.6}$$

Entonces f tiene un mínimo global en $\mathcal{U}$.

Demostración: Consideremos primero el caso $\mathcal{U}$ acotado, la demostración es muy similar a la del Teorema 2.3.1. Sea $\mu = \inf_{x \in \mathcal{U}} f(x)$ y tomemos una sucesión $\{x_n\} \subset \mathcal{U}$ minimizante, i.e. tal que

$$\lim_{n \mapsto \infty} f(x_n) = \mu. \tag{2.7}$$

Como $\mathcal{U}$ es acotado, la sucesión será por tanto acotada, luego por el Teorema 1.10.4, de compacidad débil, existen entonces una subsucesión $\{x_{n_k}\}$ de $\{x_n\}$ y un elemento $\hat{x} \in X$ tales que $\{x_{n_k}\}$ converge débilmente a $\hat{x}$. Como $\mathcal{U}$ es secuencialmente débilmente cerrado, tendremos que $\hat{x}$ pertenece a $\mathcal{U}$ y como f es s.d.s.c.i., se tendrá

$$f(\hat{x}) \leq \liminf_{k \to \infty} f(x_{n_k}) = \lim_{n \to \infty} f(x_n) = \mu,$$

y por tanto $f(\hat{x}) = \mu$.

Consideremos ahora el caso en que $\mathcal{U}$ no es acotado pero se verifica (2.6). Sabemos entonces que existe $R > 0$ tal que si $x \in \mathcal{U}$ y $\|x\| > R$, se tiene $f(x) > \max\{\mu + 1, 0\}$ (μ definido como en el caso anterior, que podría ser $-\infty$). En particular, denotando $\bar{B}_R$ a la bola cerrada de centro cero y radio R, se tiene

$$\mu = \inf_{x\in\mathcal{U}} f(x) = \inf_{x\in\mathcal{U}\cap\bar{B}_R} f(x).$$

Por otra parte, como el conjunto $\bar{B}_R$ es convexo y cerrado, se deduce que es secuencialmente débilmente cerrado y por tanto es fácil ver que la intersección $\mathcal{U}\cap\bar{B}_R$ es secuencialmente débilmente cerrado. Puesto que claramente también es acotado, podemos aplicar el resultado anterior y obtener la existencia de $\hat{x}\in\mathcal{U}\cap\bar{B}_R\subset\mathcal{U}$ tal que

$$f(\hat{x}) = \inf_{x\in\mathcal{U}\cap\bar{B}_R} f(x) = \mu,$$

lo que termina la demostración del resultado. □

A la hora de aplicar el resultado anterior conviene recordar el Teorema 1.9.2 y la Proposición 2.3.7. Así se tiene en particular

Teorema 2.3.9 *Sean X un espacio reflexivo, $\mathcal{U}\subset X$ un subconjunto convexo, cerrado, no vacío y $f:\mathcal{U}\to\mathbb{R}$ una función convexa y s.c.i. Suponemos además que se verifica una de las condiciones i) o ii) del Teorema 2.3.8. Entonces f tiene un mínimo en $\mathcal{U}$.*

Obsérvese que el resultado se aplica en particular si f es convexa y continua.

Respecto de la unicidad de solución al problema de minimización (2.1), se tiene:

Proposición 2.3.10 *Sean X un espacio vectorial sobre $\mathbb{R}$, $\mathcal{U}\subset X$ convexo y $f:\mathcal{U}\to\mathbb{R}$, estrictamente convexa. Entonces existe a lo más una solución al problema (2.1).*

Demostración: Supongamos $\hat{x}_1$ y $\hat{x}_2$ dos elementos distintos de $\mathcal{U}$ soluciones de (2.1), i.e.

$$f(\hat{x}_1) = f(\hat{x}_2) = \min\{f(x)\,/\,x\in\mathcal{U}\},$$

entonces, como $\mathcal{U}$ es convexo, $\hat{x} = \frac{1}{2}(\hat{x}_1+\hat{x}_2)\in\mathcal{U}$ y, como f es estrictamente convexa, se verifica

$$f(\hat{x}) < \frac{1}{2}(f(\hat{x}_1)+f(\hat{x}_2)) = \min\{f(x)\,/\,x\in\mathcal{U}\},$$

lo cual es un absurdo. □

Obsérvese que el resultado anterior no garantiza la existencia de solución, tan sólo el hecho de que si existe solución entonces a lo más hay una.

2.4. Espacios de Sobolev unidimensionales

A fin de aplicar el Teorema 2.3.8 surge la dificultad de que el espacio en el que se encuentra planteado el problema de mínimo debe ser reflexivo. En el caso de espacios de funciones, es entonces preferible considerar por ejemplo los espacios $L^p(a,b)$ que por ejemplo el espacio $C^0([a,b])$. Sin embargo la elección del espacio debe también tener en cuenta el funcional que queremos minimizar, así por ejemplo si queremos resolver el problema de mínima distancia entre dos puntos:

$$\min\left\{\int_0^1 \sqrt{1+\dot{x}(t)^2}\,dt\,/\,x:[0,1]\to\mathbb{R},\ x(0)=0,\ x(1)=1\right\} \tag{2.8}$$

no tiene mucho sentido escoger un espacio de tipo $L^p(0,1)$ ya que las funciones de $L^p(0,1)$ no son derivables. Entre los ejemplos de espacios que vimos en el tema anterior nos veríamos entonces obligados a considerar $C^1([0,1])$ pero este espacio no es reflexivo y por tanto la elección tampoco resulta adecuada. Si queremos pues disponer de herramientas que nos permitan probar la existencia de problemas tales como (2.8) (el cual es un caso muy simple de los problemas del Cálculo de Variaciones) tendremos que introducir espacios con una estructura similar a L^p pero en los cuales tenga sentido derivar. Esta idea es la que conduce a la teoría de los espacios de S.L. Sobolev, los cuales son también de gran utilidad para resolver por ejemplo problemas de Ecuaciones en Derivadas Parciales o Problemas de Control. En el presente curso nos limitaremos a una exposición muy elemental de tales espacios que se limitará al caso de la dimensión 1, centrándonos además en el caso de intervalos acotados. Para una exposición más avanzada puede consultarse por ejemplo [2] o [6].

Definición 2.4.1 *Sea $I=(a,b)$ un intervalo abierto acotado de $\mathbb{R}$ y $p\in(1,+\infty)$. Se define el espacio $W^{1,p}(I)$ como el conjunto de funciones $u\in C^0(\bar{I})$ tales que existe una función $v\in L^p(I)$ verificando*

$$u(t)=u(a)+\int_a^t v(s)\,ds,\ \forall\, t\in\bar{I}. \tag{2.9}$$

Se pone

$$H^1(I)=W^{1,2}(I). \tag{2.10}$$

Obsérvese que la integral que aparece en el segundo miembro de (2.9) está bien definida ya que $L^p(I)$ está contenido en $L^1(I)$ y por tanto v es una función integrable. La igualdad (2.9) implica también

$$u(t)-u(s)=\int_t^s v(r)\,dr,\ \forall\, s,t\in\bar{I}. \tag{2.11}$$

Por tanto, si existen dos funciones $v_1,v_2\in L^p(I)$ verificando (2.9) con v reemplazado por v_1 y v_2 se tendría que cumplir

$$\int_t^s (v_1(r)-v_2(r))\,dr=0,\ \forall\, s,t\in\bar{I},$$

lo que implica que $v_1=v_2$ cpd I (recordar que: si $v\in L^1(\mathbb{R})$ y $\int_E v=0,\forall E$ medible, entonces $v=0$ cpd en $\mathbb{R}$) y entonces la función v que aparece en (2.9) es única como función de $L^p(I)$.

Definición 2.4.2 *Dada $u\in H^1(I)$, la única función $v\in L^2(I)$ que verifica (2.9) se llama derivada débil de u y se denota por $\dot{u}$. La igualdad (2.9) se suele escribir por tanto*

$$u(t)=u(a)+\int_a^t \dot{u}(s)\,ds,\ \forall\, t\in\bar{I}. \tag{2.12}$$

Observación 2.4.3 *Si u es una función de $C^1(\bar{I})$, la regla de I. Barrow implica que la derivada de u (en el sentido usual) verifica (2.12). Se tiene por tanto que las funciones de $C^1(\bar{I})$ pertenecen a $W^{1,p}(I)$ y además para toda función $u\in C^1(\bar{I})$ la derivada usual y la débil coinciden.*

Recíprocamente, obsérvese que si $u\in W^{1,p}(I)$ es tal que su derivada débil admite un representante continuo (recuérdese que las funciones de $L^p(I)$ son en realidad clases de funciones), entonces por (2.12) y por el Teorema Fundamental del Cálculo Integral u pertenece a $C^1(\bar{I})$.

Vamos ahora a obtener otra interpretación de la derivada débil. Se tiene los siguientes lemas fundamentales del Cálculo de Variaciones:

Lema 2.4.4 *(Lema fundamental del Cálculo de Variaciones) Sea $u \in L^p(I)$ tal que*

$$\int_a^b u(t)\varphi(t)\,dt = 0, \quad \forall\, \varphi \in C^0(\bar{I}). \tag{2.13}$$

Entonces $u(t) = 0$ cpd $t \in I$.

Demostración: Suponemos conocido el hecho que $C^0(\bar{I})$ es denso en $L^q(I)$, para todo $q \in [1, +\infty)$. (ver por ejemplo [2] o [6]). El lema se tiene por (2.13) y el Corolario 1.5.4 de Hahn-Banach con $X = L^{p'}(I)$ y $M = C^0(\bar{I})$. □

Observación 2.4.5 *Para un resultado más general puede consultarse el Lema IV.2 de [2].*

Definición 2.4.6 *Denotamos por $C_0^1(\bar{I})$ el espacio*

$$C_0^1(\bar{I}) = \{\varphi \in C^1(\bar{I}) \,/\, \varphi(a) = \varphi(b) = 0\}. \tag{2.14}$$

Lema 2.4.7 *Sea $u \in L^p(I)$ tal que*

$$\int_a^b u(t)\dot{\varphi}(t)\,dt = 0, \quad \forall\, \varphi \in C_0^1(\bar{I}). \tag{2.15}$$

Entonces existe $C \in \mathbb{R}$ tal que $u(t) = C$ cpd $t \in I$.

Demostración: En principio, para cada función $v \in L^1(I)$, se define la media $\bar{v}$ de v por

$$\bar{v} = \frac{1}{b-a}\int_a^b v(t)\,dt.$$

Fácilmente se prueba:

$$\int_a^b u(t)[v(t) - \bar{v}]\,dt = \int_a^b v(t)[u(t) - \bar{u}]\,dt, \quad \forall u \in L^p(I), v \in L^{p'}(I). \tag{2.16}$$

Sea $\psi \in C^0(\bar{I})$ y definamos φ por $\varphi(t) = \int_a^t \big(\psi(s) - \bar{\psi}\big)\,ds$, para todo $t \in \bar{I}$. Entonces $\varphi \in C_0^1(\bar{I})$ y

$$\dot{\varphi}(t) = \psi(t) - \bar{\psi}, \quad \forall\, t \in \bar{v}.$$

Por tanto, gracias a (2.13) y a (2.16) se deduce

$$0 = \int_a^b u(t)\left(\psi(t) - \bar{\psi}\right)\,dt = \int_a^b (u(t) - \bar{u})\psi(t)\,dt.$$

Teniendo en cuenta el Lema 2.4.4 anterior se deduce que $u = \bar{u}$ cpd en I. □

Como consecuencia del resultado anterior se tiene ahora el siguiente resultado

Teorema 2.4.8 *Sea $u \in L^p(I)$ tal que existe $v \in L^p(I)$ verificando*

$$\int_a^b u(t)\dot{\varphi}(t)\,dt = -\int_a^b v(t)\varphi(t)\,dt, \quad \forall\, \varphi \in C_0^1(\bar{I}). \tag{2.17}$$

Entonces existe $w \in W^{1,p}(I)$ tal que $w(t) = u(t)$ y $\dot{w}(t) = v(t)$ cpd $t \in I$.

Recíprocamente, si $u \in W^{1,p}(I)$ entonces se tiene (2.17) con $v = \dot{u}$.

Demostración: Definimos $w_0 \in W^{1,p}(I)$ por $w_0(t) = \int_a^t v(s)\,ds$, $\forall\, t \in \bar{I}$, entonces para toda $\varphi \in C_0^1(\bar{I})$, gracias al Teorema de Fubini, la regla de Barrow y $\varphi(b) = 0$ se tiene

$$\int_a^b w_0(t)\dot{\varphi}(t)\,dt = \int_a^b \left(\int_a^t v(s)\,ds\right)\dot{\varphi}(t)\,dt = \int_a^b \int_a^t v(s)\dot{\varphi}(t)\,ds\ dt =$$

$$= \int_a^b \int_s^b v(s)\dot{\varphi}(t)\,dt\ ds = \int_a^b \left(v(s)\int_s^b \dot{\varphi}(t)\,dt\right)ds = -\int_a^b v(s)\varphi(s)\,ds.$$

Usando (2.15) se tiene entonces

$$\int_a^b (w_0(t) - u(t))\dot{\varphi}(t)\,dt = 0, \quad \forall\,\varphi \in C_0^1(\bar{I})$$

y por tanto, por el Lema 2.4.7, se concluye que existe $C \in \mathbb{R}$ tal que

$$w_0(t) - u(t) = C \ \ \text{cpd } t \in I.$$

Definiendo $w = w_0 - C$ se concluye el resultado.

Recíprocamente, sea $u \in W^{1,p}(I)$. Usando (2.9), para toda $\varphi \in C_0^1(\bar{I})$ se tiene:

$$\int_a^b u(t)\dot{\varphi}(t)\,dt = \int_a^b [u(a) + \int_a^t \dot{u}(s)\,ds]\dot{\varphi}(t)\,dt = \int_a^b [\int_a^t \dot{u}(s)\,ds]\dot{\varphi}(t)\,dt =$$

$$\int_a^b [\int_s^b \dot{\varphi}(t)\,dt]\dot{u}(s)\,ds = \int_a^b [\varphi(b) - \varphi(s)]\dot{u}(s)\,ds = -\int_a^b \dot{u}(s)\varphi(s)\,ds.$$

□

Observación 2.4.9 *En lugar de definir la derivada débil de una función como lo hemos hecho en el presente tema (ver Definición 2.4.2), es usual definir la derivada débil de la siguiente forma: una función $v \in L^p(I)$ es la derivada débil de otra función $u \in L^p(I)$ si se verifica (2.17). Entonces se define $W^{1,p}(I)$ como el espacio de las funciones $u \in L^p(I)$ tales que existe su derivada débil $v \in L^p(I)$. El resultado anterior prueba entonces la existencia de una función continua w tal que $u = w$ cpd en I, con lo cual reemplazando u por w se puede admitir que las funciones de $W^{1,p}(I)$ son continuas. La ventaja de esta nueva definición, que coincide con la dada en el caso unidimensional, es que es más fácil generalizar al caso de dimensiones mayores que 1, permitiendo definir los espacios de Sobolev en abiertos de $\mathbb{R}^N$ (ver por ejemplo [2] o [6]).*

Varias de las propiedades del espacio $W^{1,p}(I)$ serán deducidas del siguiente lema.

Lema 2.4.10 *Dada $u \in W^{1,p}(I)$ y $p' \in (1, +\infty)$ el exponente conjugado de p entonces:*

$$|u(t) - u(s)| \leq |t - s|^{1/p'}\ \|\dot{u}\|_{L^p(I)}, \quad \forall\, t, s \in \bar{I}, \tag{2.18}$$

$$\|u\|_{C^0(\bar{I})} \leq \min_{s\in\bar{I}} |u(s)| + (b - a)^{1/p'}\ \|\dot{u}\|_{L^p(I)}. \tag{2.19}$$

Demostración: Gracias a (2.12), para todo $t, s \in \bar{I}$ se tiene

$$|u(t) - u(s)| = \left|\int_s^t \dot{u}(r)\,dr\right| \leq \left|\int_s^t |u(r)|\,dr\right|,$$

con lo que gracias a la desigualdad de Hölder, se tiene

$$|u(t)-u(s)| \leq \left|\int_s^t |\dot{u}(r)|^p\,dr\right|^{1/p} \left|\int_s^t 1^{p'}\,dr\right|^{1/p'} \leq \left(\int_a^b |\dot{u}(r)|^p\,dr\right)^{1/p} |t-s|^{1/p'} = \|\dot{u}\|_{L^p(I)}|t-s|^{1/p'}.$$

Esto prueba (2.18).

La demostración de (2.19) es ahora inmediata. Basta observar que (2.18) implica

$$|u(t)| \leq |u(s)| + \|\dot{u}\|_{L^p(I)}|t-s|^{1/p'} \leq |u(s)| + \|\dot{u}\|_{L^p(I)}(b-a)^{1/p'}, \quad \forall\, t,s \in \bar{I}.$$

Tomando entonces máximo en $t \in \bar{I}$ y mínimo en $s \in \bar{I}$ se deduce (2.19) □

Al espacio $W^{1,p}(I)$ se le da estructura de e.n. y a $H^1(I)$ de espacio prehilbertiano de la siguiente forma:

Definición 2.4.11 *En $W^{1,p}(I)$ se define la norma: $\|\cdot\|_{W^{1,p}(I)} : W^{1,p}(I) \longrightarrow \mathbb{R}$ por*

$$\|u\|_{W^{1,p}(I)} = \left(\|u\|^p_{L^p(I)} + \|\dot{u}\|^p_{L^p(I)}\right)^{1/p} = \left(\int_a^b |u(t)|^p\,dt + \int_a^b |\dot{u}(t)|^p\,dt\right)^{1/p}. \tag{2.20}$$

En $H^1(I)$ se define el producto escalar $(\cdot|\cdot)_{H^1(I)} : H^1(I) \times H^1(I) \to \mathbb{R}$ por

$$(u|v)_{H^1(I)} = (u|v)_{L^2(I)} + (\dot{u}|\dot{v})_{L^2(I)} = \int_a^b u(t)v(t)\,dt + \int_a^b \dot{u}(t)\dot{v}(t)\,dt. \tag{2.21}$$

El espacio $H^1(I)$ es por tanto un espacio prehilbertiano con la norma

$$\|u\|_{H^1(I)} = \left(\|u\|^2_{L^2(I)} + \|\dot{u}\|^2_{L^2(I)}\right)^{1/2} = \left(\int_a^b |u(t)|^2\,dt + \int_a^b |\dot{u}(t)|^2\,dt\right)^{1/2}. \tag{2.22}$$

Veamos ahora una proposición muy útil del espacio $W^{1,p}(I)$.

Proposición 2.4.12 *La inyección $i : W^{1,p}(I) \longrightarrow C^0(\bar{I})$, definida por $i(u) = u$, $\forall u \in W^{1,p}(I)$, es continua, i.e. existe una constante $C > 0$ tal que*

$$\|i(u)\|_{C^0(\bar{I})} = \|u\|_{C^0(\bar{I})} \leq C\|u\|_{W^{1,p}(I)}. \tag{2.23}$$

Demostración: Es evidente que la aplicación i es lineal.

En principio observamos que $\min\{|u(s)|^p \,/\, s \in \bar{I}\} \leq |u(t)|^p, \forall t \in \bar{I}$. Integrando ahora, se tiene: $(b-a)\min\{|u(s)|^p \,/\, s \in \bar{I}\} \leq \|u(t)\|^p_{L^p(I)}$ y por tanto

$$\min\{|u(s)| \,/\, s \in \bar{I}\} \leq (b-a)^{-1/p}\|u\|_{L^p(I)}. \tag{2.24}$$

La desigualdad (2.23) es consecuencia de (2.19), (2.24) y la desigualdad de Hölder bidimensional (i.e. $a_1b_1 + a_2b_2 \leq (a_1^{p'} + a_2^{p'})^{1/p'}(b_1^p + b_2^p)^{1/p}$). En efecto, se tiene

$$\|u\|_{C^0(\bar{I})} \leq \min_{s\in\bar{I}} |u(s)| + (b-a)^{1/p'}\|\dot{u}\|_{L^p(I)} \leq (b-a)^{-1/p}\|u\|_{L^p(I)} + (b-a)^{1/p'}\|\dot{u}\|_{L^p(I)}$$

$$\leq [(b-a)^{-p'/p} + (b-a)]^{1/p'}\,\|u\|_{W^{1,p}(I)}.$$

Teorema 2.4.13 *El espacio $W^{1,p}(I)$ es un espacio de Banach reflexivo y separable. $H^1(I)$ es un espacio de Hilbert separable.*

Demostración: a) Veamos que $W^{1,p}(I)$ es completo, para ello consideremos una sucesión de Cauchy $\{u_n\}$ en $W^{1,p}(I)$ es decir tal que

$$\lim_{m,n\to\infty} \|u_n - u_m\|_{W^{1,p}(I)} = 0,$$

y veamos que es convergente. Teniendo en cuenta

$$\|u_n - u_m\|_{W^{1,p}(I)} = \left(\|u_n - u_m\|^p_{L^p(I)} + \|\dot{u}_n - \dot{u}_m\|^p_{L^p(I)}\right)^{1/p},$$

se tendrá en particular que la sucesión $\{\dot{u}_m\}$ es de Cauchy en $L^p(I)$. De hecho también se tendría que $\{u_n\}$ es de Cauchy en $L^p(I)$ pero mejor que esto, podemos usar la Proposición 2.4.12 por la cual sabemos que existe $C > 0$ tal que

$$\|u_n - u_m\|_{C^0(\bar{I})} \leq C\|u_n - u_m\|_{W^{1,p}(I)}$$

y por tanto que $\{u_n\}$ es de Cauchy en $C^0(\bar{I})$. Usando que los espacios $L^p(I)$ y $C^0(\bar{I})$ son completos, tendremos por tanto la existencia de $u \in C^0(\bar{I})$ y $v \in L^p(I)$ tales que

$$u_n \to u \text{ en } C^0(\bar{I}), \quad \dot{u}_n \to v \text{ en } L^p(I). \tag{2.25}$$

Por otra parte, por definición de derivada débil se tiene

$$u_n(t) = u_n(a) + \int_a^t \dot{u}_n(s)\, ds, \quad \forall\, t \in \bar{I}. \tag{2.26}$$

Teniendo en cuenta que la convergencia en $C^0(\bar{I})$ (la cual es equivalente a la convergencia uniforme) implica la convergencia puntual y el hecho de que la convergencia en $L^p(I)$ implica la convergencia en $L^1(I)$ y por tanto en $L^1(a,t)$ para todo $t \in (a,b]$ se puede pasar al límite en (2.26) para deducir

$$u(t) = u(a) + \int_a^t v(s)\, ds, \quad \forall\, t \in \bar{I},$$

lo que prueba que u pertenece a $W^{1,p}(I)$ así como la igualdad $v = \dot{u}$ en $L^p(I)$. Usando ahora que la convergencia en $C^0(\bar{I})$ implica la convergencia en $L^p(I)$, gracias a (2.25) se deducirá también

$$\|u_n - u\|_{W^{1,p}(I)} = \left(\|u_n - u\|^p_{L^p(I)} + \|\dot{u}_n - \dot{u}\|^p_{L^p(I)}\right)^{1/p} \to 0 \ \text{ en } L^p(I)$$

y por tanto que u_n converge a u en $W^{1,p}(I)$.

b) Veamos ahora que $W^{1,p}(I)$ es un espacio reflexivo. Como el producto de espacios reflexivos es reflexivo (demostrarlo) y como $L^p(I)$ es reflexivo se tiene que $L^p(I) \times L^p(I)$ es reflexivo. Consideramos la aplicación $F : W^{1,p}(I) \longrightarrow L^p(I) \times L^p(I)$ definida por $F(u) = (u, \dot{u})$. Se tiene que F es una aplicación lineal isométrica de $W^{1,p}(I)$ en $L^p(I) \times L^p(I)$, por tanto $F(W^{1,p}(I))$ es un subespacio cerrado de $L^p(I) \times L^p(I)$. Por la Proposición 1.7.2 vemos que $F(W^{1,p}(I))$ es reflexivo y por tanto también $W^{1,p}(I)$.

c) Finalmente, veamos que $W^{1,p}(I)$ es un espacio separable. En efecto, de nuevo recordamos que $L^p(I)$ es separable y que el producto de espacios separables es separable (demostrarlo), luego $L^p(I) \times L^p(I)$ es separable. Como todo subconjunto de un espacio métrico separable es separable (demostrarlo), se tiene que $F(W^{1,p}(I)) \subset L^p(I) \times L^p(I)$ es separable. Como F es un isomorfismo isométrico de $W^{1,p}(I)$ sobre $F(W^{1,p}(I))$, se llega a que $W^{1,p}(I)$ es separable. □

En la demostración del próximo resultado vamos a usar el Teorema de Ascoli-Arzelà (ver por ejemplo [2],[10] o [11]). A fin de recordar este resultado necesitamos la siguiente definición:

Definición 2.4.14 *Un conjunto $U \subset C^0(\bar{I})$ se dice equicontinuo si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para toda $u \in U$ y para todo $t, s \in \bar{I}$ con $|t - s| < \delta$, se tiene $|u(t) - u(s)| < \varepsilon$.*

Teorema 2.4.15 *(Teorema de G. Ascoli-C. Arzelà) Un conjunto $U \subset C^0(\bar{I})$ es relativamente compacto en $C^0(\bar{I})$ (es decir, su clausura es compacta) si y sólo si es acotado y equicontinuo.*

Estudiamos ahora un resultado más fuerte que la Proposición 2.4.12.

