Introducción a la optimización abstracta

LX Reunión de la UMA - Tandil -2010

Pablo A. Lotito

plotito@exa.unicen.edu.ar

1. Introducción

En este curso trataremos de extender los resultados básicos de optimización no lineal en $I\!R^n$ a espacios de dimensión infinita, en particular a espacios de Banach reflexivos. Para un problema de optimización general el esquema de análisis es el siguiente:

1. Analizar existencia (y unicidad)
2. Obtener condiciones de optimalidad
3. Analizar la sensibilidad

De manera análoga se organizará este curso.

El problema de optimización más general lo podemos escribir como

$$(P) \begin{cases} \min\limits_{x} & f(x) \\ \text{s.t.} & x \in K \end{cases} \tag{1}$$

donde $f : X \to I\!R$ es una función definida en un conjunto X a valores reales y K es un subconjunto de X que define los valores admisibles.

El problema consiste en encontrar un elemento x^* del conjunto X que verifique

$$f(x^*) \leq f(x), \quad \forall x \in K. \tag{2}$$

Llamaremos **conjunto solución** de P y notaremos $\mathrm{Sol}(P)$ al conjunto de los x^* que verifiquen (2). Llamaremos **valor** de P y notaremos $\mathrm{val}(P)$ a $\inf\{f(x) : x \in K\}$.

Para explicitar un poco más el conjunto factible, a veces escribiremos

$$(P) \begin{cases} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & G(x) \in K \end{cases} \tag{3}$$

donde $G : X \to Y$ es una aplicación y $K \subset Y$.

Ejemplos

1. Si $f : I\!R^n \to I\!R$, es lineal y K un poliedro en $I\!R^n$ se trata de un **problema de optimización lineal**. En este caso existe un vector $c \in I\!R^n$ una matriz A de orden $n \times p$ y un vector $b \in I\!R^p$ y el problema se escribe

$$\begin{array}{ll} \min\limits_x & c^t x \\ \text{s.t.} & Ax \leq b \end{array} \tag{4}$$

2. Si $f : I\!R^n \to I\!R$, es convexa y K un convexo de $I\!R^n$ se trata de un **problema de optimización convexa**. Si K está definido a través de la función convexa g como $K = \{g \leq 0\}$ se tiene

$$\begin{array}{ll} \min\limits_x & f(x) \\ \text{s.t.} & g(x) \leq 0 \end{array} \tag{5}$$

3. Si $f : I\!R^n \times U \to I\!R$, y se estudia la dependencia de la solución del problema

$$(P_u) \left\{ \begin{array}{ll} \min\limits_x & f(x,u) \\ \text{s.t.} & x \in K \end{array} \right. \tag{6}$$

   para cada valor de $u \in U$ fijo, se tiene un **problema de optimización paramétrica**.

4. Consideremos $f : I\!R^n \to I\!R$, la función $G : I\!R^n \to \mathcal{S}_p$, a valores en el conjunto $\mathcal{S}_p$ de matrices simétricas de orden p, notemos $\mathcal{S}_p^+$ al subconjunto de matrices semidefinidas positivas se tiene entonces un **problema de programación semidefinida:**

$$\begin{array}{ll} \min\limits_x & f(x) \\ \text{s.t.} & G(x) \in \mathcal{S}_p^+. \end{array}$$

5. Consideremos $f : I\!R^n \to I\!R$, la función $G : I\!R^n \to \mathcal{C}(\Omega)$, a valores en el conjunto $\mathcal{C}(\Omega)$ de funciones continuas sobre el compacto Ω y sea $\mathcal{C}_+ = \{f \in \mathcal{C}(\Omega) : f \geq 0\}$ se puede considerar entonces el **problema de programación semi-infinita** que consiste en resolver

$$\begin{array}{ll} \min\limits_x & f(x) \\ \text{s.t.} & g(x,\omega) \leq 0, \forall \omega \in \Omega \end{array}$$

   se puede escribir de manera abstracta como

$$\begin{array}{ll} \min\limits_x & f(x) \\ \text{s.t.} & G(x) \in \mathcal{C}_+. \end{array}$$