Teorema 2.4.16 *Los conjuntos acotados de $W^{1,p}(I)$ son relativamente compactos en $C^0(\bar{I})$ (esto significa que la inyección de $W^{1,p}(I)$ en $C^0(\bar{I})$ es compacta).*

Demostración: Consideremos un conjunto acotado U en $W^{1,p}(I)$ y, por el Teorema de Ascoli-Arzelá (ver Teorema 2.4.15), vamos a probar que U es acotado y equicontinuo en $C^0(\bar{I})$.

El hecho de que U es acotado es una simple consecuencia de (2.23), lo que nos prueba la existencia de $C > 0$ tal que

$$\sup_{u \in U} \|u\|_{C^0(\bar{I})} \leq C \sup_{u \in U} \|u\|_{W^{1,p}(I)}.$$

Con respecto a la equicontinuidad vamos a usar la desigualdad (2.18), la cual implica

$$|v(t) - v(s)| \leq \sup_{u \in U} \|u\|_{W^{1,p}(I)} |t - s|^{1/p'}, \quad \forall\, t, s \in \bar{I}, \ \forall\, v \in U.$$

Por tanto dado $\varepsilon > 0$ y tomando

$$\delta = \left(\frac{\varepsilon}{\sup_{u \in U} \|u\|_{W^{1,p}(I)} + 1} \right)^{p'},$$

se deduce que para todo $v \in U$ y para todo $t, s \in \bar{I}$ con $|t - s| < \delta$ se tiene $|v(t) - v(s)| \leq \varepsilon$. □

Observación 2.4.17 *Por definición de la norma en $W^{1,p}(I)$ está claro que si una sucesión $\{u_n\}$ converge fuertemente en $W^{1,p}(I)$ si y sólo si $\{u_n\}$ y $\{\dot{u}_n\}$ convergen fuertemente en $L^p(I)$ a u y $\dot{u}$ respectivamente. Teniendo en cuenta la Proposición 2.4.12, el resultado es aún mejor, u_n no sólo converge en $L^p(I)$ sino que converge en $C^0(\bar{I})$. Análogamente también está claro, al menos en el caso $p = 2$, (usar (1.20) y la Proposición 1.8.7) que si u_n converge débilmente a u en $W^{1,p}(I)$ entonces $\dot{u}_n$ converge a $\dot{u}$ débilmente en $L^p(I)$ y u_n converge débilmente a u en $C^0(\bar{I})$. Gracias al Teorema 2.4.16 vamos a ver que de hecho esta última convergencia es fuerte.*

Veamos un lema de topología útil en lo que sigue:

Lema 2.4.18 *Sea (X, τ) un espacio topológico y $\{x_n\} \in X$. Si existe $x \in X$ tal que toda subsucesión de $\{x_n\}$ posee una subsucesión que converge a x entonces $\{x_n\}$ converge a x.*

Demostración: Supongamos lo contrario. Entonces existe un τ-entorno $\mathcal{E}(x)$ tal que para todo $n \in \mathbb{N}$ existe $m \geq n$ con $\{x_m\} \notin \mathcal{E}(x)$. Por tanto existe una subsucesión $\{x_{n_k}\}$ tal que $\{x_{n_k}\} \notin \mathcal{E}(x)$ para todo $k \in \mathbb{N}$. De esta subsucesión es imposible extraer una sucesión que converja a x.□

Teorema 2.4.19 *Si una sucesión $\{u_n\}$ converge débilmente en $W^{1,p}(I)$ hacia una función u, entonces $\{u_n\}$ converge fuertemente a u en $C^0(\bar{I})$ (y por tanto en $L^p(I)$ para todo $p \in [1, +\infty]$).*

Demostración: Como toda sucesión que converge débilmente está acotada, el Teorema 2.4.16 implica la existencia de una subsucesión $\{u_{n_k}\}$ de $\{u_n\}$ y de $\hat{u} \in C^0(\bar{I})$ tales que $\{u_{n_k}\}$ converge fuertemente a $\hat{u}$ en $C^0(\bar{I})$. Como por otra parte la convergencia débil en $W^{1,p}(I)$ implica la convergencia débil en $C^0(\bar{I})$ deducimos que la función $\hat{u}$ debe entonces coincidir con u. En particular el límite no depende de la subsucesión escogida y por tanto toda la sucesión $\{u_n\}$ converge fuertemente a u en $C^0(\bar{I})$. $\square$

Otra propiedad importante del espacio $W^{1,p}(I)$ viene dada por el siguiente resultado.

Proposición 2.4.20 *El espacio $C^1(\bar{I})$ es denso en $W^{1,p}(I)$.*

Demostración: Como $C^0(\bar{I})$ es denso en $L^p(I)$, dada $u \in W^{1,p}(I)$ (luego $\dot{u} \in L^p(I)$), existe una sucesión $v_n \in C^0(\bar{I})$ que converge fuertemente a $\dot{u}$ en $L^p(I)$. Definimos entonces $u_n \in C^1(\bar{I})$ por

$$u_n(t) = u(a) + \int_a^t v_n(s)\, ds, \quad \forall\, t \in \bar{I}.$$

Usando (2.12) y (1.5) deducimos por tanto que para todo $t \in \bar{I}$ se tiene

$$|u_n(t) - u(t)| = \left| \int_a^t (v_n(s) - \dot{u}(s))\, ds \right| \leq \int_a^t |v_n(s) - \dot{u}(s)|\, ds \leq (b-a)^{1/p'} \|v_n - \dot{u}\|_{L^p(I)},$$

lo que, gracias a que v_n converge a $\dot{u}$ en $L^p(I)$, prueba

$$\|u_n - u\|_{C^0(\bar{I})} = \max_{t \in \bar{I}} |u_n(t) - u(t)| \to 0, \text{ cuando } n \to \infty,$$

luego u_n converge a u en $C^0(\bar{I})$ lo que a su vez implica que u_n converge a u en $L^p(I)$. Como también sabemos que $\dot{u}_n = v_n$ converge a $\dot{u}$ en $L^p(I)$ se deduce que

$$\|u_n - u\|_{W^{1,p}(I)} = \left(\|u_n - u\|_{L^p(I)}^p + \|\dot{u}_n - \dot{u}\|_{L^p(I)}^p \right)^{1/p}$$

tiende a cero, cuando $n \to \infty$. Esto prueba que u_n converge fuertemente a u en $W^{1,p}(I)$ y por tanto la densidad de $C^1(\bar{I})$ en $W^{1,p}(I)$. $\square$

Observación 2.4.21 *La demostración de la Proposición 2.4.25 se ha basado en que $C^0(\bar{I})$ es denso en $L^p(I)$, i.e.*

$$\overline{C^0(\bar{I})}^{L^p(I)} = L^p(I), \tag{2.27}$$

para probar que $C^1(\bar{I})$ es denso en $W^{1,p}(I)$, i.e.

$$\overline{C^1(\bar{I})}^{W^{1,p}(I)} = W^{1,p}(I). \tag{2.28}$$

Para terminar con esta introducción a los espacios de Sobolev unidimensionales vamos ahora a definir y estudiar el espacio $W_0^{1,p}(I)$.

Definición 2.4.22 *Se define $W_0^{1,p}(I)$ como el subespacio de $W^{1,p}(I)$ formado por las funciones que se anulan en los extremos a, b de I i.e.*

$$W_0^{1,p}(I) = \{u \in W^{1,p}(I) \,/\, u(a) = u(b) = 0\}.$$

Las propiedades principales de $W_0^{1,p}(I)$ se resumen en el siguiente

Teorema 2.4.23 *Se tiene:*

1. *El espacio* $W_0^{1,p}(I)$ *es cerrado en* $W^{1,p}(I)$ *y por tanto es un espacio de Banach separable y reflexivo. El espacio* $H^1(I)$ *es un espacio de Hilbert separable.*

2. *La aplicación* $\|\cdot\|_{W_0^{1,p}(I)} : W_0^{1,p}(I)) \to \mathbb{R}$ *definida por*

$$\|u\|_{W_0^{1,p}(I)} = \|\dot{u}\|_{L^p(I)}, \quad \forall\, u \in W_0^{1,p}(I),$$

es una norma sobre $W_0^{1,p}(I)$ *equivalente a la de* $W^{1,p}(I)$.

3. *El espacio* $C_0^1(\bar{I})$ *es denso en* $W_0^{1,p}(I)$.

Demostración: 1. Sea $\{u_n\}$ una sucesión en $W_0^{1,p}(I)$ que converge en $W^{1,p}(I)$ hacia una función $u \in W^{1,p}(I)$. Como la convergencia en $W^{1,p}(I)$ implica la convergencia en $C^0(\bar{I})$ y ésta a su vez la convergencia puntual, la igualdad $u_n(a) = u_n(b) = 0$ para todo $n \in \mathbb{N}$, implica que $u(a) = u(b) = 0$ y por tanto $u \in W_0^{1,p}(I)$. Por tanto $W_0^{1,p}(I)$ es cerrado en $W^{1,p}(I)$.
2. Es fácil comprobar que $\|\cdot\|_{W_0^{1,p}(I)}$ proporciona una norma sobre $W_0^{1,p}(I)$. Para probar que es equivalente a la de $W^{1,p}(I)$ observemos que para toda $u \in W_0^{1,p}(I)$, se tiene

$$\|u\|_{W_0^{1,p}(I)} = \|\dot{u}\|_{L^p(I)} \leq \left(\|u\|_{L^p(I)}^p + \|\dot{u}\|_{L^p(I)}^p\right)^{1/p} = \|u\|_{W^{1,p}(I)}. \tag{2.29}$$

Recíprocamente, sea $u \in W_0^{1,p}(I)$ entonces, por definición ($u(a) = 0$, $u(b) = 0$) y por (2.4.3), se tiene

$$u(t) = \int_a^t \dot{u}(s)\, ds \quad \text{y} \quad u(t) = -\int_t^b \dot{u}(s)\, ds.$$

Tomando valores absolutos, sumando y aplicando la desigualdad de Hölder, se tiene

$$2|u(t)| \leq \int_a^t |\dot{u}(s)|\, ds + \int_t^b |\dot{u}(s)|\, ds = \int_a^b |\dot{u}(s)|\, ds \leq (b-a)^{1/p'} \|\dot{u}\|_{L^p(I)}.$$

Elevando a p e integrando cuando $t \in I$ obtenemos: $2^p \|u\|_{L^p(I)}^p \leq (b-a)^p \|\dot{u}\|_{L^p(I)}^p$ i.e.

$$\|u\|_{L^p(I)} \leq \frac{b-a}{2} \|\dot{u}\|_{L^p(I)}^p, \quad \forall\, u \in W_0^{1,p}(I) \tag{2.30}$$

con lo cual se tiene

$$\|u\|_{W^{1,p}(I)} = \left(\|u\|_{L^p(I)}^p + \|\dot{u}\|_{L^p(I)}^p\right)^{1/p} \leq \left(\left(\frac{b-a}{2}\right)^p \|\dot{u}\|_{L^p(I)}^p + \|\dot{u}\|_{L^p(I)}^p\right)^{1/p} =$$

$$= \left(1 + \left(\frac{b-a}{2}\right)^p\right)^{1/p} \|\dot{u}\|_{L^p(I)} = \left(1 + \left(\frac{b-a}{2}\right)^p\right)^{1/p} \|u\|_{W_0^{1,p}(I)}, \quad \forall\, u \in W_0^{1,p}(I),$$

lo que junto con (2.29) prueba la equivalencia de las dos normas.
3. Sea $u \in W_0^{1,p}(I) \subset W^{1,p}(I)$. Por la Proposición 2.4.25, sabemos que existe $\{w_n\} \subset C^1(\bar{I})$ que convergen a u en $W^{1,p}(I)$. Sean entonces $u_n \in W_0^{1,p}(I)$ definidas por

$$u_n(t) = w_n(t) - \frac{1}{b-a}\big((t-a)w_n(b) + (b-t)w_n(a)\big), \quad \forall\, t \in I, \quad \forall\, n \in \mathbb{N}.$$

Teniendo en cuenta que la convergencia en $W^{1,p}(I)$ implica la convergencia uniforme y por tanto la puntual, se deduce que $\{w_n(a)\}$ y $\{w_n(b)\}$ convergen a cero y por tanto la igualdad

$$u_n(t) - w_n(t) = -\frac{1}{b-a}\big((t-a)w_n(b) + (b-t)w_n(a)\big), \quad \forall\, t \in I$$

implica fácilmente que $u_n - w_n$ converge a cero en $W^{1,p}(I)$, con lo cual u_n converge a u en $W_0^{1,p}(I)$. □

Observación 2.4.24 *En el transcurso de la demostración de la Proposición 2.4.23 hemos probado la desigualdad (2.30), que se conoce como la desigualdad de H. Poincaré.*

Finalmente veamos una representación de los elementos del dual de $W^{1,p}(I)$.

Proposición 2.4.25 *Sea* $f \in (W^{1,p}(I))'$*, entonces existen* $u_0, u_1 \in L^{p'}(I)$ *tales que se tiene la representación*

$$f(w) = \int_I u_0 w + \int_I u_1 \dot{w}, \quad \forall w \in W^{1,p}(I) \tag{2.31}$$

y además se tiene la igualdad

$$\|f\|_{(W^{1,p}(I))'} = \left(\|u_0\|^{p'}_{L^{p'}(I)} + \|u_1\|^{p'}_{L^{p'}(I)}\right)^{1/p'}. \tag{2.32}$$

Demostración: Se denota por X el espacio producto $L^p(I) \times L^p(I)$ dotado de la norma producto:

$$\|(v_0, v_1)\|_X = \left(\|v_0\|^p_{L^p(I)} + \|v_1\|^p_{L^p(I)}\right)^{1/p}, \ \forall (v_0, v_1) \in L^p(I) \times L^p(I), \tag{2.33}$$

y se considera la aplicación $F : w \in W^{1,p}(I) \longmapsto F(w) = (w, \dot{w}) \in L^p(I) \times L^p(I)$. Entonces F es un isomorfismo isométrico de $W^{1,p}(I)$ sobre $M = F(W^{1,p}(I)) \subset L^p(I) \times L^p(I)$. La aplicación $g : (v_0, v_1) \in M \longmapsto f(F^{-1}(v_0, v_1)) \in \mathbb{R}$ es una forma lineal continua sobre M. Por el Teorema de Hahn-Banach se puede prolongar a una forma lineal continua $\tilde{g}$ sobre X tal que $\|\tilde{g}\|_{X'} = \|f\|$. Por el teorema de representación de Riesz se sabe que existen $u_0, u_1 \in L^{p'}(I)$ tales que

$$\tilde{g}((v_0, v_1)) = \int_I u_0 v_0 + \int_I u_1 v_1, \quad \forall (v_0, v_1) \in X, \tag{2.34}$$

con

$$\|\tilde{g}\|_{X'} = \|(u_0, u_1)\|_X = \left(\|u_0\|^{p'}_{L^{p'}(I)} + \|u_1\|^{p'}_{L^{p'}(I)}\right)^{1/p'}. \tag{2.35}$$

Es fácil ahora ver (2.31) y (2.32). □

2.5. Principios de dualidad para problemas de norma mínima

Los problemas de norma mínima, i.e. los problemas de distancia de un punto a un conjunto, en espacios de Hilbert son abordables mediante el Teorema 1.1.10 de la Proyección. En el caso de un e.n. dichos problemas pueden ser estudiados mediante los denominados teoremas de dualidad que exponemos a continuación. Estos teoremas son casos particulares del denominado Teorema de dualidad de Fenchel (ver [12]). Aquí vamos a tratar solamente el caso de distancia de un punto a un subespacio.

El primer resultado que exponemos es el siguiente

Teorema 2.5.1 *Sean X un e.n., M un subespacio vectorial de X y $x \in X$ entonces*

$$\inf_{m\in M} \|x-m\|_X = \max_{\substack{m'\in M^\perp \\ \|m'\|_{X'}\leq 1}} \langle m', x\rangle. \tag{2.36}$$

Sea $\hat{x}' \in M^\perp$ el elemento donde se alcanza el máximo del segundo derecho, si existe $\hat{x} \in M$ el elemento donde se alcanza el ínfimo del primer miembro (con lo cual será un mínimo) entonces $x-\hat{x}$ y $\hat{x}'$ están alineados, es decir

$$\langle \hat{x}', x-\hat{x}\rangle = \|\hat{x}'\|_{X'}\|x-\hat{x}\|_X. \tag{2.37}$$

Demostración: Sea

$$d = \inf_{m\in M} \|x-m\|_X.$$

Si $m \in M$ y $m' \in M^\perp$ con $\|m'\|_{X'} \leq 1$, entonces

$$\langle m', x\rangle = \langle m', x-m\rangle \leq \|x-m\|_X,$$

de donde tomando ínfimo en el miembro derecho y supremo en el izquierdo, se deduce,

$$d \geq \sup_{\substack{m'\in M^\perp \\ \|m'\|_{X'}\leq 1}} \langle m', x\rangle.$$

Si $d=0$, es entonces evidente que se verifica (2.36). Suponemos, pues, que $d>0$, o lo que es equivalente, que x no pertenece a $\overline{M}$. Para demostrar (2.36) basta probar que existe $\hat{x}' \in M^\perp$ con $\|\hat{x}'\|_{X'} \leq 1$ tal que $\langle \hat{x}', x\rangle = d$. Para ello sea N el espacio generado por x y M, es decir

$$N = \{\alpha x + m \,/\, \alpha \in \mathbb{R},\ m \in M\},$$

y definamos

$$f : \alpha x + m \in N \longmapsto \alpha d \in \mathbb{R}.$$

Obsérvese que puesto que x no pertenece a M se verifica que para todo $y \in N$ existen $\alpha \in \mathbb{R}$, $m \in M$ únicos tales que $y = \alpha x + m$ y por tanto f está bien definida, además se ve fácilmente que es lineal y que verifica $f(x) = d$, $f(m) = 0$ para todo $m \in M$. Se tiene también

$$|f(\alpha x + m)| = |\alpha| d \leq |\alpha| \left\| x + \frac{m}{\alpha} \right\|_X = \|\alpha x + m\|_X, \quad \forall\, \alpha \in \mathbb{R}\setminus\{0\},\ \forall\, m \in M.$$

Como la desigualdad anterior es también cierta si $\alpha = 0$, se deduce que f pertenece a N' y $\|f\|_{N'} \leq 1$. Por el Teorema 1.5.1 de Hahn-Banach podemos ahora encontrar $\hat{x}' \in X'$ tal que $\|\hat{x}'\|_{X'} = \|f\|_{N'} \leq 1$ y $\hat{x}'|_N = f$. Este $\hat{x}'$ verifica las condiciones deseadas.

Para probar la segunda parte del Teorema, sea $\hat{x} \in M$ el elemento donde se alcanza el mínimo del miembro izquierdo de (2.36) y sea $\hat{x}' \in M^\perp$ el elemento donde se alcanza el máximo. Se tiene entonces

$$\|x-\hat{x}\|_X = \langle \hat{x}', x\rangle = \langle \hat{x}', x-\hat{x}\rangle \leq \|\hat{x}'\|_{X'}\|x-\hat{x}\|_X \leq \|x-\hat{x}\|_X,$$

con lo cual todas las desigualdades anteriores son igualdades y en particular

$$\langle \hat{x}', x-\hat{x}\rangle = \|\hat{x}'\|_{X'}\|x-\hat{x}\|_X,$$

lo que prueba la condición de alineamiento (2.37). □

Análogamente al resultado precedente se tiene el siguiente que resulta más interesante a efectos prácticos.

Teorema 2.5.2 *Sean X un e.n., M un subespacio vectorial de X y $x' \in X'$, entonces*

$$\min_{m' \in M^\perp} \|x' - m'\| = \sup_{\substack{m \in M \\ \|m\|_X \leq 1}} \langle x', m \rangle \tag{2.38}$$

Además, si el supremo en el miembro derecho de (2.38) se alcanza en algún $\hat{x} \in M$ con $\|\hat{x}\| \leq 1$ y si $\hat{x}'$ es donde se alcanza el mínimo en el miembro izquierdo entonces $x' - \hat{x}' \in X'$ y $\hat{x} \in X$ están alineados i.e.

$$\langle x' - \hat{x}', \hat{x} \rangle = \|x' - \hat{x}'\|_{X'} \|\hat{x}\|_X. \tag{2.39}$$

Demostración: Para todo $m' \in M^\perp$ y $m \in M$ con $\|m\|_X \leq 1$ se tiene

$$\langle x', m \rangle = \langle x' - m', m \rangle \leq \|x' - m'\|_{X'} \|m\|_X \leq \|x' - m'\|_{X'}.$$

Tomando supremo en el miembro izquierdo e ínfimo en el derecho se tiene entonces

$$\inf_{m' \in M^\perp} \|x' - m'\| \geq \sup_{\substack{m \in M \\ \|m\|_X \leq 1}} \langle x', m \rangle$$

Para terminar la demostración de (2.38) basta entonces probar la existencia de $\hat{x}' \in M^\perp$ tal que

$$\|x' - \hat{x}'\|_{X'} = \sup_{\substack{m \in M \\ \|m\|_X \leq 1}} \langle x', m \rangle = \|x'\big|_M\|_{M'}.$$

Para ello usamos el Teorema de Hahn-Banach y entonces existe $\tilde{x}' \in X'$ tal que

$$\tilde{x}'\big|_M = x'\big|_M, \quad \|\tilde{x}'\|_{X'} = \|x'\big|_M\|_{M'}.$$

El elemento $\hat{x}' = x' - \tilde{x}'$ verifica entonces las condiciones deseadas.

Para demostrar la segunda parte del Teorema sea ahora $\hat{x}' \in M^\perp$ el elemento donde se alcanza el mínimo del miembro izquierdo de (2.38) y supongamos que existe $\hat{x} \in M$ con $\|\hat{x}\|_X \leq 1$ donde se alcanza el supremo, entonces

$$\|x' - \hat{x}'\|_{X'} = \langle x', \hat{x} \rangle = \langle x' - \hat{x}', \hat{x} \rangle \leq \|x' - \hat{x}'\|_{X'} \|\hat{x}\|_X \leq \|x' - \hat{x}'\|_{X'}.$$

Por tanto todas las desigualdades anteriores son igualdades y en particular se tiene la condición de alineamiento (2.39). □

Observación 2.5.3 *Como consecuencia del resultado precedente se concluye que, ante un problema concreto de norma mínima, es interesante plantear éste en el dual de un e.n. conocido, ya que en ese caso tendremos garantizada la existencia de solución al problema.*

Como corolario del Teorema 2.5.2, se tiene el siguiente resultado de gran interés en las aplicaciones

Corolario 2.5.4 *Sean X e.n., $x_1, \cdots, x_n \in X$ linealmente independientes, y $c_1, \cdots, c_n \in \mathbb{R}$. Supongamos que el conjunto*

$$U = \{x' \in X' \,/\, \langle x', x_i \rangle = c_i, \; i = 1, ..., n\}$$

es no vacío. Entonces

$$\min_{x' \in U} \|x'\|_{X'} = \max_{\substack{a_i \in \mathbb{R} \\ \left\|\sum_{i=1}^n a_i x_i\right\|_X \leq 1}} \sum_{i=1}^{n} a_i c_i \tag{2.40}$$

Además, si $\hat{x}' \in U$ *es un punto donde se alcanza el mínimo en el miembro izquierdo de (2.40) y* $\hat{a} = (\hat{a}_1, ..., \hat{a}_n) \in \mathbb{R}^N$ *tal que* $\|\sum_{i=1}^n \hat{a}_i x_i\|_X \leq 1$*, es un punto donde se alcanza el máximo en el miembro derecho, entonces* $\hat{x}'$ *y* $\sum_{i=1}^n \hat{a}_i x_i$ *están alineados, es decir,*

$$\langle \hat{x}', \sum_{i=1}^{n} \hat{a}_i x_i \rangle = \|\hat{x}'\|_{X'} \left\| \sum_{i=1}^{n} \hat{a}_i x_i \right\|_X . \tag{2.41}$$

Demostración: El hecho de que el supremo en el miembro derecho de (2.40) se alcanza es consecuencia de que el conjunto

$$\{a \in \mathbb{R}^N \,/\, \| \sum_{i=1}^{n} a_i x_i \|_X \leq 1\}$$

es un compacto de $\mathbb{R}^N$, y la aplicación $a \in \mathbb{R}^N \longmapsto f(a) = \sum_{i=1}^n a_i x_i$ es continua en $\mathbb{R}^N$.

Por otra parte, como U es no vacío por hipótesis, si fijamos $x'_0 \in U$, y denotamos por M al subespacio vectorial de X generado por los x_i entonces se tiene $U = x'_0 + M^\perp$, con lo que

$$\inf_{x' \in U} \|x'\|_{X'} = \inf_{m' \in M^\perp} \|x'_0 - m'\|.$$

Por el Teorema 2.5.2, el ínfimo anterior se alcanza y además,

$$\min_{m' \in M^\perp} \|x'_0 - m'\| = \max_{\substack{m \in M \\ \|m\|_X \leq 1}} \langle x', m \rangle = \max_{\substack{a_i \in \mathbb{R} \\ \left\|\sum_{i=1}^n a_i x_i\right\|_X \leq 1}} \sum_{i=1}^{n} a_i c_i .$$

Finalmente, la condición de alineamiento (2.39) prueba la igualdad (2.41). □

A fin de poder usar el resultado anterior es conveniente saber que significa la condición de alineamiento en algunas situaciones particulares. El primer resultado es bien conocido y por tanto no incluimos su demostración.

Proposición 2.5.5 *Sea* X *un espacio prehilbertiano y sean* $x_1, x_2 \in X$, $x_1 \neq 0$. *Entonces*

$$(x_1|x_2) = \|x_1\|_X \|x_2\|_X \quad \Leftrightarrow \quad \textit{existe} \quad \lambda \geq 0 \quad \textit{tal que} \quad x_2 = \lambda x_1. \tag{2.42}$$

La siguiente condición de alineamiento sigue directamente de las condiciones para la igualdad en la desigualdad de Hölder.

Proposición 2.5.6 *Sean* $\Omega \subset \mathbb{R}^N$ *medible,* $p \in (1, +\infty)$, $1/p + 1/p' = 1$, $f \in L^p(\Omega)$, *y* $g \in L^{p'}(\Omega)$*, entonces* f *y* g *están alineadas, es decir*

$$\int_\Omega f(x) g(x)\, dx = \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p} \left(\int_\Omega |g(x)|^{p'} \, dx \right)^{1/p'} . \tag{2.43}$$

si y sólo si

$$\left(\frac{|f(x)|}{\|f\|_p} \right)^p = \left(\frac{|g(x)|}{\|g\|_{p'}} \right)^{p'} \quad \textit{i.e.} \quad g(x) = K|f(x)|^{p/p'} \textit{sgn } f(x) \ \textit{cpd } x \in \Omega, \tag{2.44}$$

para alguna constante $K \geq 0$.

Finalmente veamos el alineamiento entre funciones integrables y acotadas.

Proposición 2.5.7 *Sea* $\Omega \subset \mathbb{R}^N$ *medible,* $f \in L^1(\Omega)$, $g \in L^\infty(\Omega)$, *entonces* f *y* g *están alineadas, es decir*

$$\int_\Omega fg\,dx = \|f\|_{L^1(\Omega)}\|g\|_{L^\infty(\Omega)}. \tag{2.45}$$

si y sólo si g *y* f *verifican*

$$g(x) = \|g\|_{L^\infty(\Omega)} sgn\ f(x) \quad i.e. \quad g(x) = K sgn\ f(x)\ cpd\ x \in \Omega, \tag{2.46}$$

para alguna constante $K \geq 0$.

Demostración: Supongamos que se verifica (2.45) entonces se tiene

$$\|g\|_{L^\infty(\Omega)}\|f\|_{L^1(\Omega)} = \int_\Omega gf\,dx \leq \int_\Omega |g||f|\,dx \leq \int_\Omega \|g\|_{L^\infty(\Omega)}|f|\,dx \leq \|g\|_{L^\infty(\Omega)}\|f\|_{L^1(\Omega)}. \tag{2.47}$$

Luego todas las desigualdades deben ser igualdades. Se tendrá por tanto

$$gf = |g||f| \text{ cpd } \Omega$$

y por tanto g y f tienen el mismo signo. Además se tiene que cumplir

$$|g||f| = \|g\|_{L^\infty(\Omega)}|f| \quad \text{cpd } \Omega$$

que junto con la condición de signo implica (2.46).

La demostración del recíproco es evidente. □

Capítulo 3

OPTIMIZACIÓN DIFERENCIABLE

3.1. Derivación en los sentidos de Gâteaux y Fréchet

En esta sección vamos a realizar una breve introducción a la teoría de derivación en e.n..

3.1.1. Definiciones y primeras propiedades

Definición 3.1.1 *Sean X, Y e.n., $U \subset X$ abierto, $F : U \to Y$, $x_0 \in U$, $h \in X$. Diremos que F es diferenciable en el sentido de R. Gâteaux (brevemente, G-diferenciable) en x_0 y en la dirección h si existe el límite en Y*

$$\lim_{\varepsilon \to 0} \frac{F(x_0 + \varepsilon h) - F(x_0)}{\varepsilon}.$$

En este caso se denota

$$\delta F(x_0, h) = \lim_{\varepsilon \to 0} \frac{F(x_0 + \varepsilon h) - F(x_0)}{\varepsilon}. \tag{3.1}$$

y al elemento $\delta F(x_0, h) \in Y$, así definido, se le denomina la G-diferencial de F en el punto x_0 y en la dirección h.

Si existe $\delta F(x_0, h)$ para toda $h \in X$, diremos que F es G-diferenciable en x_0, y a la aplicación $\delta F(x_0)$ definida por

$$\delta F(x_0) : h \in X \longmapsto \delta F(x_0, h) \in Y \tag{3.2}$$

la denominaremos la G-diferencial (o diferencial Gâteaux) de F en el punto x_0.

Observación 3.1.2 *En la definición precedente no es necesaria estructura topológica para X; basta, para que las definiciones tengan sentido, que X sea un espacio vectorial sobre $\mathbb{R}$ y que x_0 sea un punto interno de U, es decir, tal que dado $h \in X$ exista un $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ se tenga $x_0 + \varepsilon h \in U$.*

Observación 3.1.3 *Es inmediato demostrar que si F es G-diferenciable en x_0 entonces la G-diferencial es homogénea en h, es decir*

$$\delta F(x_0, \alpha h) = \alpha \delta F(x_0, h), \quad \forall\, \alpha \in \mathbb{R},\ \forall\, h \in X.$$

Además, la aplicación $t \longmapsto F(x_0 + th)$ es continua en $t = 0$.

Definición 3.1.4 *Sean X, Y e.n., $U \subset X$ abierto, $x_0 \in U$, $F : U \to Y$ G-diferenciable en x_0. Si aplicación $\delta F(x_0)$ pertenece a $\mathcal{L}(X,Y)$, entonces se dice que F es G-derivable en x_0 y a la aplicación $\delta F(x_0)$ se la denomina la derivada Gâteaux (G-derivada) de F en x_0.*

Observación 3.1.5 *Es inmediato comprobar que decir que F es G-derivable en x_0 es equivalente a decir que existe $A(x_0) \in \mathcal{L}(X,Y)$ tal que*

$$\lim_{\varepsilon\to 0} \left\| \frac{F(x_0+\varepsilon h) - F(x_0)}{\varepsilon} - A(x_0)h \right\|_Y = 0, \quad \forall\, h \in X,$$

y en tal caso $A(x_0)$ está determinado unívocamente por $A(x_0) = \delta F(x_0)$.

Es evidente que si F es G-derivable en x_0 entonces F es G-diferenciable en x_0 y entonces F es G-diferenciable en x_0 en la dirección h. En general, salvo si $X = Y = \mathbb{R}$, no se tienen las implicaciones contrarias.

Las definiciones precedentes generalizan el concepto de derivada direccional para aplicaciones de $\mathbb{R}^N$ en $\mathbb{R}$. Un concepto más restrictivo lo constituye la noción de derivada en el sentido de M. Fréchet.

Definición 3.1.6 *Sean X, Y e.n., $U \subset X$ abierto, $F : U \to Y$, $x_0 \in U$. Diremos que F es derivable en el sentido de M. Fréchet (o F-derivable) en x_0, si existe $A(x_0) \in \mathcal{L}(X,Y)$ tal que*

$$\lim_{h\to 0} \frac{\|F(x_0+h) - F(x_0) - A(x_0)h\|_Y}{\|h\|_X} = 0. \tag{3.3}$$

En tal caso $A(x_0)$ es único (ver la Proposición 3.1.7), se le denota $F'(x_0)$ y se le denomina la derivada Fréchet (o F-derivada) de F en x_0.

Proposición 3.1.7 *Si F es F-derivable en x_0, entonces es G-derivable en x_0, $\delta F(x_0) \in \mathcal{L}(X,Y)$, $F'(x_0) = \delta F(x_0)$ y F es continua en x_0.*

Demostración: Sea $h \in X \setminus \{0\}$ y $A(x_0) \in \mathcal{L}(X,Y)$ satisfaciendo (3.3), entonces para $\varepsilon \neq 0$ se tiene

$$0 \leq \left\| \frac{F(x_0+\varepsilon h) - F(x_0)}{\varepsilon} - A(x_0)h \right\|_Y = \|h\|_X \frac{\|F(x_0+\varepsilon h) - F(x_0) - A(x_0)(\varepsilon h)\|_Y}{\|\varepsilon h\|_X}.$$

Usando (3.3), podemos pasar al límite cuando $\varepsilon \to 0$ en la desigualdad precedente para probar

$$A(x_0)h = \lim_{\varepsilon\to 0} \frac{F(x_0+\varepsilon h) - F(x_0)}{\varepsilon},$$

de donde se deduce que F es derivable Gâteaux en x_0 y $\delta F(x_0) = A(x_0)$. En particular, esto prueba la unicidad de $A(x_0)$ mencionada anteriormente.

La continuidad de F en x_0 se sigue observando que (3.3) implica

$$F(x_0+h) = F(x_0) + A(x_0)h + \|h\| r(x_0,h) \tag{3.4}$$

donde

$$r(x_0,h) = \frac{F(x_0+h) - F(x_0) - A(x_0)h}{\|h\|_X} \to 0 \ \text{ en } Y, \ \text{ cuando } h \to 0.$$

$\square$

Ejemplos

1. Es inmediato comprobar que si X e Y son e.n., $A : X \to Y$ es lineal entonces A es G-diferenciable en todo punto x_0 de X, con $\delta A(x_0) = A$. Si además A pertenece a $\mathcal{L}(X,Y)$, entonces A es F-derivable en todo punto x_0 de X.

2. Supongamos X, Y, Z e.n., $a(.,.) : X \times Y \to Z$ bilineal, entonces a es G-diferenciable en todo punto (x_0, y_0) de $X \times Y$ con

$$\delta F((x_0,y_0),(h,k)) = a(x_0,k) + a(h,y_0), \quad \forall\, (x_0,y_0),(h,k) \in X \times Y.$$

Si a es continua entonces es F-derivable.

Un concepto de derivada más restrictivo que el de aplicación F-derivable y que usaremos más adelante es el siguiente.

Definición 3.1.8 *Sean X, Y e.n., $U \subset X$ abierto, $F : U \to Y$, $x_0 \in U$. Diremos que F es estrictamente derivable en x_0, si existe $A(x_0) \in \mathcal{L}(X,Y)$ verificando que para todo $\varepsilon > 0$ existe un $\delta = \delta(x_0, \varepsilon) > 0$ tal que $B(x_0, \delta) \subset U$ y*

$$(3.5) \qquad \|F(x_1) - F(x_2) - A(x_0)(x_1 - x_2)\|_Y \le \varepsilon \|x_1 - x_2\|_X, \quad \forall\, x_1, x_2 \in B(x_0, \delta).$$

En tal caso, F es derivable Fréchet en x_0 con $F'(x_0) = A(x_0)$, y escribiremos $F \in SD^1(x_0)$.

Observación 3.1.9 *F derivable Fréchet en x_0 no implica $F \in SD^1(x_0)$.*

Observación 3.1.10 *Todas las nociones de diferencial y derivada que se han introducido poseen carácter lineal en F.*

3.1.2. Regla de la cadena

Teorema 3.1.11 *(Regla de la cadena) Sean X, Y, Z e.n., $U \subset X$ abierto, $x_0 \in U$, $V \subset Y$ abierto. Consideremos dos aplicaciones $\varphi : U \to V$, $\psi : V \to Z$, y denotemos $y_0 = \varphi(x_0)$, $F = \psi \circ \varphi$.*

Si ψ es F-derivable en y_0 y φ es F-derivable (respectivamente, G-derivable, G-diferenciable, G-diferenciable en la dirección h de X) en x_0, entonces F es F-derivable (respectivamente, G-derivable, G-diferenciable o G-diferenciable en la dirección h de X) en x_0, y se satisface

$$(3.6) \qquad F'(x_0) = \psi'(y_0) \circ \varphi'(x_0),$$

(respectivamente, $\delta F(x_0) = \psi'(y_0) \circ \delta\varphi(x_0)$, $\delta F(x_0, h) = \psi'(y_0)\,(\delta\varphi(x_0, h))$). Además, si $\varphi \in SD^1(x_0)$ y $\psi \in SD^1(y_0)$, entonces $F \in SD^1(x_0)$.

Demostración: a) Supongamos que φ es G-diferenciable en x_0 en la dirección h, es decir, que existe

$$\delta\varphi(x_0, h) = \lim_{\varepsilon \to 0} \frac{\varphi(x_0 + \varepsilon h) - \varphi(x_0)}{\varepsilon}.$$

Como ψ es derivable Fréchet en $y_0 = \varphi(x_0)$, para todo y en un entorno de y_0 se satisface

$$\psi(y) = \psi(y_0) + \psi'(y_0)(y - y_0) + \|y - y_0\|_Y r(y) \quad \text{con} \lim_{y \to y_0} r(y) = 0,$$

lo que unido a

$$\lim_{\varepsilon \to 0} \varphi(x_0 + \varepsilon h) = y_0,$$

permite afirmar que si $|\varepsilon|$ es suficientemente pequeño entonces

$$\frac{F(x_0+\varepsilon h)-F(x_0)}{\varepsilon}=\frac{\psi(\varphi(x_0+\varepsilon h))-\psi(y_0)}{\varepsilon}=$$

$$=\frac{\psi'(y_0)(\varphi(x_0+\varepsilon h)-y_0)+\|\varphi(x_0+\varepsilon h)-y_0\|_Y r(\varphi(x_0+\varepsilon h))}{\varepsilon}.$$

Pero como $\psi'(y_0)\in\mathcal{L}(X,Y)$ se tiene

$$\lim_{\varepsilon\to 0}\frac{\psi'(y_0)(\varphi(x_0+\varepsilon h)-y_0)}{\varepsilon}=\psi'(y_0)\left(\lim_{\varepsilon\to 0}\frac{\varphi(x_0+\varepsilon h)-y_0}{\varepsilon}\right)=\psi'(y_0)\left(\delta\varphi(x_0,h)\right)$$

y

$$\lim_{\varepsilon\to 0}\frac{\|\varphi(x_0+\varepsilon h)-y_0\|_Y r(\varphi(x_0+\varepsilon h))}{\varepsilon}=0,$$

ya que, por continuidad, $r(\varphi(x_0+\varepsilon h))$ tiende a cero con ε, y el cociente

$$\frac{\|\varphi(x_0+\varepsilon h)-y_0\|_Y}{\varepsilon}$$

se mantiene acotado.

b) Supongamos ahora que φ y ψ son estrictamente derivables en x_0 e y_0 respectivamente. Consideremos fijado $\varepsilon>0$ y tomemos $\varepsilon_1>0$ y $\varepsilon_2>0$ tales que

$$\varepsilon_1\|\psi'(y_0)\|_{\mathcal{L}(Y,Z)}+\varepsilon_2\|\varphi'(x_0)\|_{\mathcal{L}(X,Y)}+\varepsilon_1\varepsilon_2\le\varepsilon. \tag{3.7}$$

Entonces existen $\delta_1>0$ y $\delta_2>0$ tales que $B_1=\bar{B}(x_0,\delta_1)\subset U$, $B_2=\bar{B}(y_0,\delta_2)\subset V$ y:

$$\|\varphi(x_1)-\varphi(x_2)-\varphi'(x_0)(x_1-x_2)\|_Y\le\varepsilon_1\|x_1-x_2\|_X\quad\forall\, x_1,x_2\in B_1 \tag{3.8}$$

$$\|\psi(y_1)-\psi(y_2)-\psi'(y_0)(y_1-y_2)\|_Z\le\varepsilon_2\|y_1-y_2\|_Y\quad\forall\, y_1,y_2\in B_2. \tag{3.9}$$

Tomemos

$$\delta=\min\left(\delta_1,\frac{\delta_2}{\varepsilon_1+\|\varphi'(x_0)\|_{\mathcal{L}(X,Y)}}\right)\le\delta_1.$$

Sean x_1 y $x_2\in B(x_0,\delta)\subset B_1$. Por (3.8)

$$\|\varphi(x_1)-\varphi(x_2)\|_Y\le(\varepsilon_1+\|\varphi'(x_0)\|_{\mathcal{L}(X,Y)})\|x_1-x_2\|_X,\quad\forall x_1,x_2\in B(x_0,\delta) \tag{3.10}$$

De aquí, por la definición de δ, se deduce $\varphi(x_1),\varphi(x_2)\in B_2$. Así pues

$$\|F(x_1)-F(x_2)-\psi'(y_0)\left(\varphi'(x_0)(x_1-x_2)\right)\|_Z\le$$

$$\le\|\psi(\varphi(x_1))-\psi(\varphi(x_2))-\psi'(y_0)(\varphi(x_1)-\varphi(x_2))\|_Z+\|\psi'(y_0)\left(\varphi(x_1)-\varphi(x_2)-\varphi'(x_0)(x_1-x_2)\right)\|_Z\le$$

$$\le\varepsilon_2\|\varphi(x_1)-\varphi(x_2)\|_Y+\|\psi'(y_0)\|_{\mathcal{L}(Y,Z)}\varepsilon_1\|x_1-x_2\|_X\le$$

$$\le\left(\varepsilon_2(\varepsilon_1+\|\varphi'(x_0)\|_{\mathcal{L}(X,Y)})+\varepsilon_1\|\psi'(y_0)\|_{\mathcal{L}(Y,Z)}\right)\|x_1-x_2\|_X\le\varepsilon\|x_1-x_2\|_X,$$

y en consecuencia $F=\psi\circ\varphi$ pertenece a $SD^1(x_0)$.

c) Si φ y ψ son derivables Fréchet en x_0 e y_0 respectivamente, el razonamiento se realiza como en b) tomando $x_2=x_0$ y $x_1=x_0+h$. La demostración del resto del Teorema es inmediata. □

Observación 3.1.12 *El Teorema precedente es falso si solamente se supone que ψ es G-derivable.*

Ejemplo. Sean X, Y e.n., $a : X \times X \to Y$ bilineal y continua. Vamos a ver que la aplicación $J : X \to Y$ definida por $J(x) = a(x, x)$ para todo $x \in X$, es F-derivable.