6. Si se considera un conjunto X cuyos elementos son funciones y se busca minimizar un funcional $J : X \to I\!R$ (que en general depende de la función y sus derivadas) se tiene un problema de cálculo variacional, como por ejemplo el problema de la braquistocrona donde $X = \{x \in C^1[0,1] : x(0) = 0, x(1) = 1\}$ y $J(x) = \int_0^1 \sqrt{\frac{1+\dot{x}^2}{1-x}} dx.$

7. Si se considera un conjunto X de funciones $u(t)$ y el funcional no sólo está evaluado en las funciones de X sino en trayectorias definidas a través de ecuaciones diferenciales que también dependen de dichas funciones se tiene un problema de control óptimo, como por ejemplo
$$\min_x \int_0^1 L(u(t), y(t))dt + \varphi(y(1))$$
donde y verifica $y' = f(u, y)$, $y(0) = y_0$. Dado que para cada x existe (bajo ciertas hipótesis) un único y se puede definir el funcional $f(x) = \int_0^1 L(u(t), y(t))dt + \varphi(y(1))$. Análogamente se puede considerar el control óptimo de EDP's donde ahora X es un conjunto de funciones $u(x)$ e y es solución de una EDP como por ejemplo
$$\begin{aligned} -\Delta y + \phi(y) &= u, \quad \forall x \in \operatorname{int}\Omega \\ y &= y_0, \quad \forall x \in \partial\Omega. \end{aligned}$$

2. Existencia y unicidad

Consideremos primero el problema para X de dimensión finita,

$$(\mathcal{P}) \begin{cases} \min\limits_x & f(x) \\ \text{s.t.} & x \in K \end{cases} \tag{7}$$

donde K es un subconjunto cerrado de X. Tenemos entonces el siguiente teorema:

Theorem 2.1 (Existencia en dim finita)**.** *Sea $f : X \subset I\!R^n \to \bar{I\!R}$ semicontinua inferiormente y coerciva. Si K es cerrado y $\operatorname{dom}(f) \cap K \neq \emptyset$, entonces el problema $\mathcal{P}$ tiene al menos una solución.*

Antes de demostrar el teorema recordemos que:

- $\bar{I\!R}$ es el conjunto de los reales con $+\infty$.
- $\operatorname{dom}(f) = \{x \in X : f(x) < +\infty\}$.
- f es **semicontinua inferiormente** o abreviadamente **s.c.i** en un punto x^* si se verifica que
$$\lim_{x \to x^*} f(x) \geq f(x^*) \tag{8}$$
- f es **semicontinua inferiormente** si lo es en cada punto de su dominio.
- si definimos el **epigrafo** de f como épi$(f) = \{(x, y) : y \geq f(x), \forall x \in \operatorname{dom}(f)\}$ se tiene que f es s.c.i. si y sólo si épi(f) es cerrado.
- decimos que f es **coerciva** si se verifica que $\lim_{\|x\| \to \infty} f(x) = +\infty$.
- decimos que $\{x_k\}$ es una sucesión minimizante (o minimizadora) para $\mathcal{P}$ si $\lim f(x_k) = \inf\{f(x) : x \in K\} = val(\mathcal{P})$.

Demostración. Gracias a las hipótesis podemos asegurar la existencia de una sucesión minimizante (ejercicio). La coercividad de f implica que esta sucesión es acotada y está contenida en K. Luego tiene una subsucesión convergente en K (ejercicio). Gracias a la s.c.i. de f se concluye que dicho límite es solución del problema. □

¿Cómo se extiende este teorema a un espacio de dimensión infinita? Inmediatamente no es posible ya que no se puede utilizar la compacidad de los cerrados y acotados. Deberemos relajar la topología. Sea ahora el problema planteado sobre X un espacio de Banach reflexivo (un Hilbert por ejemplo). Podemos utilizar que los convexos cerrados y acotados son débilmente compactos, pero entonces necesitaremos la semicontinuidad inferior en la topología relajada (débil) o por lo menos secuencialmente. Una topología débil tiene menos abiertos y más sucesiones convergentes y también más compactos. Sin embargo, las funciones continuas (o s.c.i) fuertemente pueden dejar de serlo en la topología más débil. La topología débil que usaremos no es metrizable, por eso usaremos también la noción de secuencialmente cerrado y secuencialmente s.c.i.

Con estas nociones el teorema de existencia en dimensión finita se extiende a:

Theorem 2.2 (Existencia en Banach reflexivo)**.** *Sea X un espacio de Banach reflexivo, $K \subset X$ un conjunto secuencialmente débilmente cerrado no vacío y sea $f : X \subset I\!R^n \to \bar{I\!R}$ propia, débilmente semicontinua inferiormente y coerciva tal que $dom(f) \cap K \neq \emptyset$. Entonces el problema $\mathcal{P}$ tiene al menos una solución.*

Antes de la demostración aclaremos algunas definiciones:

- Sea X un e.v.t. (un Banach por ejemplo) se denomina dual (topológico) de X y se nota X^* al conjunto de funcionales lineales y continuos de X en $I\!R$, es decir que

$$X^* = \{f : X \to I\!R : f \text{ lineal y continua}\}.$$

- Una sucesión $\{x_k\}$ converge débilmente a $\bar{x}$ en un espacio de Banach X, lo cual se nota $x_k \rightharpoonup \bar{x}$, si vale

$$f(x_k) \to f(\bar{x}), \quad \forall f \in X^*.$$

  Toda sucesión que converge fuertemente (en norma) converge débilmente. Toda sucesión que converge débilmente está acotada y vale

$$x_k \rightharpoonup \bar{x} \Rightarrow \|\bar{x}\| \leq \liminf \|x_k\|.$$

- En un e.v.t. X se define la topología débil como la mínima que hace continuas a las funciones de X^*. Esta topología puede ser generada considerando la familia $\{f^{-1}(A) : A \text{ abierto de } I\!R, f \in X^*\}$, o más simplemente una base de entornos del cero dada por $\{x : |f_i(x)| < \varepsilon, i = 1, \ldots, n\}$ donde n es un natural (finito) y $f_i \in X^*$. En general es una topología no metrizable (no localmente acotada ya que contiene el núcleo trasladado de f) que implica la convergencia débil secuencial (pero no al revés).

- Al tener la topología débil menos abiertos, los abiertos (y cerrados) débiles son abiertos (y cerrados) fuertes. En general no vale la recíproca (trivial dimensión finita), sin embargo para los conjuntos convexos ser cerrado fuerte o débil es equivalente. Fácilmente vemos que como f es s.c.i. si y sólo si $epi(f)$ es cerrado, si f es además convexa entonces f es débilmente s.c.i. y por lo tanto débilmente secuencialmente s.c.i. (ver lema posterior)

- Un espacio de Banach se dice reflexivo cuando la inyección canónica $J : X \to X^{**}$, que a cada x le asigna $\xi_x \in X^{**}$ tal que $\xi_x(f) = f(x), \quad \forall f \in X^*$, es sobreyectiva.

- Los espacios L^p con $1 < p < \infty$, que son los espacios de funciones medibles f tales que $|f|^p$ es integrable, son reflexivos. L^1 no es reflexivo pero es separable, L^∞ no es reflexivo ni separable.

- Los espacios de Hilbert son un caso particular de Banach reflexivo, en particualr L^2 es un espacio de Hilbert.

- En un espacio reflexivo, convexo cerrado y acotado implica débilmente compacto.

Demostración. La prueba es una extensión inmediata de la demostración del teorema anterior. Consideremos una sucesión minimizante, es decir una sucesión$\{x_k\} \subset K$ tal que $f(x_k) \to \mathrm{val}(P)$ (esta convergencia es en $I\!R$) gracias a la coercividad esta sucesión es acotada y está contenida en K. Consideremos la cápsula convexa de dicha sucesión que será (convexa) cerrada y acotada (usar desigualdad triangular) luego débilmente compacta. Existe entonces una subsucesión débilmente convergente a un punto de K (porque es débilmente sec. cerrado) y gracias a la d.s.s.c.i. se concluye. □