Claramente, $J = a \circ D$, con $D : X \to X \times X$ definida por $Dx = (x, x)$, para todo $x \in X$. Como D es lineal y continua, entonces D es F-derivable para todo $x \in X$ y se tiene

$$D'(x_0)h = Dh = (h, h), \quad \forall\, h \in X.$$

Además, sabemos que a es F-derivable y

$$a'(x_0, y_0)(h, k) = a(x_0, k) + a(h, y_0), \quad \forall\, (x_0, y_0), (h, k) \in X \times Y.$$

Usando la regla de la cadena, se tendrá que J es F-derivable en todo $x_0 \in X$ y se tiene

$$J'(x_0)h = (a \circ D)'(x_0)h = a'(Dx_0)D'(x_0)h = a'(x_0, x_0)(h, h) = a(x_0, h) + a(h, x_0).$$

En particular, si a es simétrica, se tiene $J'(x_0)h = 2a(x_0, h)$.

3.1.3. Teorema del valor medio

Vamos a tratar ahora de generalizar el Teorema del valor medio al caso de funciones G-diferenciables. Denotamos a los segmentos que une los puntos x_1, x_2 por

$$[x_1, x_2] = \{x_1 + t(x_2 - x_1) \,/\, t \in [0, 1]\}, \quad (x_1, x_2) = \{x_1 + t(x_2 - x_1) \,/\, t \in (0, 1)\}.$$

Se tiene el siguiente resultado.

Teorema 3.1.13 *(Valor medio) Sean X un e.n., $U \subset X$ abierto y x_1, x_2 dos puntos de U tales que $[x_1, x_2] \subset U$. Se tiene*

1. *Si $F : U \to \mathbb{R}$ es G-diferenciable en todos los puntos de $[x_1, x_2]$ en la dirección $x_2 - x_1$, entonces existe $x_0 \in (x_1, x_2)$ tal que*

$$F(x_2) - F(x_1) = \delta F(x_0, x_2 - x_1) \tag{3.11}$$

2. *Si Y es un e.n. y $F : U \to Y$ es G-diferenciable en todos los puntos de $[x_1, x_2]$ en la dirección $x_2 - x_1$, entonces existe $x_0 \in (x_1, x_2)$ tal que*

$$\|F(x_2) - F(x_1)\|_Y \le \|\delta F(x_0, x_2 - x_1)\|_Y. \tag{3.12}$$

Demostración: En el caso 1., se considera $\varphi(t) = F(x_1 + t(x_2 - x_1))$, definida y derivable en todo el intervalo $[0, 1]$, con derivada

$$\varphi'(t) = \delta F(x_1 + t(x_2 - x_1), x_2 - x_1).$$

Aplicando el Teorema del valor medio en $\mathbb{R}$ a φ se deduce entonces (3.11).

En el caso 2., por el Corolario 1.5.4 de Hahn-Banach, existe $y' \in Y'$ con $\|y'\|_{Y'} = 1$ tal que

$$\langle y', F(x_2) - F(x_1)\rangle_{Y',Y} = \|F(x_2) - F(x_1)\|_Y.$$

Una vez elegido este y', se define $\varphi : [0, 1] \to \mathbb{R}$ por $\varphi(t) = \langle y', F(x_1 + t(x_2 - x_1))\rangle$ y se observa que gracias a la regla de la cadena, la derivada de φ viene dada por

$$\varphi'(t) = \langle y', \delta F(x_1 + t(x_2 - x_1), x_2 - x_1)\rangle, \quad \forall\, t \in [0, 1].$$

Aplicando el Teorema del valor medio en $\mathbb{R}$ a φ se tiene entonces que existe $t \in (0,1)$ verificando

$$\|F(x_2) - F(x_1)\|_Y = \langle y', F(x_2) - F(x_1)\rangle_{Y',Y} = \varphi(1) - \varphi(0) =$$

$$= \langle y', \delta F(x_1 + t(x_2 - x_1), x_2 - x_1)\rangle \leq \|y'\|_{Y'} \|\delta F(x_1 + t(x_2 - x_1), x_2 - x_1)\|_Y,$$

lo que junto con $\|y'\|_{Y'} = 1$ prueba (3.12) con $x_0 = x_1 + t(x_2 - x_1)$. □

Como consecuencia del resultado precedente, podemos obtener un criterio de derivabilidad estricta. Para ello introducimos la definición.

Definición 3.1.14 *Sean X e Y e.n., $U \subset X$ un abierto y $F : U \to Y$. Se dice que F es de clase 1 en U, y se denota $F \in C^1(U, Y)$, si F es G-derivable en todo punto $x_0 \in U$ y la aplicación $x_0 \in U \to \delta F(x_0) \in \mathcal{L}(X, Y)$ es continua.*

Proposición 3.1.15 *Sean X e Y e.n., $U \subset X$ abierto y $F : U \to Y$. Si $F \in C^1(U, Y)$, entonces F es estrictamente derivable en todos los puntos de U (y en consecuencia es derivable Fréchet en todos los puntos de U).*

Demostración: Sean $x_0 \in U$ y $\varepsilon > 0$ fijados. Por hipótesis, existe $\delta > 0$ tal que

$$B_\delta = \bar{B}(x_0, \delta) \subset U, \quad \|\delta F(x) - \delta F(x_0)\|_{\mathcal{L}(X,Y)} \leq \varepsilon \quad \forall\, x \in B_\delta.$$

Si x_1 y x_2 pertenecen a B_δ, entonces $[x_1, x_2] \subset B_\delta$. Aplicando el Teorema del valor medio a $F - \delta F(x_0)$, podemos afirmar la existencia de $\hat{x} \in (x_1, x_2) \subset B_\delta$ tal que

$$\|F(x_2) - F(x_1) - \delta F(x_0)(x_2 - x_1)\|_Y \leq \|\delta F(\hat{x}) - \delta F(x_0)\| \|x_2 - x_1\| \leq \varepsilon \|x_2 - x_1\|.$$

□

3.1.4. Derivación de orden superior

Observación 3.1.16 *Si $F : U \to Y$ es derivable Fréchet en todos los puntos de U, podemos considerar la aplicación F' definida por*

$$F' : x \in U \longmapsto F'(x) \in \mathcal{L}(X, Y).$$

Como $\mathcal{L}(X, Y)$ es un e.n., se puede hablar del concepto de derivada Fréchet de F'; caso de que exista dicha derivada en un punto $x_0 \in U$, la denotaremos $(F')'(x_0)$ y diremos que F es dos veces derivable Fréchet en x_0. Evidentemente, en tal caso $(F')'(x_0) \in \mathcal{L}(X, \mathcal{L}(X, Y))$, y en consecuencia es identificable con una aplicación bilineal continua de $X \times X \to Y$ que denotaremos $F''(x_0)$. Este razonamiento puede ser generalizado a derivadas Fréchet de orden mayor que dos, y da lugar a un Cálculo diferencial para derivadas Fréchet. Nosotros nos vamos a contentar con algo más modesto.

Definición 3.1.17 *Sean X e Y e.n., $U \subset X$ abierto y $F : U \to Y$. Dados $x_0 \in U$ y $h \in X$, se dice que F es dos veces G-diferenciable en x_0 en la dirección h, si existe $\varepsilon_0 > 0$ tal que F es G-diferenciable en la dirección h en todo punto de la forma $x_0 + \varepsilon h$ con $|\varepsilon| < \varepsilon_0$, y existe $\delta^2 F(x_0, h, h)$ definido como el límite*

$$\delta^2 F(x_0, h, h) = \lim_{\varepsilon \to 0} \frac{\delta F(x_0 + \varepsilon h, h) - \delta F(x_0, h)}{\varepsilon}. \tag{3.13}$$

Al elemento $\delta F^2(x_0, h, h) \in Y$ así definido lo denominaremos la G-diferencial segunda de F en el punto x_0 en la dirección h.

Ejemplo Si X, Y son e.n., $a(.,.) : X \times X \to Y$ es una forma bilineal simétrica y definimos J por $J(x) = a(x, x)$, entonces es fácil comprobar que

$$\delta^2 F(x_0, h, h) = 2a(h, h) \quad \forall\, h \in X.$$

Observación 3.1.18 *Está claro que si para ciertos $x_0 \in U$ y $h \in X$ existe $\delta^2 J(x_0, h, h)$, entonces existe $\delta^2 F(x_0, \alpha h, \alpha h)$ para todo $\alpha \in \mathbb{R}$ y se verifica*

$$\delta^2 F(x_0, \alpha h, \alpha h) = \alpha^2 \delta^2 F(x_0, h, h).$$

Observación 3.1.19 *En el caso en que el espacio de llegada es $\mathbb{R}$, caso de que existan, es fácil comprobar a partir de las definiciones que*

$$\delta F(x_0, h) = \frac{d}{d\varepsilon} F(x_0 + \varepsilon h)_{|\varepsilon=0} \qquad \delta^2 F(x_0, h, h) = \frac{d^2}{d\varepsilon^2} F(x_0 + \varepsilon h)_{|\varepsilon=0}.$$

Utilizando esta observación, la fórmula de Taylor para funciones reales de variable real y la existencia, por el Corolario 1.5.4, de un elemento $y' \in Y'$ con $\|y'\|_{Y'} = 1$ alineado con un elemento dado en Y, es fácil demostrar el siguiente resultado

Teorema 3.1.20 *Sean X, Y e.n., $U \subset X$ abierto y $F : U \to Y$. Supongamos dados $x_1, x_2 \in U$ tales que $[x_1, x_2] \subset U$, y F es dos veces G-diferenciable en todo punto de $[x_1, x_2]$ en la dirección $x_2 - x_1$. Entonces*

1. *Si $Y = \mathbb{R}$, existe $x_0 \in (x_1, x_2)$ tal que*

$$F(x_2) = F(x_1) + \delta F(x_1, x_2 - x_1) + \frac{1}{2}\delta^2 F(x_0, x_2 - x_1, x_2 - x_1). \tag{3.14}$$

2. *Si Y no es necesariamente $\mathbb{R}$, existe $x_0 \in (x_1, x_2)$ tal que*

$$\|F(x_2) - F(x_1) - \delta F(x_1, x_2 - x_1)\|_Y \leq \frac{1}{2}\|\delta^2 F(x_0, x_2 - x_1, x_2 - x_1)\|_Y. \tag{3.15}$$

3.2. Derivación y convexidad

Vamos ahora a mostrar como se puede aplicar la G-diferenciabilidad para estudiar la convexidad de una función

Proposición 3.2.1 *Sea X un e.n., $U \subset X$ abierto, convexo y $F : U \to \mathbb{R}$. Supongamos que para todo par de puntos x_1 y x_2 del abierto U existe la G-diferencial $\delta F(x_1, x_2 - x_1)$. Entonces se tiene*

1. *F es convexa si y sólo si se verifica*

$$F(x_2) - F(x_1) \geq \delta F(x_1, x_2 - x_1) \quad \forall\, x_1, x_2 \in U. \tag{3.16}$$

2. *F es estrictamente convexa si y sólo si se verifica*

$$F(x_2) - F(x_1) > \delta F(x_1, x_2 - x_1) \quad \forall\, x_1, x_2 \in U,\ x_1 \neq x_2. \tag{3.17}$$

Demostración: Supongamos primero F convexa, entonces dado $\varepsilon \in (0,1)$, se tiene

$$\frac{F(x_1+\varepsilon(x_2-x_1))-F(x_1)}{\varepsilon} \leq \frac{\varepsilon F(x_2)+(1-\varepsilon)F(x_1)-F(x_1)}{\varepsilon} = F(x_2)-F(x_1).$$

La demostración de (3.16) se sigue entonces sin más que tomar el límite cuando ε tiende a cero.

Recíprocamente, si se satisface (3.16), entonces para todo par $x_1, x_2 \in U$ y todo $\alpha \in [0,1]$ se tiene:

$$F(x_1) \geq F(x_1+\alpha(x_2-x_1)) + \delta F(x_1+\alpha(x_2-x_1), -\alpha(x_2-x_1))$$

y

$$F(x_2) \geq F(x_1+\alpha(x_2-x_1)) + \delta F(x_1+\alpha(x_2-x_1), (1-\alpha)(x_2-x_1)).$$

Por la homogeneidad de la G-diferencial en la variable h, multiplicando la primera de las dos desigualdades por $1-\alpha$, la segunda por α y sumando, se prueba

$$F(\alpha x_1 + (1-\alpha)x_2) \leq \alpha F(x_1) + (1-\alpha)F(x_2)$$

y por tanto F es convexa. El argumento anterior sirve también para probar que (3.17) implica la convexidad estricta de F.

Para probar el recíproco, supongamos F estrictamente convexa. Como en particular F es convexa, se tendrá por tanto (3.16). Si ahora $x_1, x_2 \in U$ $x_1 \neq x_2$ y $\alpha \in (0,1)$ entonces, por el carácter estrictamente convexo de F se tiene

$$F(x_2)-F(x_1) > \frac{F(x_1+\alpha(x_2-x_1))-F(x_1)}{\alpha} \geq \frac{\delta F(x_1, \alpha(x_2-x_1))}{\alpha} = \delta F(x_1, x_2-x_1).$$

□

Proposición 3.2.2 *En las hipótesis de la Proposición 3.2.1 con F G-diferenciable en $[x_1, x_2]$ y en la dirección $x_2 - x_1$. F es convexa si y sólo si*

$$\delta F(x_2, x_2-x_1) - \delta F(x_1, x_2-x_1) \geq 0, \quad \forall\, x_1, x_2 \in U. \tag{3.18}$$

Demostración: De (3.16) se tiene

$$F(x_2)-F(x_1) \geq \delta F(x_1, x_2-x_1)$$

$$F(x_1)-F(x_2) \geq \delta F(x_2, x_1-x_2).$$

Sumando estas desigualdades y teniendo en cuenta la homogeneidad de la G-diferencial en la dirección se obtiene (3.18).

Sean $x_1, x_2 \in U$ y consideremos la función $\varphi(t) = F(x_1+t(x_2-x_1))$. Por las hipótesis, φ es derivable en $[0,1]$ y $\varphi'(t) = \delta F(x_1+t(x_2-x_1), x_2-x_1)$.

Sea ahora $0 \leq s < t \leq 1$ y llevamos $x_s = x_1 + s(x_2-x_1)$ y $x_t = x_1 + t(x_2-x_1)$ a (3.18) en lugar de x_1 y x_2. Se obtiene:

$$\delta F(x_t, (t-s)(x_2-x_1)) \geq \delta F(x_s, (t-s)(x_2-x_1)),$$

dividiendo por $t-s$ se deduce $\varphi'(t) \geq \varphi'(s)$, i.e. φ' es creciente luego, por cálculo elemental, φ es convexa y por tanto $\varphi(\alpha) \leq \alpha\varphi(1) + (1-\alpha)\varphi(0)$ i.e. F es convexa. □

Proposición 3.2.3 *Sean X un e.n., $U \subset X$ abierto, convexo y $F : U \to \mathbb{R}$. Supongamos que existe la G-diferencial segunda $\delta^2 F(x_1, x_2 - x_1, x_2 - x_1)$ para todo par de puntos x_1, x_2 de U, entonces se tiene*

1. *F es convexa si y sólo si se verifica*

$$\delta^2 F(x_1, x_2 - x_1, x_2 - x_1) \geq 0, \quad \forall\, x_1, x_2 \in U. \tag{3.19}$$

2. *Si se verifica*

$$\delta^2 F(x_1, x_2 - x_1, x_2 - x_1) > 0, \quad \forall\, x_1, x_2 \in U. \tag{3.20}$$

entonces F es estrictamente convexa.

Demostración: Gracias a la Proposición 3.2.1, a (3.14) y a

$$\delta^2 F(\xi, x_2 - x_1, x_2 - x_1) = \delta^2 F(\xi, -\frac{\|x_2 - x_1\|_X}{\|\xi - x_1\|_X}(x_1 - \xi), -\frac{\|x_2 - x_1\|_X}{\|\xi - x_1\|_X}(x_1 - \xi)) =$$

$$= \frac{\|x_2 - x_1\|_X^2}{\|\xi - x_1\|_X^2} \delta^2 F(\xi, x_2 - \xi, x_2 - \xi),$$

está claro que (3.19) y (3.20) implican la convexidad y la convexidad estricta respectivamente de F.

Para terminar la demostración basta probar que si F es convexa entonces se verifica (3.19). Para ello sean $x_1, x_2 \in U$ con $x_1 \neq x_2$ (el caso $x_1 = x_2$ es trivial). Como F es convexa entonces se verifica (3.18) luego reemplazamos x_2 por $x_1 + t(x_2 - x_1)$, para cada $t \in (0,1)$, y se tiene

$$\delta F(x_1 + t(x_2 - x_1), t(x_2 - x_1)) - \delta F(x_1, t(x_2 - x_1)) \geq 0.$$

Dividiendo por t^2, teniendo en cuenta la homogeneidad de la G-diferencial en la dirección, y haciendo tender t a cero se obtiene (3.19). $\square$

Observación 3.2.4 *Si F es estrictamente convexa, no tiene por qué verificarse (3.20), basta por ejemplo con considerar la aplicación $F : \mathbb{R} \to \mathbb{R}$ definida por $F(t) = t^4$ para todo $t \in \mathbb{R}$, la cual es estrictamente convexa y sin embargo $F''(0) = 0$.*

Ejemplo Si X es un e.n., $a(.,.) : X \times X \to \mathbb{R}$ es una forma bilineal simétrica y definimos J por $J(x) = a(x,x)$, usando el resultado anterior se tiene que J es convexa si y sólo si es semidefinida positiva y que si a es definida positiva entonces J es estrictamente convexa. De hecho en este caso, se tiene que J es estrictamente convexa si y sólo si a es definida positiva ya que si existe $h \neq 0$ con $a(h,h) = 0$, entonces para todo $\alpha \in (0,1)$ se tiene

$$J(\alpha 0 + (1-\alpha)h) = (1-\alpha)^2 a(h,h) = 0 = \alpha J(0) + (1-\alpha)J(h).$$

3.3. Teoremas de la función inversa y de la función implícita

Para terminar con esta introducción a la teoría de derivación en e.n. veamos ahora un Teorema de la función inversa de gran generalidad, que nos servirá más adelante para obtener una generalización del Teorema de los multiplicadores de J.L. Lagrange en el caso de espacios de Banach

Teorema 3.3.1 *(Teorema de L.A. Liusternik de la función inversa) Sean X, Y dos espacios de Banach, $U \subset X$ abierto, $x_0 \in U$, y $F : U \to Y$ tal que $F \in SD^1(x_0)$ y $F'(x_0)$ sea sobreyectiva. Entonces existen dos números $r_0 > 0$ y $K > 0$ tales que para todo $y \in \overset{\circ}{B}(F(x_0), r_0))$ existe $x \in U$ con*

$$F(x) = y, \quad \|x - x_0\|_X \leq K\|y - F(x_0)\|_Y \tag{3.21}$$

Demostración: Suponemos, sin pérdida de generalidad, que $x_0 = 0$ y $F(x_0) = 0$ (bastaría con reemplazar F por la función G definida por $G(x) = F(x_0 + x) - F(x_0)$). Para simplificar la notación, denotaremos $A = F'(0)$. Como A es sobreyectiva, por el Teorema 1.5.8, del inverso a la derecha de Banach, sabemos que existe $C > 0$ y una aplicación $M : Y \to X$ tal que para todo $y \in Y$ se tiene

$$A(M(y)) = y, \quad \|M(y)\|_X \leq C\|y\|_X,$$

(se dice que M es un inverso a la derecha de A). Como por hipótesis $F \in SD^1(0)$, podemos afirmar la existencia de un $r_1 > 0$ tal que $\bar{B}(0, r_1) \subset U$, y

$$\|F(x_1) - F(x_2) - A(x_1 - x_2)\|_Y \leq \frac{1}{2C}\|x_1 - x_2\|_X \quad \forall\, x_1,\, x_2 \in \bar{B}(0, r_1). \tag{3.22}$$

Tomemos

$$r_0 = \frac{r_1}{2C}.$$

Fijado $y \in \overset{\circ}{B}(0,\, r_0)$ definamos

$$x_0 = 0, \quad x_{n+1} = x_n + M(y - F(x_n)) \quad \forall\, n \geq 0.$$

En primer lugar, vamos a demostrar por inducción que, con esta fórmula, los elementos $x_n \in X$ que se obtienen verifican

$$\|x_n\|_X \leq 2C\|y\|_Y, \quad \forall\, n \in \mathbb{N} \tag{3.23}$$

con lo que en particular, $x_n \in B(0, r_1) \subset U$. De ahí que tenga sentido $F(x_n)$ y por tanto que exista x_{n+1}, es decir la sucesión $\{x_n\}$ estará bien definida.

Vamos a probar la condición (3.23) por inducción. Claramente, la condición se verifica para x_0. Para x_1, se tiene también $x_1 = M(y)$ y por tanto $\|x_1\|_X \leq C\|y\|_Y \leq 2C\|y\|_Y$. Sea entonces $n \geq 1$ y supongamos $\|x_k\|_X \leq 2C\|y\|_Y$ para $k = 0, 1, ..., n$. En tal caso se tiene

$$x_{n+1} - x_n = M(y - F(x_n)) \Rightarrow \|x_{n+1} - x_n\|_X \leq C\|y - F(x_n)\|_Y$$

y

$$x_n - x_{n-1} = M(y - F(x_{n-1})) \Rightarrow y = F(x_{n-1}) - A(x_{n-1} - x_n),$$

con lo que, por (3.22), se obtiene

$$\|x_{n+1} - x_n\|_X \leq C\|F(x_{n-1}) - F(x_n) - A(x_{n-1} - x_n)\|_Y \leq \frac{1}{2}\|x_n - x_{n-1}\|_X.$$

En consecuencia $\|x_{n+1} - x_n\|_X \leq \frac{1}{2^n}\|x_1\|_Y$, y por tanto

$$\|x_{n+1}\|_X \leq \left(\frac{1}{2^n} + \frac{1}{2^{n-1}} + ... + 1\right)\|x_1\|_X \leq 2C\|y\|_X. \tag{3.24}$$

Además, del razonamiento precedente, se tiene que si $m > n$ entonces

$$\|x_m - x_n\|_X \leq \left(\frac{1}{2^{m-1}} + ... + \frac{1}{2^n} \right) \|x_1\|_X \to 0$$

si m y n tienden a infinito. En consecuencia, la sucesión $\{x_n\}$ así construida es de Cauchy y, como X es completo, podemos afirmar que existe $x \in X$ tal que

$$\lim_{n\to\infty} x_n = x, \quad \text{con } \|x\|_X \leq 2C\|y\|_Y.$$

Finalmente, de la igualdad $x_{n+1} = x_n + M(y - F(x_n))$ se tiene

$$Ax_{n+1} = Ax_n + y - F(x_n),$$

con lo que

$$y = \lim_{n\to\infty} F(x_n),$$

pero como $\|x\|_X \leq 2C\|y\|_Y \leq r_1$, se deduce de (3.22)

$$\|F(x_n) - F(x)\|_Y \leq \|F(x_n) - F(x) - A(x_n - x)\|_Y + \|A\|_{\mathcal{L}(X,Y)}\|x_n - x\|_X \leq$$

$$\leq \left(\frac{1}{2C} + \|A\|_{\mathcal{L}(X,Y)} \right) \|x_n - x\|_X,$$

con lo que

$$\lim_{n\to\infty} F(x_n) = F(x),$$

lo que junto con (3.23) prueba el Teorema. □

En el Teorema precedente, a causa de la debilidad en las hipótesis, la afirmación final que se obtiene dista bastante del Teorema clásico de la función inversa. Bajo hipótesis más fuertes y usando el Teorema de Liusternik se puede obtener el siguiente teorema que damos sin demostración.

Teorema 3.3.2 *Sean X, Y dos espacios de Banach, $U \subset X$ abierto, $x_0 \in U$ y $F \in C^1(U,Y)$ tal que $F'(x_0)$ sea biyectiva. Entonces existe $\varepsilon > 0$ tal que $B(x_0,\varepsilon) \subset U$, el conjunto $V = F\big(B(y_0,\varepsilon)\big)$ es un entorno abierto de $F(x_0)$, y se tiene*

1. *F es biyectiva de $B(x_0,\varepsilon)$ en V.*
2. *La aplicación $F^{-1} : V \to B(x_0,\varepsilon)$ es de clase 1 y verifica*

$$(F^{-1})'(y) = \big(F'(F^{-1}(y))\big)^{-1}, \quad \forall\, y \in V.$$

Se puede también probar la siguiente versión del Teorema de la función implícita

Teorema 3.3.3 *Sean X, Y, Z tres espacios de Banach, con $U \subset X \times Y$ abierto, $(x_0,y_0) \in U$. Supongamos $F \in C^1(U,Z)$ tal que la derivada de la aplicación $F(x_0,\cdot) : Y \to Z$ (derivada parcial) en y_0 sea biyectiva de Y sobre Z, y sea $z_0 = F(x_0,y_0)$, entonces existen $r_0 > 0$, $r_1 > 0$ y $\varphi \in C^1\big(B(x_0,r_0), B(y_0,r_1)\big)$ tales que*

1. *$\varphi(x_0) = y_0$,*
2. *$F(x,\varphi(x)) = z_0$, para todo $x \in B(x_0,r_0)$,*
3. *Si $(x,y) \in B(x_0,r_0) \times B(y_0,r_1)$ y $F(x,y) = z_0$ entonces $y = \varphi(x)$.*

3.4. Condiciones necesarias de extremo no condicionado.

Usando la derivación en e.n., vamos ahora a obtener algunas condiciones necesarias y suficientes que debe verificar un punto en el cual se alcanza el mínimo de una función. Comenzamos recordando parte de la siguiente definición 2.1.1.

Definición 3.4.1 *Sean X un e.n. y $\mathcal{U}$ un subconjunto de X. Diremos que en $\hat{x} \in \mathcal{U}$ se alcanza un mínimo local de f en $\mathcal{U}$ si existe $\varepsilon > 0$ tal que*

$$f(\hat{x}) \leq f(x) \quad \forall\, x \in \mathcal{U} \cap B(\hat{x}, \varepsilon).$$

Evidentemente todo mínimo global de f en $\mathcal{U}$, es mínimo local. El siguiente resultado muestra que de hecho, en el caso de funcionales convexos ambos conceptos de mínimo son equivalentes

Proposición 3.4.2 *Sean X un e.n., $\mathcal{U} \subset X$ convexo y $f : \mathcal{U} \to \mathbb{R}$ convexa. En estas condiciones f alcanza mínimo global en $\hat{x} \in \mathcal{U}$ si y sólo si alcanza mínimo local en $\hat{x}$.*

Demostración: Supongamos que f alcanza mínimo local en $\hat{x} \in \mathcal{U}$, y consideremos fijado $\varepsilon > 0$ tal que $f(\hat{x}) \leq f(x) \quad \forall\, x \in \mathcal{U} \cap B(\hat{x}, \varepsilon)$.

Por la convexidad de $\mathcal{U}$, dado $x \in \mathcal{U}$, se tiene que para todo $\alpha \in (0, 1)$, $\hat{x} + \alpha(x - \hat{x})$ pertenece a $\mathcal{U}$. Además, si α es suficientemente pequeño, se tiene $\|\alpha(x - \hat{x})\| < \varepsilon$ y por tanto $\hat{x} + \alpha(x - \hat{x})$ pertenece a $B(\hat{x}, \varepsilon)$. Para dicho α se tendrá por tanto

$$f(\hat{x}) \leq f(\hat{x} + \alpha(x - \hat{x})) \leq (1 - \alpha) f(\hat{x}) + \alpha f(x),$$

lo que implica $f(\hat{x}) \leq f(x)$ para todo $x \in \mathcal{U}$. □

Si f no es convexa, el resultado anterior es en general falso, sin embargo como todo mínimo global también lo es local, un paso para encontrar los mínimos globales consiste en buscar los mínimos locales. De manera análoga al caso de funciones reales de variable real, se tiene

Teorema 3.4.3 *Sean X un e.n., $\mathcal{U} \subset X$ con $\mathring{\mathcal{U}} \neq \emptyset$ y $f : \mathcal{U} \to \mathbb{R}$ dados. En estas condiciones se tiene*

a) Si f alcanza un mínimo local en $\hat{x} \in \mathring{\mathcal{U}}$ y f es G-diferenciable en $\hat{x}$, entonces

$$\delta f(\hat{x}, h) = 0, \quad \forall\, h \in X. \tag{3.25}$$

b) Si $\mathcal{U}$ es convexo y f es convexa en $\mathcal{U}$ y G-diferenciable en un punto $\hat{x} \in \mathring{\mathcal{U}}$, entonces f alcanca un mínimo global en $\hat{x}$ si y sólo si satisface (3.25).

Demostración: Para probar a), consideremos $h \in X$ fijado. Sabemos que existe $\varepsilon_0 > 0$ tal que para todo ε con $|\varepsilon| < \varepsilon_0$ se satisface $f(\hat{x}) \leq f(\hat{x} + \varepsilon h)$, con lo cual

$$f(\hat{x} + \varepsilon h) - f(\hat{x}) \geq 0 \quad \forall\, \varepsilon \text{ tal que } 0 < |\varepsilon| < \varepsilon_0.$$

Dividiendo por ε en la desigualdad anterior y tomando límite cuando $\varepsilon \to 0$ se prueba $\delta f(\hat{x}, h) \geq 0$ y $\delta f(\hat{x}, h) \leq 0$, con lo cual se deduce (3.25).

Para probar el apartado b), basta observar que si $\hat{x} \in \mathcal{U}$ satisface (3.25) entonces gracias a (3.16) se tiene

$$f(x) \geq f(\hat{x}) + \delta f(\hat{x}, x - \hat{x}) = f(\hat{x}) \quad \forall\, x \in \mathcal{U}.$$

□

Observando la demostración anterior, introducimos la definición

Definición 3.4.4 *Sea $x \in \mathcal{U}$, un $h \in X$ no nulo, se dice que es una dirección admisible en x para el conjunto $\mathcal{U}$ si existe $\varepsilon_0 > 0$ tal que para todo ε con $|\varepsilon| < \varepsilon_0$ se satisface $x + \varepsilon h \in \mathcal{U}$.*

Al conjunto de direcciones admisibles lo denotamos por $D_{ad}(x,\mathcal{U})$.

Entonces se tiene:

Corolario 3.4.5 *Sean X un e.n., $\mathcal{U} \subset X$ y $f : \mathcal{U} \to \mathbb{R}$ dados. Si f alcanza un mínimo local en $\hat{x} \in \mathcal{U}$ y f es G-diferenciable en $\hat{x}$ y en cualquier dirección de $D_{ad}(\hat{x},\mathcal{U})$, entonces*

$$\delta f(\hat{x}, h) = 0, \quad \forall\, h \in D_{ad}(\hat{x},\mathcal{U}). \tag{3.26}$$

En el caso $\mathcal{U}$ convexo, el resultado anterior se puede mejorar de la siguiente forma.

Supongamos X un espacio vectorial (no necesariamente normado), $\mathcal{U} \subset X$ un subconjunto convexo no vacío de X y $f : \mathcal{U} \to \mathbb{R}$ una función dada. Bajo estas condiciones, fijados dos puntos $\hat{x}$ y x en $\mathcal{U}$, sabemos que para todo $\varepsilon \in (0, 1]$ se satisface que $\hat{x} + \varepsilon(x - \hat{x})$ pertenece a $\mathcal{U}$, y en consecuencia está definido el cociente

$$\frac{f(\hat{x} + \varepsilon(x - \hat{x})) - f(\hat{x})}{\varepsilon}.$$

Por tanto tiene sentido la siguiente definición.

Definición 3.4.6 *En las condiciones anteriores, si existe*

$$\lim_{\varepsilon \to 0^+} \frac{f(\hat{x} + \varepsilon(x - \hat{x})) - f(\hat{x})}{\varepsilon},$$

se dice que f es G-diferenciable por la derecha en $\hat{x}$, en la dirección $x - \hat{x}$, y se denota al límite anterior por $\delta^+ f(\hat{x}, x - \hat{x})$.

Con esta definición, razonando como en el Teorema 3.4.3 es fácil probar el siguiente resultado

Teorema 3.4.7 *Sean X un espacio vectorial, $\mathcal{U} \subset X$ convexo no vacío y $f : \mathcal{U} \to \mathbb{R}$. Bajo estas condiciones, se tiene*

a) Si f alcanza un mínimo local en $\hat{x} \in \mathcal{U}$ y f es G-diferenciable por la derecha en $\hat{x}$, entonces

$$\delta^+ f(\hat{x}, x - \hat{x}) \geq 0 \quad \forall\, x \in \mathcal{U}. \tag{3.27}$$

b) Si f es convexa y G-diferenciable por la derecha en $\hat{x} \in \mathcal{U}$, entonces f alcanza un mínimo global en $\hat{x}$ si y sólo si satisface (3.27).

Observación 3.4.8 *Se puede introducir de forma natural el concepto de dirección admisible lateral en x para el conjunto $\mathcal{U}$ y generalizar la parte a) del anterior Teorema en el sentido del Corolario 3.4.5.*

En el caso de funciones dos veces diferenciables se tiene también la siguiente condición suficiente de mínimo local.

Teorema 3.4.9 *Sean X un e.n., $\mathcal{U} \subset X$, y $f : \mathcal{U} \to \mathbb{R}$ dada. Supongamos que f alcanza un mínimo local en $\hat{x} \in \mathring{\mathcal{U}}$, y que f es dos veces G-diferenciable en $\hat{x}$ en la dirección h. Entonces, se verifica*

$$\delta^2 f(\hat{x}, h, h) \geq 0 \tag{3.28}$$

(y por (3.25), $\delta f(\hat{x}, h) = 0$).

Demostración: Basta considerar la función φ definida en un entorno de $0 \in \mathbb{R}$ por $\varphi(t) = f(\hat{x}+th)$ y observar que $\varphi'(0) = \delta f(\hat{x}, h)$, $\varphi''(0) = \delta^2 f(\hat{x}, h, h)$ y que φ alcanza mínimo local en $t = 0$. $\square$

Observación 3.4.10 *Análogamente al caso de funciones de variable real, se pueden también conseguir condiciones suficientes de mínimo local usando derivadas segundas, pero para ello hay que considerar la derivada segunda en el sentido Fréchet, la cual no hemos estudiado.*

3.5. Problemas con restricciones

3.5.1. Multiplicadores de Lagrange

En muchas situaciones, el conjunto $\mathcal{U}$ de las restricciones de un problema de mínimos se escribe de la forma

$$\mathcal{U} = \{x \in O \,/\, G(x) = 0\},$$

con $O \subset X$ un conjunto abierto que se conoce de manera explícita, y G una aplicación definida sobre O con valores en un e.n. Y, aplicación que define de manera implícita el resto de las restricciones del problema. En tal caso escribimos el problema en la forma

$$\text{ínf}\{f(x) \,/\, x \in O, \quad G(x) = 0\} \tag{3.29}$$

En el caso en que los espacios X y Y son de dimensión finita, es bien conocido que se verifica el Teorema de los multiplicadores de Lagrange. En el presente tema vamos a ver una generalización de este resultado al caso de e.n.. Se tiene

Teorema 3.5.1 *(Teorema de multiplicadores de Lagrange) Sean X, Y dos espacios de Banach, $O \subset X$ un conjunto abierto, $f : O \to \mathbb{R}$, $G : O \to Y$. Denotemos*

$$\mathcal{U} = \{x \in O \,/\, G(x) = 0\}$$

y supongamos que en $\hat{x} \in \mathcal{U}$ se alcanza un mínimo local de f en $\mathcal{U}$. Si f y G son estrictamente derivables en $\hat{x}$ y $\mathcal{R}(G'(\hat{x})) = G'(\hat{x})(X)$ es cerrado en Y, entonces existen $\lambda_0 \in \mathbb{R}$, $y' \in Y'$ tales que $(\lambda_0, y') \neq (0, 0)$ y

$$\lambda_0 f'(\hat{x}) + y' \circ G'(\hat{x}) = 0, \quad \text{en } X' \tag{3.30}$$

o lo que es equivalente

$$\lambda_0 \langle f'(\hat{x}), h\rangle_{X',X} + \langle y', G'(\hat{x})h\rangle_{Y',Y} = 0, \quad \forall\, h \in X. \tag{3.31}$$

Además, si $G'(\hat{x})$ es sobreyectiva, entonces $\lambda_0 \neq 0$ y se puede en consecuencia tomar $\lambda_0 = 1$.

Demostración: En el caso en que $\mathcal{R}(G'(\hat{x}))$ está estrictamente contenido en Y, la demostración es bastante simple. Como por hipótesis este espacio es cerrado, sabemos que existe $y' \in \left[\mathcal{R}(G'(\hat{x}))\right]^{\perp}$, no nulo. Es decir y' satisface

$$\langle y', G'(\hat{x})h\rangle_{Y',Y} = 0, \quad \forall\, h \in X.$$

Tomando entonces $\lambda_0 = 0$ se tiene (3.30).

Consideremos ahora el caso en que $G'(\hat{x})$ es sobreyectivo. Vamos a probar que en este caso se verifica la siguiente condición (que en algunos casos puede ser importante por sí misma)

$$\mathcal{N}\big(G'(\hat{x})\big) \subset \mathcal{N}\big(f'(\hat{x})\big). \tag{3.32}$$

Razonamos por reducción al absurdo. Supongamos $h_0 \in \mathcal{N}\big(G'(\hat{x})\big)$, el cual no pertenece a $\mathcal{N}\big(f'(\hat{x})\big)$. Sea entonces $H : O \to \mathbb{R} \times Y$ definida por $H(x) = (f(x), G(x))$, para todo $x \in O$. Fácilmente se puede ver que H es estrictamente derivable en $\hat{x}$, y su derivada viene dada por $H'(\hat{x}) = (f'(\hat{x}), G'(\hat{x}))$. Además se puede probar que H' es sobreyectiva, para ello sea $(t, k) \in \mathbb{R} \times Y$ arbitrario. Como $G'(\hat{x}$ es sobreyectivo, existe $h \in X$ tal que $G'(\hat{x})h = k$. Un simple cálculo prueba entonces

$$H'(\hat{x}) \left(h + \frac{t - \langle f'(\hat{x}), h \rangle_{X',X}}{\langle f'(\hat{x}), h_0 \rangle_{X',X}} h_0 \right) = (t, k)$$

y por tanto H es sobreyectiva. Al ser $H(\hat{x}) = (f(\hat{x}), 0)$, podemos afirmar, por el Teorema 3.3.1, de la función inversa de Liusternik, que existen dos números $r_0 > 0$ y $K > 0$ tales que para todo $\varepsilon \in (0, r_0)$ existe $x_\varepsilon \in O$ verificando

$$H(x_\varepsilon) = (f(\hat{x}) - \varepsilon, 0), \quad \|x_\varepsilon - \hat{x}\|_X \le K\varepsilon.$$

En consecuencia, en todo entorno de $\hat{x}$ existe un punto de $\mathcal{U}$ donde f toma un valor estrictamente menor que $f(\hat{x})$, en contradicción con el hecho de que en $\hat{x}$ se alcanza un mínimo local de f en $\mathcal{U}$. Esto prueba (3.32).

Sea ahora $y' : Y \to \mathbb{R}$ definida por

$$y'(k) = -\langle f'(\hat{x}), h \rangle_{X',X}, \quad \forall\, h \in X \text{ tal que } G'(\hat{x})h = k, \ \forall\, k \in Y.$$

La aplicación y' está bien definida debido a que $G'(\hat{x})$ es sobreyectiva y al hecho de que si $h_1, h_2 \in X$ son tales que $G'(\hat{x})h_1 = G'(\hat{x})h_2 = k$, para $k \in X$. Entonces $h_1 - h_2$ pertenece al núcleo de $G'(\hat{x})$ y por tanto, gracias a (3.32) se tiene $-\langle f'(\hat{x}), h_1 \rangle_{X',X} = -\langle f'(\hat{x}), h_2 \rangle_{X',X}$. Fácilmente se prueba que y' es lineal. Veamos también que es continua y por tanto un elemento de Y'. Para ello usamos que por el Teorema 1.5.8 del inverso a la derecha de Banach, existe $C > 0$ verificando que para todo $k \in Y$ existe $h \in X$ con $G'(\hat{x})h = k$, $\|h\|_X \le C\|k\|_Y$. Tomando este h se tiene entonces

$$|y'(k)| = |\langle f'(\hat{x}), h \rangle_{X',X}| \le \|f'(\hat{x})\|_{X'} \|h\|_X \le \|f'(\hat{x})\|_{X'} C \|k\|_Y,$$

lo que al ser y' lineal implica su continuidad. Teniendo en cuenta que por definición y' satisface $y' \circ G'(\hat{x}) = -f'(\hat{x})$ se deduce ahora (3.30) con $\lambda_0 = 1$. □

Observación 3.5.2 *La condición (3.30) o (3.31) del Teorema se obtiene introduciendo el lagrangiano asociado definido por*

$$\mathcal{L}(x; \lambda_0, y') = \lambda_0 f(x) + \langle y', G(x) \rangle_{Y',Y}$$

e igualando a cero la derivada de éste respecto a x y manteniendo fijo (λ_0, y').

En las aplicaciones que veremos en el curso (especialmente en el caso de problemas del Cálculo de Variaciones), aunque el espacio X es de dimensión infinita, el espacio Y es finito dimensional. Teniendo en cuenta que en $\mathbb{R}^n$ todo subespacio vectorial es cerrado así como el hecho de que el dual de $\mathbb{R}^n$ se identifica con el propio $\mathbb{R}^n$ es inmediato entonces el siguiente resultado.

Corolario 3.5.3 *Sean X un espacio de Banach, $O \subset X$ un conjunto abierto, $\hat{x} \in O$, $f : O \to \mathbb{R}$, $g_1, \cdots, g_m : O \to \mathbb{R}$ y supongamos que $f, g_1, \cdots, g_n$ son estrictamente derivables en $\hat{x}$. Denotemos*

$$\mathcal{U} = \{x \in O \,/\, g_i(x) = 0,\ 1 \leq i \leq n\}.$$

Si en $\hat{x}$ se alcanza un mínimo local de f en $\mathcal{U}$, entonces existen $\lambda_0, \lambda_1, \cdots, \lambda_n \in \mathbb{R}$ no todos nulos tales que

$$\lambda_0 f'(\hat{x}) + \sum_{i=1}^{n} \lambda_i g_i'(\hat{x}) = 0. \tag{3.33}$$

A los números $\lambda_0, \lambda_1, ..., \lambda_n$ se le denominan los multiplicadores de Lagrange del problema. Si los vectores $g_i'(\hat{x})$ son linealmente independientes, entonces λ_0 es no nulo y se puede tomar $\lambda_0 = 1$.

Observación 3.5.4 *Como en la Observación 3.5.2 la condición (3.33) se obtiene introduciendo el lagrangiano asociado definido por*

$$\mathcal{L}(x; \lambda_0, \lambda_1, \cdots, \lambda_n) = \lambda_0 f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$$

e igualando a cero la derivada de éste respecto a x y manteniendo fijo λ_i.

3.5.2. Teorema de H.W. Kuhn- A.W. Tucker

En el Teorema 3.5.1, de los multiplicadores de Lagrange, hemos considerado restricciones de tipo igualdad. En numerosas ocasiones hay que considerar restricciones de tipo desigualdad. En el caso convexo, un resultado de gran interés en esta dirección es el teorema siguiente.

Teorema 3.5.5 *(Teorema de Kuhn-Tucker). Sean X un espacio vectorial, $K \subset X$ convexo, $f : K \to \mathbb{R}$ convexa, $g_1, \cdots, g_n : K \to \mathbb{R}$ convexas, $h_1, \cdots, h_m : X \to \mathbb{R}$ afines. Definiendo $\mathcal{U}$ por*

$$\mathcal{U} = \big\{x \in K \,/\, g_i(x) \leq 0,\ 1 \leq i \leq n,\ h_j(x) = 0,\ 1 \leq j \leq m\big\},$$

y suponemos que existe $\hat{x} \in \mathcal{U}$, solución del problema

$$\min\{f(x) \,/\, x \in \mathcal{U}\}. \tag{3.34}$$

Entonces, existen $\lambda_i \geq 0$, $0 \leq i \leq n$, $\mu_j \in \mathbb{R}$, $1 \leq j \leq m$, no todos nulos, tales que $\lambda_i g_i(\hat{x}) = 0$, $1 \leq i \leq n$ y $\hat{x}$ es solución del problema

$$\min\{\mathcal{L}(x; \lambda_0, \lambda_1, \cdots, \lambda_m, \mu_1, \cdots, \mu_m) \,/\, x \in K\}. \tag{3.35}$$

con $\mathcal{L}(x; \lambda_0, \lambda_1, \cdots, \lambda_m, \mu_1, \cdots, \mu_m) = \lambda_0 f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{j=1}^{m} \mu_j h_j(x)$.