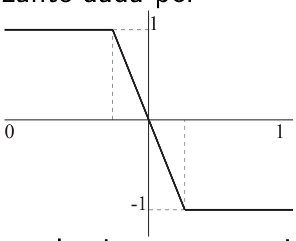
Remark 2.3. En general la condición de reflexividad no puede relajarse sin agregar hipótesis. Por ejemplo, consideremos el espacio de funciones continuas $C[0,1]$ y el subconjunto convexo y cerrado

$$K = \{u \in C[0,1] : \int_0^{\frac{1}{2}} u(t)dt - \int_{\frac{1}{2}}^1 u(t)dt = 1\}$$

Consideremos el funcional $f(u) = \|u\|_\infty$ que es trivialmente continuo y coercivo. Fácilmente se ve que

$$f(u) = \|u\|_\infty \geq \int_0^1 |u(t)|dt \geq \int_0^{\frac{1}{2}} u(t)dt - \int_{\frac{1}{2}}^1 u(t)dt = 1 \tag{9}$$

con lo cual $\inf_{u\in K} f(u) \geq 1$ pero se puede demostrar la igualdad ya que existe una sucesión minimizante dada por



$$u_k(x) = \begin{cases} 1 & \text{si } x \in [1, \frac{1}{2} - \frac{1}{k}] \\ -kx + \frac{k}{2} & \text{si } x \in [\frac{1}{2} - \frac{1}{k}, \frac{1}{2} + \frac{1}{k}] \\ -1 & \text{si } x \in [\frac{1}{2} + \frac{1}{k}, 1] \end{cases} \tag{10}$$

Con lo cual existe una sucesión minimizante pero no existe solución ya que ninguna función continua u puede verificar a la vez $\|u\|_\infty = 1$ y $u \in K$.

Veamos ahora cómo la hipótesis suplementaria de convexidad para la función f permite evitar el uso de topología débil.

Diremos que una función $f : X \to \bar{I\!R}$ propia [1] definida sobre un espacio normado X es:

- **convexa:** si $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \ \ \forall\alpha \in (0,1), \ \forall x, y \in \mathrm{dom}(f)$.
- **estrictamente convexa:** si vale la desigualdad estricta, es decir, $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y), \ \ \forall\alpha \in (0,1), \ \forall x, y \in \mathrm{dom}(f)$.
- **fuertemente convexa:**si existe $\beta \in I\!R_+$ tal que

  $$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) - \tfrac{1}{2}\beta\alpha(1-\alpha)\|x-y\|^2,$$

  para todo $\alpha \in (0,1), \ \forall x, y \in \mathrm{dom}(f)$.

Primero veamos este resultado de existencia y unicidad para f fuertemente convexa que no depende de la dimensión de X (pero es menos general ya que pide más hipótesis sobre el criterio).

Theorem 2.4 (Existencia y unicidad)**.** *Sea $f : X \to \bar{I\!R}$ fuertemente convexa y s.c.i. y sea K un convexo cerrado de X tales que $\mathrm{dom}(f) \cap K \neq \emptyset$. Entonces f alcanza su mínimo sobre K en un único punto.*

Demostración. Primero veamos que la condición $\mathrm{dom}(f) \cap K \neq \emptyset$, implica que $-\infty < \mathrm{val}(\mathcal{P}) + \infty$. Para eso tomemos $x_0 \in \mathrm{dom}(f) \cap K$ e $y_0 := f(x_0) < +\infty$ (y obviamente mayor que $-\infty$). Como (x_0, y_0) está en el gráfico de f, $z_0 := (x_0, y_0 - 1)$ no pertenece a épi(f) que es un convexo cerrado en X por ser f convexa s.c.i., luego puede separarse estrictamente por un hiperplano (t. de H.-B.) y tenemos que existe un funcional $\xi = (\xi_x, \xi_y)$ tal que

$$\langle \xi, z_0 \rangle < \langle \xi, z \rangle, \quad \forall z \in \text{épi}(f),$$

luego se tiene que para todo $x \in \mathrm{dom}(f)$,

$$\langle \xi_x, x_0 \rangle + \langle \xi_y, y_0 - 1 \rangle \leq \langle \xi_x, x) \rangle + \langle \xi_y, y \rangle, \quad \forall (x,y) : x \in \mathrm{dom}(f), y \geq f(x). \tag{11}$$