Supongamos además que se verifica la condición siguiente (condición de Slater):

Todas las restricciones son de desigualdad (es decir no aparecen las funciones h_j) y para todo $i \in \{1, \cdots, n\}$ existe $x_{i_0} \in \mathcal{U}$ verificando $g_i(x_{i_0}) < 0$,. Entonces $\lambda_0 \neq 0$.

Recíprocamente, supongamos que existen $\hat{x} \in \mathcal{U}$, $\lambda_1, \cdots, \lambda_n \geq 0$, $\mu_1, \cdots, \mu_n \in \mathbb{R}$ tales que $\lambda_i g_i(\hat{x}) = 0$, $1 \leq i \leq n$ y $\hat{x}$ es solución del problema

$$\min\{\mathcal{L}(x; 1, \lambda_1, \cdots, \lambda_m, \mu_1, \cdots, \mu_m) \,/\, x \in K\}.$$

Entonces $\hat{x}$ es solución de (3.34).

Demostración: La demostración usa el siguiente resultado de separación de convexos que es una consecuencia inmediata del Corolario 1.5.6. Dado $C \subset \mathbb{R}^N$ convexo, $(\hat{\tau}_1, \cdots, \hat{\tau}_N) \in \mathbb{R}^N \setminus C$, existe $(\lambda_1, \cdots, \lambda_N) \in \mathbb{R}^N \setminus \{0\}$, tal que

$$\sum_{i=1}^N \lambda_i \hat{\tau}_i \leq \sum_{i=1}^N \lambda_i \tau_i, \quad \forall (\tau_1, \cdots, \tau_N) \in C.$$

a) Comenzamos suponiendo $g_i(\hat{x}) = 0$, $i \in \{1, \cdots, n\}$. Definimos entonces $C \subset \mathbb{R}^{n+m+1}$ por

$$C = \{(\tau_0, \cdots, \tau_{n+m}) \in \mathbb{R}^{n+m+1} \,/\, \exists\, x \in K \text{ con}$$

$$f(x) < \tau_0,\ g_i(x) \leq \tau_i,\ 1 \leq i \leq n,\ h_j(x) = \tau_{n+j},\ 1 \leq j \leq m\}.$$

Fácilmente se prueba que este conjunto es convexo. Además gracias a que $\hat{x}$ es solución de (3.34), se deduce que el punto

$$(f(\hat{x}), g_1(\hat{x}), \cdots, g_n(\hat{x}), h_1(\hat{x}), \cdots, h_m(\hat{x})) = (f(\hat{x}), 0, \cdots, 0)$$

no pertenece a C. Existen por tanto $\lambda_0, \cdots, \lambda_n, \mu_1, \cdots, \mu_n \in \mathbb{R}$ no todos nulos tales que

$$\lambda_0 f(\hat{x}) + \sum_{i=1}^n \lambda_i g_i(\hat{x}) + \sum_{j=1}^m \mu_j h_j(\hat{x}) \leq \sum_{i=0}^n \lambda_i \tau_i + \sum_{j=1}^m \mu_j \tau_{n+j}, \quad \forall (\tau_0, \cdots, \tau_{n+m}) \in C. \tag{3.36}$$

Sea por otra parte $\tilde{C}$ definido por

$$\tilde{C} = \{(\tau_0, \cdots, \tau_{n+m}) \in \mathbb{R}^{n+m+1} \,/\, \exists\, x \in K \text{ con}$$

$$f(x) \leq \tau_0,\ g_i(x) \leq \tau_i,\ 1 \leq i \leq n,\ h_j(x) = \tau_{n+j},\ 1 \leq j \leq m\}.$$

Entonces para todo $(\tau_0, \cdots, \tau_{n+m}) \in \tilde{C}$ y para todo $\delta > 0$, se tiene que el elemento $(\tau_0 + \delta, \tau_1, \cdots, \tau_{n+m}) \in C$. Sustituyendo este elemento en (3.36) y tomando límite cuando δ tiende a cero se deduce que (3.36) es cierto para todo $(\tau_0, \cdots, \tau_{n+m}) \in \tilde{C}$.

Dado $x \in K$, y teniendo en cuenta que $(f(x), g_1(x), \cdots, g_n(x), h_1(x), \cdots, h_m(x)) \in \tilde{C}$ se deduce entonces que $\hat{x}$ es solución de (3.35).

Dado $i \in \{0, \cdots, n\}$ y definiendo $e^i = (e^i_0, \cdots, e^i_{n+m}) \in \mathbb{R}^{n+m+1}$ por

$$e^i_l = 1 \text{ si } l = i, \quad e^i_l = 0 \text{ si } l \neq i,$$

y tomando en (3.36)

$$(\tau_0, \cdots, \tau_{n+m}) = (f(\hat{x}), 0, \cdots, 0) + e^i \in \tilde{C},$$

se deduce $\lambda_i \geq 0$, $i \in \{0, \cdots, n\}$. Como estamos suponiendo $g_i(\hat{x}) = 0$, está claro también que se tiene $\lambda_i g_i(\hat{x}) = 0$ para $i \in \{1, \cdots, n\}$. Esto prueba el resultado cuando $g_i(\hat{x}) = 0$, $i \in \{1, \cdots, n\}$.

b) Supongamos ahora el caso general y sea $I = \{i_1, \cdots, i_l \,/\, i_1 \leq \cdots \leq i_l\}$, el conjunto de índices de $\{1, \cdots, n\}$, en los cuales la desigualdad $g_i(\hat{x}) \leq 0$ es estricta. Consideramos ahora el conjunto de restricciones:

$$\mathcal{U}^* = \Big\{x \in K \,/\, g_i(x) \leq 0,\ i \notin I,\ h_j(x) = 0,\ j \in \{1, \cdots, m\}\Big\},$$

y veamos que $\hat{x}$ es solución del problema

$$\text{mín}\{f(x) \,/\, x \in \mathcal{U}^*\}, \tag{3.37}$$

Sea $x \in \mathcal{U}^*$. Entonces, para todo $\varepsilon \in (0,1)$, también se tiene $g_i(\hat{x} + \varepsilon(x - \hat{x})) \leq 0$, $i \notin I$, $h_j(\hat{x} + \varepsilon(x - \hat{x})) = 0$, $j \in \{1, \cdots, m\}$. Además, si $i \in I$, se tiene

$$g_i(\hat{x} + \varepsilon(x - \hat{x})) \leq (1 - \varepsilon) g_i(\hat{x}) + \varepsilon g_i(x) \leq (1 - \varepsilon) \max_{k \in I} g_k(\hat{x}) + \varepsilon \max_{k \in I} g_k(x).$$

Como $\max_{k \in I} g_k(\hat{x}) < 0$, se deduce, que si $\varepsilon > 0$ es suficientemente pequeño, entonces $g_i(\hat{x} + \varepsilon(x - \hat{x})) \leq 0$, $i \in I$, es decir $\hat{x} + (1 - \varepsilon)x$ pertenece a $\mathcal{U}$, se tiene entonces

$$f(\hat{x}) \leq f(\hat{x} + \varepsilon(x - \hat{x})) \leq (1 - \varepsilon) f(\hat{x}) + \varepsilon f(x),$$

de donde se deduce $f(\hat{x}) \leq f(x)$, luego $\hat{x}$ es solución del problema (3.37).

Aplicando el resultado ya probado a este problema, deducimos la existencia de $\lambda_0, \lambda_{i_1}, \cdots, \lambda_{i_l}$, $\mu_1, \cdots, \mu_m$ en las condiciones del Teorema aplicado a (3.37). Tomando $\lambda_i = 0$ si $i \in I$, es fácil ver que se cumplen todas las condiciones del Teorema.

Supongamos ahora que se verifica la condición de Slater. Tenemos que probar $\lambda_0 \neq 0$. Por reducción al absurdo, supongamos $\lambda_0 = 0$, sabemos entonces que debe existir $i_0 \in \{1, \cdots, n\}$ tal que $\lambda_{i_0} \neq 0$. Dado este i_0, consideremos $x_{i_0} \in \mathcal{U}$ tal que $g_{i_0}(x_{i_0}) < 0$. Usando que en $\hat{x}$ se alcanza el mínimo que aparece en (3.35), $\lambda_i \geq 0$ y $\lambda_i g_i(\hat{x}) = 0$, $i \in \{1, \cdots, n\}$, se tendrá entonces

$$0 = \sum_{i=1}^{n} \lambda_i g_i(\hat{x}) \leq \sum_{i=1}^{n} \lambda_i g_i(x_{i_0}) \leq \lambda_{i_0} g_{i_0}(x_{i_0}) < 0,$$

lo que da la contradicción.

Para terminar supongamos ahora $\hat{x} \in \mathcal{U}$ tal que existen λ_i, $i \in \{0, \cdots, n\}$, μ_j, $j \in \{1, \cdots, m\}$ en las condiciones del Teorema con $\lambda_0 = 1$. Dado $x \in \mathcal{U}$, y usando las propiedades de los λ_i y μ_j, se tiene entonces

$$f(\hat{x}) = f(\hat{x}) + \sum_{i=1}^{n} \lambda_i g_i(\hat{x}) + \sum_{i=1}^{m} \mu_i h_i(\hat{x}) \leq f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{i=1}^{m} \mu_i h_i(x) \leq f(x).$$

Esto prueba por tanto que $\hat{x}$ es solución de (3.34). $\square$

Observación 3.5.6 *Si en el Teorema anterior existen*

$$\delta^+ f(\hat{x}, x - \hat{x}), \quad \delta^+ g_i(\hat{x}, x - \hat{x}),\ i \in \{1, \cdots, n\}, \quad \forall\, x \in K$$

(la existencia de $\delta^+ h_j(\hat{x}, x - \hat{x})$, $j \in \{1, \cdots, m\}$ se verifica siempre ya que son afines), entonces, usando el Teorema 3.4.7, se tiene que decir que $\hat{x}$ es solución de (3.35) es equivalente a

$$\lambda_0 \delta^+ f(\hat{x}, x - \hat{x}) + \sum_{i=1}^{n} \lambda_i \delta^+ g_i(\hat{x}, x - \hat{x}) + \sum_{i=1}^{m} \mu_i \delta^+ h_i(\hat{x}, x - \hat{x}) \geq 0.$$

Observación 3.5.7 *Para el estudio de un problema de minimización con infinitas restricciones:*

$$\min\{f(x) \,/\, G(x) \leq 0\}, \tag{3.38}$$

donde f es definida sobre un espacio vectorial X y G es una aplicación de X en un e.n. Z puede consultarse por ejemplo el libro de D. Luenberger.

Capítulo 4

ELEMENTOS DEL CÁLCULO DE VARIACIONES

Como aplicación de los resultados expuestos vamos ahora a realizar una pequeña introducción a los problemas del Cálculo de Variaciones clásico. El Cálculo de Variaciones constituye la primera fuente histórica de problemas de extremos planteados en espacios de dimensión infinita (espacios de funciones).

4.1. Ejemplos clásicos

Son ejemplos clásicos de este tipo de problemas:

a) Hallar la curva plana de clase C^1 que da la menor distancia a dos puntos dados (x_0, y_0) y (x_1, y_1) del plano (problema de respuesta bien conocida), el cual se formula

$$\text{mín}\left\{\int_{x_0}^{x_1} \sqrt{1+\dot{y}^2(x)}\, dx \,/\, y \in C^1[x_0, x_1], y(x_0) = y_0,\, y(x_1) = y_1\right\}.$$

b) Hallar la curva plana de clase C^1 que une dos puntos dados (x_0, y_0) y (x_1, y_1) del plano, y que al girar alrededor del eje OX engendra una superficie de área mínima, el cual se formula

$$\text{mín}\left\{\int_{x_0}^{x_1} y(x)\sqrt{1+\dot{y}^2(x)}\, dx \,/\, y \in C^1[x_0, x_1],\, y(x_0) = y_0,\, y(x_1) = y_1\right\}.$$

c) Problema de la braquistocrona.- Hallar la curva plana de clase C^1 que une dos puntos dados $P_0 = (x_0, y_0)$ y $P_1 = (x_1, y_1)$ y que produce la caída de P_0 a P_1, bajo la sola acción de la gravedad, en el menor tiempo posible.

De la igualdad de energías

$$mgy_0 = mgy(t) + \frac{1}{2}m|v|^2(t)$$

se obtiene $(v = (1, \dot{y}))$

$$T = \int_0^T dt = \int_{P_0}^{P_1} \frac{d\ell}{|v|} = \int_{x_0}^{x_1} \frac{\sqrt{1+\dot{y}^2(x)}}{\sqrt{2g(y_0 - y(x))}}\, dx.$$

Por consiguiente, el problema se formula

$$\text{mín}\left\{\int_{x_0}^{x_1} \frac{\sqrt{1+\dot{y}^2(x)}}{\sqrt{2g(y_0 - y(x))}}\, dx \,/\, y \in C^1[x_0, x_1],\, y(x_0) = y_0,\, y(x_1) = y_1\right\}$$

d) El problema de la reina Dido.- Hallar la curva plana de clase C^1 que une dos puntos dados $(x_0, 0)$ y $(x_1, 0)$ del plano, que tiene una longitud prefijada l y que encierra con el eje OX un área máxima, el cuál se formula:

$$\text{máx}\left\{\int_{x_0}^{x_1} y(x)\,dx \,/\, y\in C^1[x_0,x_1],\; y(x_0)=0,\; y(x_1)=0,\; \int_{x_0}^{x_1}\sqrt{1+\dot y^2(x)}\,dx = l\right\}.$$

Es éste un problema de los que se denominan de tipo isoperimétrico.

En la presente tema vamos a ver como en algunos casos se pueden resolver este tipo de problemas. En realidad, a fin de disponer de resultados que nos garanticen la existencia de solución, sería preferible trabajar con espacios de tipo Sobolev en lugar del espacio C^1. Como nuestro principal objetivo va a ser no la existencia sino el obtener condiciones necesarias que debe verificar una función para ser solución, nos vamos a limitar al caso de soluciones C^1, lo cual simplifica un poco la exposición.

Antes de ver cómo se resuelven estos problemas vamos a estudiar directamente tres ejemplos de problemas de minimización en el conjunto:

$$\mathcal{U} = \left\{x\in H^1([0,4]) \,/\, x(0)=0,\; x(4)=2\right\}.$$

$$\text{ínf}\left\{\int_0^4 \dot x(t)^2\,dt \,/\, x\in\mathcal{U}\right\}. \tag{4.1}$$

$$\text{ínf}\left\{\int_0^4 |1-\dot x(t)^2|\,dt \,/\, x\in\mathcal{U}\right\}. \tag{4.2}$$

$$\text{ínf}\left\{\int_0^4 \arctan^2(\dot x(t)^2)\,dt \,/\, x\in\mathcal{U}\right\}. \tag{4.3}$$

El problema (4.1) tiene una única solución y es $\hat x(t) = t/2$. En efecto $\hat x\in\mathcal{U}$ y para todo $x\in\mathcal{U}$ se tiene (omitiendo la variable muda t):

$$\int_0^4 \dot x^2\,dt - \int_0^4 \dot{\hat x}^2\,dt = \int_0^4 (\dot x-\dot{\hat x})(\dot x+\dot{\hat x})\,dt = \int_0^4 (\dot x-\dot{\hat x})(\dot x-\dot{\hat x}+2\dot{\hat x})\,dt = \int_0^4 (\dot x-\dot{\hat x})^2\,dt + \int_0^4 (\dot x-\dot{\hat x})\,dt =$$

$$= \int_0^4 (\dot x-\dot{\hat x})^2\,dt + [x-\hat x]_0^4 = \int_0^4 (\dot x-\dot{\hat x})^2\,dt \geq 0,$$

luego $\hat x(t) = t/2$ es una solución de (4.1), además, esta última integral se anula si y sólo si $\dot x - \dot{\hat x} = 0$ i.e. $x-\hat x = C$, constante, pero como $\hat x$ y $x\in\mathcal{U}$ entonces $C=0$ i.e. $x=\hat x$.

El problema (4.2) tiene infinitas soluciones. Obviamente el ínfimo es no negativo y para las funciones de $\mathcal{U}$ que tienen pendiente ±1 es nulo. Ahora bien, existen infinitas soluciones que verifican esto: las funciones lineales a trozos de pendiente ±1 cuyas gráficas están situadas en el rectángulo de vértices $(0,0), (1,-1), (4,2), (3,3)$ y que unen los vértices $(0,0)$ y $(4,2)$.

El problema (4.3) no tiene solución. Obviamente el ínfimo es no negativo. Veamos que es cero, para ello sea $n\in\mathbb{N}$ y definimos la sucesión minimizante: $x_n(t)=0$ si $0\leq t\leq 4-1/n$ y $x_n(t) = 2n(1/n-4+t)$ si $4-1/n\leq t\leq 4$, entonces se tiene:

$$0 \leq \int_0^4 \arctan^2(\dot x_n(t)^2)\,dt = \int_{4-1/n}^4 \arctan^2(\dot x_n(t)^2)\,dt \leq \frac{1}{n}\frac{\pi^2}{4} \longrightarrow 0,$$

luego el ínfimo es cero. Si existe una función $\hat x\in\mathcal{U}$ que lo alcanza entonces $\dot{\hat x}=0$ i.e. $\hat x$ es contante y las funciones constantes no están en $\mathcal{U}$.

4.2. Algunos resultados importantes de derivabilidad.

Vamos a ver en esta sección dos resultados referentes a la derivabilidad de los funcionales que aparecen usualmente en el Cálculo de Variaciones y que nos permitirán aplicar los teoremas estudiados en el Capítulo anterior a este tipo de problemas.

Como suele ser usual en el Cálculo de Variaciones, los puntos de $\mathbb{R}^{1+N}$ los vamos a denotar por (t,x), con $t \in \mathbb{R}$, $x \in \mathbb{R}^N$ aunque la letra x se usará también para denotar funciones.

Teorema 4.2.1 *Sea $\Omega \subset \mathbb{R}^{1+N}$ abierto, y sea I un intervalo cerrado y acotado de $\mathbb{R}$. Entonces, el conjunto*

$$\mathcal{O} = \big\{x \in C^0(I;\mathbb{R}^N) \,/\, (t,x(t)) \in \Omega,\ \forall\, t \in I\big\}$$

(que podría ser vacío) es un abierto de $C^0(I)$.

Sea $f \in C^0(\Omega)$, tal que existe $\nabla_x f : \Omega \to \mathbb{R}^N$ (el gradiente con respecto a las N últimas variables) y $\nabla_x f \in C^0(\Omega;\mathbb{R}^N)$, entonces el funcional (de V.V. Nemitskii) $N : \mathcal{O} \to C^0(I)$, definido por

$$N(x)(t) = f(t,x(t)), \quad \forall\, t \in I,\ \forall\, x \in \mathcal{O}$$

es de clase C^1 y se verifica

$$N'(x)h(t) = \nabla_x f(t,x(t)) \cdot h(t), \quad \forall\, t \in I,\ \forall\, x \in \mathcal{O},\ \forall\, h \in C^0(I;\mathbb{R}^N).$$

Demostración: Vamos a probar que $\mathcal{O}$ es abierto, para ello consideramos $x_0 \in \mathcal{O}$ y vamos a ver que existe un entorno de este punto contenido en $\mathcal{O}$. Denotamos

$$\mathcal{G}(x_0) = \big\{(t,x_0(t)) \,/\, t \in I\big\} \subset \Omega, \tag{4.4}$$

el grafo de x_0. Como x_0 es continua e I compacto, se verifica que $\mathcal{G}(x_0)$ es compacto y por tanto

$$d = \mathrm{dist}\big(\mathcal{G}(x_0), \partial\Omega\big) > 0. \tag{4.5}$$

Sea entonces $x \in C^0(I,\mathbb{R}^N)$ tal que $\|x - x_0\|_{C^0} < d$ y $t \in I$, se tiene

$$\mathrm{dist}\big((t,x(t)),\mathcal{G}(x_0)\big) \le |(t,x(t)) - (t,x_0(t))|_{\mathbb{R}^{1+N}} = |x(t) - x_0(t)|_{\mathbb{R}^N} \le \|x - x_0\|_{C^0} < d,$$

lo que implica que $(t,x(t))$ pertenece a Ω, para todo $t \in I$, y por tanto que x pertenece a $\mathcal{O}$. Esto prueba que $\mathcal{O}$ es abierto.

Vamos ahora a probar que N es de clase C^1, comenzamos probando que para toda $x_0 \in \mathcal{O}$ y para toda $h \in C^0(I;\mathbb{R}^N)$, se verifica

$$\begin{aligned} &\mathrm{lím}_{\varepsilon\to 0}\left\|\frac{N(x_0+\varepsilon h) - N(x_0)}{\varepsilon} - \nabla_x f(\cdot,x_0(\cdot)) \cdot h\right\|_{C^0(I)} = \\ &= \mathrm{lím}_{\varepsilon\to 0}\,\mathrm{máx}_{t\in I}\left|\frac{f(t,x_0(t)+\varepsilon h(t)) - f(t,x_0(t))}{\varepsilon} - \nabla_x f(t,x_0(t)) \cdot h(t)\right| = 0 \end{aligned} \tag{4.6}$$

lo que probará que N es diferenciable Gâteaux en x_0 en la dirección h y

$$\delta N(x_0,h)(t) = \nabla_x f(t,x_0(t)) \cdot h(t), \quad \forall\, t \in I.$$

Denotamos

$$K = \Big\{(t,x) \in \Omega \,/\, \mathrm{dist}\big((t,x),\mathcal{G}(x_0)\big) \le \mathrm{mín}\{1,\frac{d}{2}\}\Big\} \tag{4.7}$$

conjunto que se prueba fácilmente que es cerrado y acotado y por tanto compacto. Como $\nabla_x f$ es continua en Ω, se tiene entonces que $\nabla_x f$ es uniformemente continua en K. Sea entonces $\rho > 0$ y consideremos $\delta > 0$ tal que

$$|\nabla_x f(t_1, x_1) - \nabla_x f(t_2, x_2)|_{\mathbb{R}^N} < \rho,$$

para todo $(t_1, x_1), (t_2, x_2) \in K$, con $|(t_1 - t_2, x_1 - x_2)|_{\mathbb{R}^{1+N}} < \delta$. Entonces, para todo $\varepsilon \in \mathbb{R}$ con

$$0 < \varepsilon < \frac{\text{mín}\{\frac{d}{2}, \delta\}}{\|h\|_{C^0} + 1},$$

es fácil ver que en particular se verifica que $(t, x_0(t) + \theta\varepsilon h(t)) \in K$ para todo $t \in I$ y para todo $\theta \in [0, 1]$. Ello permite aplicar el Teorema del valor medio para probar que para todo $t \in I$, existe $\theta(t) \in (0, 1)$ verificando

$$\left|\frac{f(t, x_0(t) + \varepsilon h(t)) - f(t, x_0(t))}{\varepsilon} - \nabla_x f(t, x_0(t)) \cdot h(t)\right| =$$

$$= \big|\big(\nabla_x f(t, x_0(t) + \varepsilon\theta(t)) \cdot h(t) - \nabla_x f(t, x_0(t))\big) \cdot h(t)\big|.$$

Como por las condiciones impuestas a ε, se tiene

$$(t, x_0(t) + \varepsilon\theta(t)h(t)),\ (t, x_0(t)) \in K, \quad |(0, \varepsilon\theta(t)h(t))|_{\mathbb{R}^{1+N}} < \delta,$$

se tiene entonces

$$\big|\nabla_x f(t, x_0(t) + \varepsilon\theta(t)) \cdot h(t) - \nabla_x f(t, x_0(t))\big) \cdot h(t)\big| \leq \rho,$$

para todo $t \in I$. Esto prueba (4.6).

Para probar ahora que N es derivable Gâteaux hay que probar ahora que la aplicación $\delta N(x_0) : h \in C^0(I; \mathbb{R}^N) \longmapsto \delta N(x_0, h) \in C^0(I)$ es lineal y continua. La linealidad es evidente, respecto a la continuidad, basta observar que gracias a que $\mathcal{G}(x_0)$ es compacto, para todo $h \in C^0(I; \mathbb{R}^N)$ se tiene

$$\|\delta N(x_0, h)\|_{C^0(I)} = \max_{t\in I} |\nabla_x f(t, x_0(t)) \cdot h(t)| \leq (\max_{\mathcal{G}(x_0)} |\nabla_x f|_{\mathbb{R}^N}) \max_{t\in I} |h(t)|_{\mathbb{R}^N} = (\max_{\mathcal{G}(x_0)} |\nabla_x f|_{\mathbb{R}^N}) \|h\|_{C^0},$$

lo que prueba la continuidad de $\delta N(x_0)$.

Para terminar la demostración de que N es de clase C^1 hace falta probar que la aplicación $\delta N : x \in \mathcal{O} \longmapsto \delta N(x) \in \mathcal{L}(C^0(I; \mathbb{R}^N), C^0(I))$ es continua en todo punto $x_0 \in \mathcal{O}$. Para ello, dado $x_0 \in \mathcal{O}$, definimos $\mathcal{G}(x_0)$, d y K por (4.4), (4.5) y (4.7). Dado $\varepsilon > 0$, tomamos $\tilde{\delta} > 0$ tal que

$$|\nabla_x f(t_1, x_1) - \nabla_x f(t_2, x_2)|_{\mathbb{R}^N} < \varepsilon,$$

para todo $(t_1, x_1), (t_2, x_2) \in K$, con $|(t_1 - t_2, x_1 - x_2)|_{\mathbb{R}^{1+N}} < \tilde{\delta}$. Tomando entonces $\delta = \text{mín}\{\frac{d}{2}, \tilde{\delta}\}$, es fácil ver que si $\|x - x_0\|_{C^0} < \delta$ entonces $(t, x(t)) \in K$, para todo $t \in I$ y se tiene

$$\|\delta N(x) - \delta N(x_0)\|_{\mathcal{L}(C^0(I;\mathbb{R}^N), C^0(I))} = \max_{\|h\|_{C^0} \leq 1} \|\delta N(x, h) - \delta N(x_0, h)\|_{C^0(I)} =$$

$$= \max_{\|h\|_{C^0} \leq 1} \max_{t\in I} \big|\big(\nabla_x f(t, x(t)) - \nabla_x f(t, x_0(t))\big) \cdot h(t)\big| \leq$$

$$\leq \max_{t\in I} \big|\big(\nabla_x f(t, x(t)) - \nabla_x f(t, x_0(t))\big)\big|_{\mathbb{R}^N} < \varepsilon,$$

y por tanto δN es continua en x_0. □

Observación 4.2.2 *Se tiene un resultado totalmente análogo cuando $f \in C^0(\Omega, \mathbb{R}^M)$, cambiando el vector gradiente $\nabla_x f$ por la matriz de Jacobi:*

$$\left(\frac{\partial f_i}{\partial x_j}\right)_{i=1,\cdots,M,j=1,\cdots,N}.$$

Teorema 4.2.3 *Sean $\Omega \subset \mathbb{R}^3$ abierto, cuyos puntos denotamos por $(t, x, \dot{x})$, $I \subset \mathbb{R}$ cerrado y acotado. Entonces, el conjunto*

$$\mathcal{O} = \{x \in C^1(I) \,/\, (t, x(t), \dot{x}(t)) \in \Omega, \ \forall\, t \in I\}$$

es un abierto de $C^1(I)$ (que podría ser vacío).

Sea $f \in C^0(\Omega)$, tal que existen $f_x, f_{\dot{x}} \in C^0(\Omega)$. Entonces el funcional (del Cálculo de Variaciones) $J : \mathcal{O} \longrightarrow \mathbb{R}$, definido por

$$J(x) = \int_I f(t, x(t), \dot{x}(t))\, dt, \forall x \in \mathcal{O}$$

es de clase C^1 en $\mathcal{O}$ y se verifica

$$J'(x)h = \int_I \big(f_x(t, x(t), \dot{x}(t))h(t) + f_{\dot{x}}(t, x(t), \dot{x}(t))\dot{h}(t)\big)\, dt, \quad \forall\, x \in \mathcal{O}, \forall\, h \in C^1(I).$$

Demostración: Consideramos la aplicación $D : C^1(I) \to C^0(I; \mathbb{R}^2)$ definida por

$$D(x)(t) = (x(t), \dot{x}(t)), \quad \forall\, t \in I, \ \forall\, x \in C^1(I),$$

la cual se prueba fácilmente que es lineal y continua. Definiendo entonces

$$\mathcal{V} = \big\{(x_1, x_2) \in C^0(I; \mathbb{R}^2) \,/\, (t, x_1(t), x_2(t)) \in \Omega, \forall\, t \in I\big\},$$

el cual es abierto gracias al Teorema 4.2.1 y teniendo en cuenta la igualdad $\mathcal{O} = D^{-1}(\mathcal{V})$, se deduce que $\mathcal{O}$ es la imagen inversa por una aplicación continua de un abierto y por tanto abierto.

Dada ahora f en las condiciones del Teorema, definimos $N : \mathcal{V} \to C^0(I)$ por

$$N(x_1, x_2)(t) = f(t, x_1(t), x_2(t)), \quad \forall\, t \in I, \ \forall\, (x_1, x_2) \in \mathcal{V},$$

e $\mathcal{I} : C^0(I) \to \mathbb{R}$ por

$$\mathcal{I}(y) = \int_I y(t)\, dt, \quad \forall\, y \in C^0(I),$$

se verifica $J = \mathcal{I} \circ N \circ D$. Como, por el Teorema 4.2.1, N es de clase C^1, e $\mathcal{I}$, D son lineales y continuas y por tanto C^1 se deduce, gracias a la regla de la cadena, que J es de clase C^1. Además, usando la regla de la cadena, los resultados vistos para la derivación de aplicaciones lineales y continuas y el Teorema 4.2.1 se tiene

$$J'(x)h = \mathcal{I}'(N(D(x)))N'(Dx)D'(x)h = \mathcal{I}N'(x, \dot{x})Dh = \mathcal{I}(N'(x, \dot{x})(h, \dot{h})) =$$

$$= \mathcal{I}(f_x(\cdot, x(\cdot), \dot{x}(\cdot))h(\cdot) + f_{\dot{x}}(\cdot, x(\cdot), \dot{x}(\cdot)) = \int_I \big(f_x(t, x(t), \dot{x}(t))h(t) + f_{\dot{x}}(t, x(t), \dot{x}(t))\dot{h}(t)\big)\, dt.$$

□

4.3. El problema más elemental del Cálculo de Variaciones

El problema que vamos a estudiar en esta sección se enuncia como sigue: Entre todas las curvas planas C^1 que unen dos puntos (t_0, x_0) y (t_1, x_1) dados, encontrar la o las que minimizan (o maximizan) un criterio de la forma

$$\int_I f(t, x(t), \dot{x}(t))\, dt,$$

con $I = [t_0, t_1]$. Por el Teorema 4.2.3, sabemos que si $f \in C^0(\Omega)$ con $\Omega \subset \mathbb{R}^3$ abierto es tal que existen las parciales f_x y $f_{\dot{x}}$ y pertenecen a $C^0(\Omega)$, entonces el conjunto

$$\mathcal{O} = \{x \in C^1(I) \,/\, (t, x(t), \dot{x}(t)) \in \Omega \ \forall\, t \in I\},$$

es un abierto de $C^1(I)$ y la aplicación $J : \mathcal{O} \to \mathbb{R}$ definida por

$$J(x) = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))\, dt$$

es de clase C^1. Denotando

$$\mathcal{U} = \{x \in \mathcal{O} \,/\, x(t_0) = x_0,\ x(t_1) = x_1\}$$

(que supondremos no vacío) el problema más elemental del Cálculo de Variaciones se enuncia

$$\text{mín}\{J(x) \,/\, x \in \mathcal{U}\} \tag{4.8}$$

(o máx en vez de mín).

Nosotros nos vamos a interesar en encontrar condiciones necesarias que debe satisfacer una función $x \in \mathcal{U}$ para que se alcance un mínimo local en x.

Como $\mathcal{O}$ es abierto, se tiene que si $x \in \mathcal{U}$ y $h \in C_0^1(I)$, entonces existe un $\varepsilon_0 > 0$, dependiente de x y h, tal que

$$x + \varepsilon h \in \mathcal{U},\ \forall\, \varepsilon \text{ con } |\varepsilon| < \varepsilon_0.$$

Claramente, si en x se alcanza un mínimo local de f, entonces la aplicación $\varepsilon \in [-|\varepsilon_0|, |\varepsilon_0|] \longmapsto f(x + \varepsilon h)$ alcanza un mínimo local en 0, con lo que gracias al Teorema 4.2.3 se deduce

$$J'(x)h = \int_{t_0}^{t_1} \big(f_x(t, x(t), \dot{x}(t))h(t) + f_{\dot{x}}(t, x(t), \dot{x}(t))\dot{h}(t)\big)\, dt = 0, \quad \forall\, h \in C_0^1(I) \tag{4.9}$$

Vamos a ver como esta condición implica que x satisface una cierta ecuación diferencial que se conoce como la ecuación de L. Euler-J.L. Lagrange de f.

Lema 4.3.1 *(de Du Bois-Reymond) Sean $p, q \in C^0([t_0, t_1])$ tales que*

$$\int_{t_0}^{t_1} \Big(p(t)h(t) + q(t)\dot{h}(t)\Big)\, dt = 0,\ \forall h \in C_0^1(I),$$

entonces $q \in C^1(I)$ y se tiene

$$\dot{q}(t) = p(t), \quad \forall\, t \in I. \tag{4.10}$$

Demostración: Gracias al Teorema 2.4.8 se deduce que q pertenece al espacio $H^1(I)$ y se verifica $\dot{q} = p$ en el sentido débil de las derivadas. Como además p es continua la Observación 2.4.3 nos dice que de hecho $q \in C^1(I)$ y su derivada débil coincide con la usual. □

Como consecuencia inmediata de este resultado y de (4.9), se tiene

Teorema 4.3.2 *Si x es una solución local de (4.8), entonces la aplicación*

$$t \in I \longmapsto f_{\dot{x}}(t, x(t), \dot{x}(t)),$$

es de clase C^1 y se verifica la ecuación de Euler-Lagrange

$$\frac{d}{dt}\Big(f_{\dot{x}}(t, x(t), \dot{x}(t))\Big) = f_x(t, x(t), \dot{x}(t)), \quad \forall\, t \in I. \tag{4.11}$$

Observación 4.3.3 *A fin de calcular x, a la ecuación de segundo orden (4.11) hay que añadirle además las dos condiciones de contorno $\hat{x}(t_0) = x_0$, $\hat{x}(t_1) = x_1$.*

Observación 4.3.4 *Se puede probar, mediante el Teorema de la función implícita, que si x es una solución de la ecuación diferencial (4.11), $f_{\dot{x}} \in C^1(\Omega)$ y $f_{\dot{x}\dot{x}}$ no se anula en Ω, entonces x pertenece a $C^2[t_0, t_1]$.*

Observación 4.3.5 *Cuando la función f no depende explícitamente de t ($f = f(x, \hat{x})$) y x es una solución de (4.11) que además es clase C^2, un simple cálculo prueba que la derivada, respecto de t de la expresión*

$$f(x(t), \dot{x}(t)) - \dot{x}(t) f_{\dot{x}}(x(t), \dot{x}(t)),$$

es idénticamente nula en I y por tanto existe $C \in \mathbb{R}$ tal que

$$f(x(t), \dot{x}(t)) - \dot{x}(t) f_{\dot{x}}(x(t), \dot{x}(t)) = C, \quad \forall\, t \in I,$$

con lo que en vez de una ecuación de segundo orden, lo que hay que resolver ahora es una de primer orden.

4.4. Otros tipos de problemas del Cálculo de Variaciones

Vamos a ver ahora algunos ejemplos de problemas más complejos, donde el conjunto de restricciones no se reduce a $x(t_0) = x_0$, $x(t_1) = x_1$. La forma usual que emplearemos para resolver estos problemas consistirá en aplicar bien el Teorema de los multiplicadores de Lagrange, bien el Teorema de Kuhn-Tucker. Así por ejemplo observemos que para estudiar el problema más elemental del Cálculo de Variaciones considerando lo que se conoce como la lagrangiana del problema

$$\mathcal{L}(x) = \mathcal{L}(x; \lambda_0, \lambda_1, \lambda_2) = \lambda_0 \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))\, dt + \lambda_1(x(t_0) - x_0) + \lambda_2(x(t_1) - x_1),$$

el Teorema de los multiplicadores de Lagrange nos dice que si x es solución del problema entonces existen $\lambda_0, \lambda_1, \lambda_2 \in \mathbb{R}$ no todos nulos tales que $\mathcal{L}'(x; \lambda_0, \lambda_1, \lambda_2) = 0$, es decir

$$\mathcal{L}'(x; \lambda_0, \lambda_1, \lambda_2)h = 0, \quad \forall\, h \in C^1([t_0, t_1]).$$

A partir de este resultado, es fácil ahora concluir (4.11).

Obsérvese que a fin de poder aplicar el Teorema de los multiplicadores de Lagrange necesitamos que las aplicaciones consideradas sean de clase C^1 (o al menos estrictamente derivables). En gran parte esto se deducirá del Teorema 4.2.3, si bien en algunos casos nos encontraremos con funcionales que no son del tipo considerado en este Teorema, cuyas propiedades de derivabilidad admitiremos.

4.4.1. Problemas Isoperimétricos

Se trata de problemas en los que aparecen restricciones de tipo integral, como ejemplo consideramos el siguiente:

Encontrar la solución de problema

$$\text{mín}\left\{\int_0^\pi \dot{x}(t)^2\,dt + x(\pi)\,/\,x\in C^1([0,\pi]),\ x(0)=0,\ \int_0^\pi x(t)sen\,t\,dt=0,\right\}. \tag{4.12}$$

probando que la función obtenida es efectivamente solución

Solución.- Puesto que el funcional a minimizar es convexo y las restricciones son lineales, podemos aplicar tanto el Teorema de los multiplicadores de Lagrange como el de Kuhn-Tucker. En cualquier caso definimos la lagrangiana del problema por

$$\mathcal{L}(x)=\lambda_0\left(\int_0^\pi \dot{x}(t)^2\,dt + x(\pi)\right)+\lambda_1 x(0)+\lambda_2\int_0^\pi x(t)\text{sen}\,t\,dt=$$

$$\int_0^\pi \left(\lambda_0\dot{x}(t)^2+\lambda_2 x(t)\text{sen}\,t\right)dt+\lambda_0 x(\pi)+\lambda_1 x(0),$$

(es preferible que aparezca una sola integral). Entonces, si x es solución, existen $\lambda_0,\lambda_1,\lambda_2$ no todos nulos tales que

$$\mathcal{L}'(x)h=0,\quad \forall\, h\in C^1([0,\pi]). \tag{4.13}$$

El Teorema de Kuhn-Tucker nos diría también $\lambda_0\geq 0$ aunque puesto que no sabemos nada sobre el signo de λ_1, λ_2 está claro que esto siempre se puede suponer y por tanto no aporta ninguna información. Una ventaja a destacar del Teorema de Kuhn-Tucker es que no necesita que los funciones sean C^1, bastaría con diferenciables Gâteaux.

Usamos ahora

$$\mathcal{L}'(x)h=\frac{d}{d\varepsilon}\mathcal{L}(x+\varepsilon h)_{|\varepsilon=0},$$

donde

$$\mathcal{L}(x+\varepsilon h)=\int_0^\pi \left(\lambda_0(\dot{x}(t)+\varepsilon\dot{h}(t))^2+\lambda_2(x(t)+\varepsilon h(t))\text{sen}\,t\right)dt+$$
$$+\lambda_0(x(\pi)+\varepsilon h(\pi))+\lambda_1(x(0)+\varepsilon h(0)).$$

Por tanto

$$\frac{d}{d\varepsilon}\mathcal{L}(x+\varepsilon h)=\int_0^\pi\left(2\lambda_0(\dot{x}(t)+\varepsilon\dot{h}(t))\dot{h}(t)+\lambda_2 h(t)\text{sen}\,t\right)dt+\lambda_0 h(\pi)+\lambda_1 h(0).$$

La condición (4.13), se escribe por tanto

$$\int_0^\pi\left(2\lambda_0\dot{x}(t)\dot{h}(t)+\lambda_2 h(t)\text{sen}\,t\right)dt+\lambda_0 h(\pi)+\lambda_1 h(0)=0,\quad \forall\, h\in C^1([0,\pi]). \tag{4.14}$$

En particular, se tiene

$$\int_0^\pi\left(2\lambda_0\dot{x}(t)\dot{h}(t)+\lambda_2 h(t)\text{sen}\,t\right)dt=0,\quad \forall\, h\in C_0^1([0,\pi]). \tag{4.15}$$

El Lema de Du Bois Raymond nos dice entonces que se verifica la ecuación diferencial

$$\frac{d}{dt}(2\lambda_0\dot{x}(t))=\lambda_2\text{sen}\,t,\quad \forall\, t\in[0,\pi]. \tag{4.16}$$

Teniendo en cuenta esta ecuación e integrando por partes, se tiene que para toda $h \in C^1([0,\pi])$, se verifica

$$\int_0^\pi 2\lambda_0 \dot{x}(t)\dot{h}(t)\,dt = 2\lambda_0\dot{x}(\pi)h(\pi) - 2\lambda_0\dot{x}(0)h(0) - \int_0^\pi \frac{d}{dt}(2\lambda_0\dot{x}(t))h(t)\,dt =$$

$$= 2\lambda_0\dot{x}(\pi)h(\pi) - 2\lambda_0\dot{x}(0)h(0) - \int_0^\pi \lambda_2 \operatorname{sen} t\, h(t)\,dt.$$

Sustituyendo esta relación en (4.14), se obtiene

$$\lambda_0(1+2\dot{x}(\pi))h(\pi) + (\lambda_1 - 2\lambda_0\dot{x}(0))h(0) = 0, \quad \forall\, h \in C_0^1([0,\pi]).$$

Tomando h tal que $h(0) = 0$, $h(\pi) = 1$ se deduce

$$\lambda_0(1+2\dot{x}(\pi)) = 0. \tag{4.17}$$

Tomando h tal que $h(0) = 1$, $h(\pi) = 0$ se deduce

$$\lambda_1 - 2\lambda_0\dot{x}(0) = 0. \tag{4.18}$$

La condición (4.14) se reduce por tanto a (4.16), (4.17) y (4.18). A estas condiciones hay que añadirles el hecho de que x verifica las restricciones del problema, es decir

$$x(0) = 0 \tag{4.19}$$

$$\int_0^\pi x(t)\operatorname{sen} t\,dt = 0. \tag{4.20}$$

Como los multiplicadores λ_0, λ_1, λ_2 están definidos salvo una constante multiplicativa, alguno de ellos, que no sea nulo, se podrá tomar igual a 1. Vamos a ver que $\lambda_0 \neq 0$. Razonamos por reducción al absurdo, si fuera cero, (4.18) implicaría $\lambda_1 = 0$. La condición (4.16) nos daría también $\lambda_2 = 0$ pero entonces todos los multiplicadores son nulos. Por tanto $\lambda_0 \neq 0$ y podemos suponer $\lambda_0 = 1$. La condición (4.18) se escribe por tanto $\lambda_1 = 2\dot{x}(0)$, ecuación que ahora no aporta nada, su única utilidad sería la de permitirnos calcular λ_1 en caso de que nos hiciera falta.

Resolviendo la ecuación (4.16), se deduce

$$x(t) = -\frac{\lambda_2}{2}\operatorname{sen} t + At + B, \quad \forall\, t \in [0,\pi],$$

con $A, B \in \mathbb{R}$. Las ecuación (4.19) nos da $B = 0$ mientras que (4.17) (4.20) nos proporciona el sistema

$$\frac{\lambda_2}{2} + A = -\frac{1}{2}; \quad -\frac{\lambda_2}{4}\pi + A\pi = 0,$$

que implica $A = -1/6$, $\lambda_2 = -2/3$. Por tanto se tiene finalmente

$$x(t) = \frac{1}{3}\operatorname{sen} t - \frac{1}{6}t, \quad \forall\, t \in [0,\pi].$$

El problema nos pide ahora probar que esta función es efectivamente solución del problema. En realidad ello se deduce inmediatamente de Kuhn-Tucker ya que $\lambda_0 \neq 0$. Otra demostración de este resultado algo más directa consiste en considerar otra función arbitraria satisfaciendo las restricciones del problema, la cual la podemos escribir en la forma $x+h$, con x la solución encontrada y $h \in C^1([0,\pi])$. Como tanto x como h verifican las restricciones del problema, se deduce que h satisface

$$h(0)=0, \quad \int_0^\pi h(t)\operatorname{sen} t\, dt = 0. \tag{4.21}$$

Para ver que x es solución del problema tendremos entonces que probar

$$\int_0^\pi (\dot{x}(t)+\dot{h}(t))^2\, dt + x(\pi) + h(\pi) \geq \int_0^\pi \dot{x}(t)^2\, dt + x(\pi), \tag{4.22}$$

para toda $h \in C^1([0,\pi])$ que verifique (4.21). Desarrollando, se tiene

$$\int_0^\pi (\dot{x}(t)+\dot{h}(t))^2\, dt + x(\pi)+h(\pi) = \int_0^\pi \dot{x}(t)^2\, dt + x(\pi) + 2\int_0^\pi \dot{x}(t)h(t)\, dt + h(\pi) + \int_0^\pi \dot{h}(t)^2\, dt \tag{4.23}$$

Ahora bien, como h satisface (4.21) y $\lambda_0 = 1$, la condición (4.14) prueba

$$2\int_0^\pi \dot{x}(t)h(t)\, dt + h(\pi) = 0$$

y por tanto (4.22), se deduce inmediatamente de (4.23).