Es fácil ver que $\xi_y > 0$ (ejercicio) y de la ecuación 11 se tiene que

$$f(x) \geq \langle \frac{\xi_x}{\xi_y}, x_0 - x \rangle + y_0 - 1. \tag{12}$$

Usando (12) y la fuerte convexidad con parámetro β y $\alpha = \frac{1}{2}$ se tiene que

$$f(x) \geq 2(y_0 - 1) - \frac{\|\xi_x\|}{\xi_y}\|x_0 - x\| + \frac{1}{4}\beta\|x_0 - x\|^2. \tag{13}$$

Tomando ínfimo en ambos miembros se tiene a la derecha que minimizar una función cuadrática convexa de variable real que obviamente tiene un mínimo finito en el cero de su derivada. Se prueba así que $\mathrm{val}(\mathcal{P}) > -\infty$. Sobre X y con más razón sobre K.

[1] Una función es propia si su dominio es no vacío y nunca toma el valor $-\infty$. Una función convexa que asume el valor $-\infty$ es idénticamente $-\infty$ por eso las excluimos de la definición.

Lo anterior nos garantiza la existencia de una sucesión minimizante $\{x_k\}$ que demostraremos que es de Cauchy, con lo cual, al estar en un espacio completo (Banach), convergirá en K (porque éste es cerrado) y gracias a la propiedad s.c.i. será la solución buscada.

Veamos entonces que la sucesión es de Cauchy, para eso utilizamos que f es fuertemente convexa y entonces podemos acotar la diferencia $\|x_k - x_j\|$ tomando $\alpha = \frac{1}{2}$ como

$$\frac{1}{8}\beta\|x_k - x_j\|^2 \leq \tfrac{1}{2}f(x_k) + \tfrac{1}{2}f(x_j) - f(\frac{x_k + x_j}{2}). \tag{14}$$

Pero sabemos que $f(x_k) \to \operatorname{ínf}_{x\in K} f(x) = \mathrm{val}(\mathcal{P})$ que para abreviar llamaremos $z := \mathrm{val}(\mathcal{P})$. Luego en la ecuación (14) tenemos

$$\frac{1}{8}\beta\|x_k - x_j\|^2 \leq \tfrac{1}{2}f(x_k) + \tfrac{1}{2}f(x_j) - z. \tag{15}$$

que podemos escribir como

$$\frac{1}{8}\beta\|x_k - x_j\|^2 \leq \tfrac{1}{2}(f(x_k) - z) + \tfrac{1}{2}(f(x_j) - z). \tag{16}$$

Como el miembro de la derecha tiende a cero, queda demostrado que la sucesión minimizante es de Cauchy. □

Este teorema válido para todo espacio de Banach tiene como corolario el teorema de representación de Riesz para espacios de Hilbert.

Corollary 2.5. *Sea X un espacio de Hilbert y x^* un elemento de su dual entonces existe un único $\bar{x}$ tal que $\langle x^*, x\rangle = (\bar{x}, x), \quad \forall x \in X$.*

Demostración. Es fácil ver la unicidad.

Para ver la existencia notemos que la función $f : X \to I\!R$ definida como $f(x) := \frac{1}{2}\|x\|^2 - \langle x^*, x\rangle$ es fuertemente convexa. Para verlo basta escribir la definición de fuerte convexidad (con $\beta = 2$) y utilizar la ley del paralelogramo para reescribir el producto interno.

Además, f es s.c.i. por ser continua y es propia, luego tiene mínimo en X por teorema anterior, con lo cual existe $\bar{x}$ tal que $f(\bar{x}) \leq f(x), \quad \forall x \in X$. Entonce para $x \in X$, se verifica que $f(\bar{x}) \leq f(\bar{x} + \varepsilon x), \ \forall \varepsilon \in I\!R$ luego

$$\tfrac{1}{2}\|\bar{x}\|^2 - \langle x^*, \bar{x}\rangle \leq \tfrac{1}{2}\|\bar{x} + \varepsilon x\|^2 - \langle x^*, \bar{x} + \varepsilon x\rangle \tag{17}$$

y simplificando tenemos

$$\tfrac{1}{2}\|\bar{x}\|^2 - \tfrac{1}{2}\|\bar{x} + \varepsilon x\|^2 \leq -\varepsilon\langle x^*, x\rangle \tag{18}$$

desarrollando la norma de la suma tenemos

$$-\varepsilon(\bar{x}, x) - \tfrac{1}{2}\varepsilon^2\|x\|^2 \leq -\varepsilon\langle x^*, x\rangle \tag{19}$$

Si consideramos $\varepsilon > 0$ podemos dividir y obtener

$$(\bar{x}, x) + \tfrac{1}{2}\varepsilon|x\|^2 \geq \langle x^*, x\rangle \tag{20}$$

hacemos tender ε a cero obtenemos que $(\bar{x}, x) \geq \langle x^*, x\rangle$. Considerando $\varepsilon < 0$ se demuestra la desigualdad contraria y queda demostrado el corolario. □

El teorema 2.4 garantiza existencia y unicidad sin topologías débiles pero con la condición de convexidad fuerte que frecuentemente no se tiene (por ejemplo criterios lineales). Es necesario entonces aplicar el teorema 2.2 y para eso verificar que la función del criterio sea d.s.s.c.i. En general una función continua (s.c.i.) fuerte no lo es en el sentido débil, el siguiente teorema da una condición suficiente para que una función s.c.i. fuerte lo siga siendo en el sentido débil por sucesiones.

Theorem 2.6. *Sea X un espacio de Banach y $f : X \to I\!R$ convexa y s.c.i. Entonces f es d.s.s.c.i.*

Demostración. Primero veamos que una función $f : X \to I\!R$ de un espacio topológico metrizable X a valores reales es s.c.i. si y sólo si su epigrafo es cerrado en la topología producto. Si $x_k \to \bar{x}$ y $\liminf f(x_k) = z$ existe una subsucesión que notamos $x_{k'}$ tal que $\lim f(x_k) = z$ y $x_{k'} \to \bar{x}$ tenemos entonces que el par $(x_{k'}, f(x_{k'})) \to (\bar{x}, z)$ y si el epigrafo de f es cerrado se tiene que $(\bar{x}, z) \in$ épi(f) con lo cual $f(\bar{x}) \leq \liminf f(x_k)$. Supongamos ahora que $(x_k, z_k) \to (\bar{x}, \bar{z})$ con $z_k \geq f(x_k), \forall k$ entonces por la topología producto $x_k \to \bar{x}$ y $z_k \to \bar{z}$ si f es s.c.i. se tiene que $\bar{z} = \lim z_k \geq \liminf f(z_k) \geq f(\bar{x})$ y por lo tanto $(\bar{x}, \bar{z}) \in$ épi(f) con lo cual el epigrafo es cerrado.

Veamos ahora el caso general para X espacio topológico. Sea entonces épi(f) cerrado y supongamos que existe $\bar{x} \in X$ tal que f no es s.c.i. en $\bar{x}$, es decir que

$$f(x) > z := \liminf_{x \to \bar{x}} f(x) = \sup_{V \in \mathcal{V}_{\bar{x}}} \inf_{x \in V} f(x)$$

donde $\mathcal{V}_{\bar{x}}$ es una base de entornos de $\bar{x}$. Entonces como $z < f(\bar{x})$ se tiene que el par $(\bar{x}, z) \notin$ épi(f) y por ser éste último cerrado existe un entorno (en la topología producto de $X \times I\!R$) que lo separa de épi(f). Tenemos entonces que existe $V \in \mathcal{V}_{\bar{x}}$ y $\delta > 0$ tales que $x \in V, y \in (z - \delta, z + \delta) \Rightarrow (x, y) \notin$ épi(f) con lo cual $\liminf_{x \to \bar{x}} f(x) \geq z + \delta > z$ que es un absurdo. Para ver el recíproco supongamos que f es s.c.i. y que $(\bar{x}, z) \notin$ épi(f) se tiene entonces que $f(x) > z)$ y por la s.c.i. también

$$\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x}) > z$$

por definición de límite inferior, existe entonces un entorno V de $\bar{x}$ donde $\inf_{x \in V} f(x) \geq f(\bar{x}) - \varepsilon > z + \varepsilon$ con $\varepsilon > 0$. Pero entonces el conjunto $V \times (z - \varepsilon, z + \varepsilon)$ es un entorno de $(\bar{x}, z)$ no contenido en épi(f). Esto prueba que es cerrado.