Podemos razonar también como en el ejemplo 4.1. Sabemos que un candidato a solución es $\hat{x}(t) = \frac{1}{3}\operatorname{sen} t - \frac{1}{6}t, \quad \forall\, t \in [0,\pi]$. Sea x un punto admisible i.e. $x \in \mathcal{U} = \{x \in C^1([0,\pi]) \,/\, x(0) = 0,\ \int_0^\pi x(t)\operatorname{sen} t\, dt = 0\}$ entonces

$$\int_0^\pi \dot{x}^2\, dt + x(\pi) - \int_0^\pi \dot{\hat{x}}^2\, dt - \hat{x}(\pi) = \int_0^\pi (\dot{x}-\dot{\hat{x}})(\dot{x}+\dot{\hat{x}})\, dt + x(\pi) - \hat{x}(\pi) =$$

$$= \int_0^\pi (\dot{x}-\dot{\hat{x}})(\dot{x}-\dot{\hat{x}}+2\dot{\hat{x}})\, dt + x(\pi) - \hat{x}(\pi) = \int_0^\pi (\dot{x}-\dot{\hat{x}})^2\, dt + 2\int_0^\pi \dot{\hat{x}}(\dot{x}-\dot{\hat{x}})\, dt + x(\pi) - \hat{x}(\pi) =$$

$$= \int_0^\pi (\dot{x}-\dot{\hat{x}})^2\, dt + 2\int_0^\pi (\frac{1}{3}\cos t - \frac{1}{6})(\dot{x}-\dot{\hat{x}})\, dt + x(\pi) - \hat{x}(\pi) =$$

$$= \int_0^\pi (\dot{x}-\dot{\hat{x}})^2\, dt + \left[(\frac{2}{3}\cos t - \frac{1}{3})(x(t)-\hat{x}(t))\right]_0^\pi - \int_0^\pi (-\frac{2}{3}\operatorname{sen} t)(\hat{x}(t)-x(t))\, dt + x(\pi) - \hat{x}(\pi) =$$

$$= \int_0^\pi (\dot{x}-\dot{\hat{x}})^2\, dt \geq 0,$$

luego $\hat{x}(t) = \frac{1}{3}\operatorname{sen} t - \frac{1}{6}t$ es una solución de (4.1), además, esta última integral se anula si y sólo si $\dot{x}-\dot{\hat{x}} = 0$ i.e. $x - \hat{x} = C$, constante, pero como $\hat{x}$ y $x \in \mathcal{U}$ entonces $C = 0$ i.e. $x = \hat{x}$, luego $\hat{x}$ es la única solución del problema (4.12).

4.4.2. Problemas de punto final variable

Estos problemas se caracterizan porque al menos uno de los extremos del intervalo en el cual está planteado el problema es también una incógnita. Consideramos el ejemplo siguiente (el cual no sólo es de punto final variable sino también isoperimétrico).

Supuesta la existencia de solución, resolver el problema

$$\min\left\{\int_0^T x(t)^2\dot{x}(t)^2\, dt \,/\, T>0,\ x \in C^1([0,T]),\, x(0)=0,\, x(T)=1,\, \int_0^T x(t)^2\, dt = 1.\right\}$$

Solución.- Supongamos que tenemos una solución (T,x). Usando que dada $S\in(0,T+1)$ e $y\in C^1([0,S])$ existe $\tilde y\in C^1([0,T+1])$ tal que su restricción a $[0,S]$ coincide con y, está claro que la pareja formada por T y una prolongación $\tilde x$ de x a $C^1[0,T+1])$ es también solución del problema

$$\text{mín}\left\{\int_0^S \tilde y(t)^2\dot{\tilde y}(t)^2\,dt \,/\, S\in(0,T+1),\tilde y\in C^1([0,T+1]),\tilde y(0)=0,\tilde y(S)=1,\int_0^S \tilde y(t)^2\,dt=1\right\}.$$

Este problema tiene la ventaja de estar planteado en el abierto $\mathcal{O}=(0,T+1)\times C^1([0,T+1])$ que es un abierto del espacio de Banach $\mathbb{R}\times C^1([0,T+1])$. Además $(T,\tilde x)\in\mathcal{O}$. El Teorema de los multiplicadores nos dice entonces que existen λ_0, λ_1, λ_2, λ_3 no todos nulos tales que definiendo la lagrangiana

$$\mathcal{L}(S,\tilde y)=\lambda_0\int_0^S \tilde y(t)^2\dot{\tilde y}(t)^2\,dt+\lambda_1\tilde y(0)+\lambda_2(\tilde y(S)-1)+\lambda_3\left(\int_0^S \tilde y(t)^2\,dt-1\right)=$$

$$=\int_0^S \tilde y(t)^2\big(\lambda_0\dot{\tilde y}(t)^2+\lambda_3\big)\,dt+\lambda_1\tilde y(0)+\lambda_2(\tilde y(S)-1)-\lambda_3,\quad \forall(S,\tilde y)\in\mathbb{R}\times C^1([0,T+1]),$$

se tiene

$$\mathcal{L}'(T,\tilde x)(\beta,\tilde h)=\frac{d}{d\varepsilon}\mathcal{L}(T+\varepsilon\beta,\tilde x+\varepsilon\tilde h)_{|\varepsilon=0}=0,\quad \forall\,(\beta,\tilde h)\in\mathbb{R}\times C^1([0,T+1]). \tag{4.24}$$

Usando que $\mathcal{L}(T+\varepsilon\beta,\tilde x+\varepsilon\tilde h)$ viene dado por

$$\mathcal{L}(T+\varepsilon\beta,\tilde x+\varepsilon\tilde h)=\int_0^{T+\varepsilon\beta}(\tilde x(t)+\varepsilon\tilde h(t))^2\big(\lambda_0(\dot{\tilde x}(t)+\varepsilon\dot{\tilde h}(t))^2+\lambda_3\big)\,dt+$$
$$+\lambda_1(\tilde x(0)+\varepsilon\tilde h(0))+\lambda_2(\tilde x(T+\varepsilon\beta)+\varepsilon\tilde h(T+\varepsilon\beta)-1)-\lambda_3,$$

se deduce

$$\frac{d}{d\varepsilon}\mathcal{L}(T+\varepsilon\beta,\tilde x+\varepsilon\tilde h)=(\tilde x(T+\varepsilon\beta)+\varepsilon\tilde h(T+\varepsilon\beta))^2\big(\lambda_0(\dot{\tilde x}(T+\varepsilon\beta)+\varepsilon\dot{\tilde h}(T+\varepsilon\beta))^2+\lambda_3\big)\beta+$$
$$+\int_0^{T+\varepsilon\beta}[2(\tilde x(t)+\varepsilon\tilde h(t))\tilde h(t)\big(\lambda_0(\dot{\tilde x}(t)+\varepsilon\dot{\tilde h}(t))^2+\lambda_3\big)+(\tilde x(t)+\varepsilon\tilde h(t))^2\lambda_0 2(\dot{\tilde x}(t)+\varepsilon\dot{\tilde h}(t))\dot{\tilde h}(t)]\,dt+$$
$$+\lambda_1\tilde h(0)+\lambda_2\big(\dot{\tilde x}(T+\varepsilon\beta)\beta+\tilde h(T+\varepsilon\beta)+\varepsilon\dot{\tilde h}(T+\varepsilon\beta)\beta\big).$$

Haciendo entonces $\varepsilon=0$, la condición (4.24) se escribe entonces

$$\begin{gathered}\tilde x(T)^2\big(\lambda_0\dot{\tilde x}(T)^2+\lambda_3\big)\beta+\textstyle\int_0^T\Big(2\tilde x(t)\tilde h(t)\big(\lambda_0\dot{\tilde x}(t)^2+\lambda_3\big)+\tilde x(t)^2\lambda_0 2\dot{\tilde x}(t)\dot{\tilde h}(t)\Big)\,dt+\\ +\lambda_1\tilde h(0)+\lambda_2\big(\dot{\tilde x}(T)\beta+\tilde h(T)\big)=0,\quad \forall(\beta,\tilde h)\in\mathbb{R}\times C^1([0,T+1]).\end{gathered} \tag{4.25}$$

Como vemos en la expresión anterior sólo aparecen los valores de $\tilde x$ en el intervalo $[0,T]$ que es donde coincide con x, de la misma forma para cada función $\tilde h\in C^1([0,T])$, en (4.25) sólo aparece la restricción h de $\tilde h$ a $[0,T]$. Esto ocurre siempre en este tipo de problemas, obsérvese que en todo lo anterior en vez de $T+1$ podríamos haber considerado cualquier número mayor que T. De esta forma, (4.25) se escribe

$$\begin{gathered}x(T)^2\big(\lambda_0\dot x(T)^2+\lambda_3\big)\beta+\textstyle\int_0^T\Big(2x(t)h(t)\big(\lambda_0\dot x(t)^2+\lambda_3\big)+2\lambda_0 x(t)^2\dot x(t)\dot h(t)\Big)\,dt+\\ \lambda_1 h(0)+\lambda_2\big(\dot x(T)\beta+h(T)\big)=0,\qquad \forall\,(\beta,h)\in\mathbb{R}\times C^1([0,T]).\end{gathered} \tag{4.26}$$

Tomando en la expresión anterior $\beta = 0$, $h \in C_0^1([0,T])$, se deduce

$$\int_0^T \Big(2x(t)h(t)\big(\lambda_0 \dot{x}(t)^2 + \lambda_3\big) + 2\lambda_0 x(t)^2 \dot{x}(t)\dot{h}(t)\Big)\,dt = 0, \qquad \forall\, h \in C_0^1([0,T]).$$

Por el lema de Du Bois-Raymond se deduce entonces que x verifica la ecuación diferencial

$$\frac{d}{dt}\Big(2\lambda_0 x(t)^2 \dot{x}(t)\Big) = 2x(t)\big(\lambda_0 \dot{x}(t)^2 + \lambda_3\big), \qquad \forall\, t \in [0,T]. \tag{4.27}$$

Integrando por partes y usando (4.27), se deduce que para toda $h \in C^1([0,T])$, se tiene

$$\int_0^T 2\lambda_0 x(t)^2 \lambda_0 2\dot{x}(t)\dot{h}(t)\,dt = 2\lambda_0 x(T)^2 \dot{x}(T)h(T) - 2\lambda_0 x(0)^2 \dot{x}(0)h(0) - \int_0^T \frac{d}{dt}\Big(2\lambda_0 x(t)^2 \dot{x}(t)\Big)h(t)\,dt =$$

$$= 2\lambda_0 x(T)^2 \dot{x}(T)h(T) - 2\lambda_0 x(0)^2 \dot{x}(0)h(0) - \int_0^T 2x(t)\big(\lambda_0 \dot{x}(t)^2 + \lambda_3\big)h(t)\,dt.$$

Sustituyendo esta expresión en (4.36) y tomando $\beta = 0$, se obtiene

$$\big(2\lambda_0 x(T)^2 \dot{x}(T) + \lambda_2\big)h(T) + \big(-2\lambda_0 x(0)^2 \dot{x}(0) + \lambda_1\big)h(0) = 0, \quad \forall h \in C^1([0,T]),$$

lo que como en el ejemplo anterior implica

$$2\lambda_0 x(T)^2 \dot{x}(T) + \lambda_2 = 0 \tag{4.28}$$

$$-2\lambda_0 x(0)^2 \dot{x}(0) + \lambda_1 = 0. \tag{4.29}$$

Tomando ahora $h = 0$, $\beta = 1$ en (4.36) se tiene también

$$x(T)^2\big(\lambda_0 \dot{x}(T)^2 + \lambda_3\big) + \lambda_2 \dot{x}(T) = 0. \tag{4.30}$$

La condición (4.36) resulta entonces ser equivalente a (4.27), (4.28), (4.29) y (4.30). A estas condiciones tenemos que añadirle las restricciones del problema

$$x(0) = 0 \tag{4.31}$$

$$x(T) = 1 \tag{4.32}$$

$$\int_0^T x(t)^2\,dt = 1. \tag{4.33}$$

Como en el ejemplo anterior, vamos a ver que λ_0 es distinto de cero y que por tanto se puede tomar igual a 1. Por reducción al absurdo, supongamos $\lambda_0 = 0$, entonces las ecuaciones (4.27), (4.28), (4.29) prueban respectivamente $\lambda_3 = 0$, $\lambda_2 = 0$, $\lambda_1 = 0$ pero entonces todos los multiplicadores son nulos, lo cual es una contradicción. Tomamos entonces $\lambda_0 = 1$. Como en el ejemplo anterior podemos ahora prescindir de la ecuación (4.29), que sólo serviría para calcular λ_1. Sin embargo éste no es el caso de (4.28) (que está asociada al extremo variable T) ya que λ_2 también aparece en (4.30). Lo que se hace usualmente es eliminar este multiplicador de las dos ecuaciones correspondientes obteniéndose lo que se conoce como la condición de transversalidad del problema. En nuestro caso, despejando λ_2 de (4.28) y sustituyendo en (4.30) se obtiene la condición de transversalidad

$$x(T)^2(-\dot{x}(T)^2 + \lambda_3) = 0, \tag{4.34}$$

Tendríamos ahora que resolver la ecuación diferencial (4.27), mejor que tratar de hacerlo directamente, es usar que esta ecuación no es otra que la ecuación de Euler-Lagrange del funcional

$$f(x, \dot{x}) = x^2\big(\dot{x}^2 + \lambda_3\big)$$

(que era lo que aparecía como integrando en la lagrangiana, usando $\lambda_0 = 1$). Por tanto, gracias a la observación 4.3.5 sabemos que existe $C \in \mathbb{R}$ tal que

$$x(t)^2\big(-\dot{x}(t)^2 + \lambda_3\big) = C, \qquad \forall\, t \in [0, T]. \tag{4.35}$$

Usando bien $x(0) = 0$ bien (4.34), deducimos $C = 0$. Obsérvese que entonces la ecuación (4.34) está contenida en (4.35) y por tanto ya no es necesaria. La ecuación (4.35) nos llevaría a que (recuérdese que $x \in C^1([0, T])$) o bien $x(t) = 0$ para todo $t \in [0, T]$ o bien $-\dot{x}(t)^2 + \lambda_3 = 0$ para todo $t \in [0, T]$. Como $x(T) = 1$, la primera condición no puede darse, luego

$$\dot{x}(t)^2 = \lambda_3, \qquad \forall\, t \in \mathbb{R}$$

y por tanto x describe una recta que debe pasar por $(0, 0)$ y $(T, 1)$, es decir

$$x(t) = \frac{t}{T}, \qquad \forall\, t \in [0, T].$$

La condición (4.33) nos da finalmente

$$T = 3, \qquad x(t) = \frac{t}{3}, \qquad \forall\, t \in [0, 3].$$

4.4.3. Problemas con datos convexos y restricciones de desigualdad

Este tipo de problemas se pueden estudiar usando el Teorema de Kuhn-Tucker. Consideramos el ejemplo

Resolver el problema

$$\text{mín}\left\{x(1) \,/\, x \in C^1([0, 1]),\, x(0) = 0,\, \int_0^1 \big(x(t)^2 + \dot{x}(t)^2\big)\, dt \le 1,\right\}$$

comprobando que la función encontrada es efectivamente solución.

Solución.- Definimos la lagrangiana

$$\mathcal{L}(x) = \mathcal{L}(x, \lambda_0, \lambda_1, \mu) = \lambda_0 x(1) + \lambda_1 \left(\int_0^1 \big(x(t)^2 + \dot{x}(t)^2\big)\, dt - 1\right) + \mu x(0).$$

Entonces, si x es solución del problema, existe $\lambda_0 \in \mathbb{R}$, $\lambda_1 \in \mathbb{R}$, $\mu \in \mathbb{R}$, no todos nulos tales que

$$\left\{\begin{array}{ll} a) & \mathcal{L}'(x)h = 0, \qquad \forall\, h \in C^1([0, 1]) \\ b) & \lambda_i \ge 0 \quad i = 0, 1 \\ c) & \lambda_1 \Big(\int_0^1 \big(x(t)^2 + \dot{x}(t)^2\big)\, dt - 1\Big) = 0, \end{array}\right. \tag{4.36}$$

Calculando esta derivada se obtiene entonces

$$\lambda_0 h(1) + 2\lambda_1 \int_0^1 \big(x(t)h(t) + \dot{x}(t)\dot{h}(t)\big)\,dt + \mu h(0) = 0, \quad \forall\, h \in C^1([0,1]), \tag{4.37}$$

lo que en particular prueba

$$\lambda_1 \int_0^1 \big(x(t)h(t) + \dot{x}(t)\dot{h}(t)\big)\,dt = 0, \qquad \forall\, h \in C_0^1([0,1]). \tag{4.38}$$

Por el lema de Du Bois-Raymond, se tiene entonces

$$\frac{d}{dt}\big(\lambda_1 \dot{x}(t)\big) = \lambda_1 x(t), \qquad \forall\, t \in [0,1]. \tag{4.39}$$

Usando esta ecuación e integrando por partes se tiene

$$\lambda_1 \int_0^1 \dot{x}(t)\dot{h}(t)\,dt = \lambda_1 \dot{x}(1)h(1) - \lambda_1 \dot{x}(0)\dot{h}(0) - \int_0^1 \frac{d}{dt}\big(\lambda_1 \dot{x}(t)\big)h(t)\,dt =$$

$$= \lambda_1 \dot{x}(1)h(1) - \lambda_1 \dot{x}(0)\dot{h}(0) - \lambda_1 \int_0^1 x(t)h(t)\,dt.$$

Por tanto, la condición (4.38), se reduce a

$$\big(\lambda_0 + 2\lambda_1 \dot{x}(1)\big)h(1) + \big(\mu - 2\lambda_1 \dot{x}(0)\big)h(0) = 0, \qquad \forall\, h \in C^1([0,1]),$$

lo que implica

$$\lambda_0 + 2\lambda_1 \dot{x}(1) = 0 \tag{4.40}$$

$$\mu - 2\lambda_1 \dot{x}(0) = 0. \tag{4.41}$$

Además tendremos que considerar las condiciones

$$x(0) = 0 \tag{4.42}$$

$$\int_0^1 \big(x(t)^2 + \dot{x}(t)^2\big)\,dt \le 1. \tag{4.43}$$

Ahora deberíamos eliminar algún multiplicador viendo que es distinto de cero y que por tanto lo podemos suponer igual a 1. En este caso, las ecuaciones (4.40) y (4.41) implican inmediatamente $\lambda_1 \neq 0$ y por tanto suponemos $\lambda_1 = 1$. En particular, la condición c) de (4.36) se escribe entonces

$$\int_0^1 \big(x(t)^2 + \dot{x}(t)^2\big)\,dt = 1. \tag{4.44}$$

Como en los otros ejemplos, podemos ahora olvidar la condición (4.41) que sólo serviría para calcular μ. Sin embargo, la ecuación (4.40) si que nos da alguna información ya que $\lambda_0 \ge 0$ y por tanto se tiene

$$\dot{x}(1) \le 0, \tag{4.45}$$

La ecuación (4.39), se convierte en la ecuación linea homogénea con coeficientes constantes de segundo orden

$$\ddot{x}(t) = x(t) \qquad \forall\, t \in [0, T].$$

Resolviendo la ecuación característica de esta ecuación ($r^2 = 1$), se deduce entonces que existen $C_1, C_2 \in \mathbb{R}$ tales que

$$x(t) = C_1 e^t + C_2 e^{-t}, \qquad \forall\, t \in [0, 1].$$

las condiciones (4.42), (4.44) permiten ahora calcular C_1 y C_2 obteniéndose

$$C_1 = \frac{1}{\sqrt{e^2 - e^{-2}}},\ C_2 = -\frac{1}{\sqrt{e^2 - e^{-2}}} \quad \text{o} \quad C_1 = -\frac{1}{\sqrt{e^2 - e^{-2}}},\ C_2 = \frac{1}{\sqrt{e^2 - e^{-2}}}.$$

La condición (4.45) nos dice que la segunda condición es la correcta, se tiene por tanto

$$x(t) = -\frac{e^t - e^{-t}}{\sqrt{e^2 - e^{-2}}}, \qquad \forall\, t \in [0, 1].$$

Para probar que esta función es efectivamente solución, basta tener en cuenta que (4.40) implica

$$\lambda_0 = -\frac{e + e^{-1}}{\sqrt{e^2 - e^{-2}}} > 0,$$

y por tanto, el Teorema de Kuhn-Tucker nos dice que x es efectivamente solución.

Bibliografía

[1] V. M. Alexeev-V.M. Tikhomirov-S.V. Fomin: *Optimal Control,* Consultants Bureau, (1987).

[2] H. Brézis: *Análisis Funcional. Teoría y aplicaciones,* Alianza Editorial, (1984).

[3] J. Céa: *Optimisation. Théorie et algorithmes,* Dunod, Gauthiers-Villars, (1969).

[4] I. M. Gelfand-S.V. Fomin: *Calculus of Variations,* Prentice-Hall, (1963).

[5] A. D. Ioffe-V. Tikhomirov: *Theory of extremal problems,* North-Holland, (1979).

[6] J.Jost-X. Li-Jost: *Calculus of Variations,* Cambridge Univ. Press, (2003).

[7] M. L. Krasnov-G. I. Makarenko-A. I. Kiseliov: *Cálculo variacional (ejemplos y problemas)*, Mir, (1976).

[8] E. Kreiszig: *Introductory Functional Analysis and applications,* John Wiley and Sons, (1978).

[9] P.L. Lax *Functional Analysis,* John Wiley & Sons, (2002).

[10] L. A. Liusternik-V. J. Sobolev: *Elements of Functional Analysis,* F. Ungar Publ. Co., (1965).

[11] L. P. Lebedev-M. J. Cloud: *The Calculus of Variations and Functional Analysis,* World Sci. Pub Co. , (2003).

[12] D. G. Luenberger: *Optimization by vector space methods,* John Wiley and Sons, (1969).

[13] J. T. Oden-L. F. Demkowicz: *Applied Functional Analysis,* CRC Press, (1996).

[14] E. R. Pinch: *Optimal control and the calculus of variations,* Oxford Science Publications, (1993).

[15] V. M. Tikhomirov: *Fundamental Principles of the Theory of Extremal Problems,* John Wiley and Sons, (1986).

[16] V. A. Trenoguin: *Analyse Fonctionnelle,* Mir, (1985).

[17] V. A. Trenoguin-B. M. Pisarievski-T. S. Sóboleva: *Problemas y ejercicios de Análisis Funcional,* Mir (1987).

[18] E. Zeidler: *Nonlinear Functional Analysis and its applications III: Variational methods and optimization,* Springer-Verlag, (1985).