Demostremos ahora el teorema. Si f es convexa y s.c.i. su epigrafo es convexo y fuertemente cerrado con lo cual también es cerrado débil y también cerrado débil por sucesiones. Por lo probado anteriormente tenemos que en el caso de funciones convexas la noción de s.c.i. coincide para la topología fuerte, la débil y ambas implican la s.c.i. por sucesiones (débil y fuerte).

□

Un resultado intermedio existe para funciones cuasi convexas. Se dice que f es cuasi convexa si los conjuntos de nivel inferior $L_\alpha(f) := \{x \in X : f(x) \leq \alpha\}$ son convexos.

Theorem 2.7. *Sea X un espacio de Banach, $K \subset X$ un conjunto secuencialmente débilmente cerrado no vacío y sea $f : X \subset I\!R^n \to \bar{I\!R}$ cuasi convexa y semicontinua inferiormente* *tal que $L_\alpha(f) \cap K$ es no vacío y acotado para algún $\alpha \in I\!R$. Entonces el problema $\mathcal{P}$ tiene al menos una solución y el conjunto de soluciones es acotado.*

Demostración. Sea $x_k \rightharpoonup \bar{x}$ minimizante en K. Podemos suponer que $\alpha_k := f(x_k)$ es decreciente, entonces tenemos que $\bar{x} \in L_{\alpha_k}(f)$ para todo k. Pero además los conjuntos $L_{\alpha_k}(f)$ son convexos (cuasi convexidad) y cerrados (s.c.i.) por lo que son también débilmente cerrados (y por sucesiones) entonces $\bar{x} \in L_{\alpha_k}(f)$ para todo k, con lo que $f(\bar{x}) \leq f(x_k)$ para todo k, y por lo tanto

$$f(\bar{x}) \leq \lim f(x_k) = \inf_{x \in K} f(x).$$

□

Remark 2.8. En el caso de espacios de Banach reflexivos pudimos demostrar existencia sin necesidad de nociones de convexidad. En el caso de un espacio no reflexivo (L^∞ por ejemplo) será necesaria alguna condición de convexidad para poder obtener resultados. Es más, se puede demostrar por ejemplo, que en el caso del cálculo de variaciones, el funcional $J : W^{1,p}(\Omega) \to I\!R$ definido como

$$J(u) := \int_\Omega f(x, u(x), \nabla u(x))dx \tag{21}$$

es débilmente s.c.i. **si y sólo si** f es convexa en la tercera variable. El espacio $W^{1,p}$ es el de las funciones de L^∞ cuya derivada (en el sentido de las distribuciones) está en L^p. Más precisamente se tiene el siguiente teorema

Theorem 2.9. *Sea $\Omega \subset I\!R^n$ un abierto acotado y sea $f : \Omega \times I\!R^m \times I\!R^{nm} \to I\!R$ una función continua que satisface*

$$|f(x, u, \xi)| \leq a(x, |u|, |\xi|) \tag{22}$$

*pata todo $(x, u, \xi) \in \Omega \times I\!R^m \times I\!R^{nm}$, donde a es creciente con respecto a u y $|\xi|$ y localmente integrable en x. Si $n = 1$ o $m = 1$ y el funcional $I(u) = \int_\Omega f(x, u(x), \nabla u(x))dx$ es *-débilmente inferiormente semicontinuo en $W^{1,\infty}(\Omega, I\!R^m)$, entonces f es convexa en su tercera variable.*

Como toda función débilmente s.c.i en $W^{1,p}$ es *-débilmente s.c.i. en $W^{1,\infty}$ se ve que la convexidad es necesaria para la d.s.c.i. en $W^{1,p}$ para todo p. Con hipótesis adicionales sobre f se muestra el recíproco en $W^{1,p}$.

3. Condiciones de optimalidad

Lo visto anteriormente permite garantizar la existencia de soluciones pero no brinda ideas para el cálculo de mínimos o minimizadores. La definición de mínimo es una condición sobre todos los puntos de X con lo cual es imposible verificar directamente (probando con todos los puntos). Afortunadamente existen condiciones necesarias que permiten reducir la búsqueda y a veces condiciones suficientes que simplifican la tarea.

Consideremos primero la búsqueda de mínimo de un funcional sin restricciones. Es dir sea $f : X \to I\!R$ de un e.v.t. X en los reales y consideremos el problema

$$(\mathcal{P}) \min_{x \in X} f(x) \tag{23}$$

Las condiciones de optimalidad en el caso de $X = I\!R^n$ y f diferenciable son bien conocidas. Sabemos que si $X = I\!R^n$ y f es diferenciable en un punto x^* que minimiza f localmente entonces $\nabla f(x^*) = 0$.

Podemos demostrar el siguiente resultado más general:

Theorem 3.1. *Sea X un espacio de Banach, $f : X \to I\!R$ Gateaux diferenciable y sea $\bar{x}$ un minimizador de f en X, entonces $Df(\bar{x}) = 0$. Si f es sólamente directionalmente diferenciable, se tiene que $f'(\bar{x}; v) \geq 0,\ \forall v \in X$.*

Antes de demostrarlo aclaremos noción de diferenciable Gâteaux.

Definition 3.2 (Diferenciabilidad)**.** Sea $f : X \to Y$ con X e Y espacios de Banach.

1. f es **direccionalmente** derivable en x en la dirección $v \in X$ si existe el límite
$$f'(x, v) := \lim_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t}. \tag{24}$$
2. f es **Gateaux** diferenciable en x si para toda dirección existe la derivada direccional y la aplicación $v \mapsto f'(x, v)$ es linear y continua (en v). Se nota $Df(x)$ y es un elemento de $L(X, Y)$. Por lo tanto, cuando f es Gateaux diferenciable se tiene que $f'(x, v) = \langle Df(x), h\rangle$ que abusando del lenguaje denotaremos con $Df(x)h$.

Notemos que si f es Gateaux diferenciable entonces $Df(x)v$ puede calcularse como la derivada clásica de la función $f(x + \alpha v)$ respecto de α evaluada en 0, es decir que

$$Df(x)v = \frac{d}{d\alpha} f(x + \alpha v)\bigg|_{\alpha=0}. \tag{25}$$

Veamos algunos ejemplos de derivación:

- Consideremos $\varphi : L^\infty(0, 1) \to I\!R$ tal que $\varphi(u) = \int_0^1 u^2(t)dt$ aplicando la definición tenemos que

$$\varphi'(u; v) = \lim_{\alpha \downarrow 0} \int_0^1 \frac{(u(t) + \alpha v(t))^2 - u^2(t)}{\alpha} dt = \lim_{\alpha \downarrow 0} \int_0^1 \frac{\alpha v(t)(2u(t) + \alpha v(t))}{\alpha} dt \tag{26}$$

luego

$$\varphi'(u;v) = \int_0^1 2u(t)v(t)dt \tag{27}$$

Además es fácil ver que para cada u la aplicación $\varphi'(u;v)$ es lineal y continua en v ya que

$$|\varphi'(u;v)| \leq 2\|u\|_\infty \|v\|_\infty \tag{28}$$

- Sea $\Phi : L^\infty(0,1) \to: W^{1,\infty}(0,1)$ definida como $\Phi(u) = y_u(t)$ donde y_u es solución de la ODE (parametrizada en u)

$$\begin{cases} y' = y + u \\ y(0) = 1 \end{cases}$$

Si aplicamos la definición tenemos que calcular

$$\Phi'(u;v) = \lim_{\alpha \downarrow 0} \frac{y_{u+\alpha v} - y_u}{\alpha} \tag{29}$$

pero es fácil ver que el numerador verifica la ODE con condiciones homogéneas

$$\begin{cases} y' = y + \alpha v \\ y(0) = 0 \end{cases}$$

y por lo tanto

$$\Phi'(u;v) = z_v \quad \text{donde } z_v \text{ es solución de} \begin{cases} y' = y + v \\ y(0) = 0 \end{cases} \tag{30}$$

Para analizar la diferenciabilidad Gateaux, o la continuidad de la aplicación, en este caso sencillo podemos integrar la ODE y tenemos que

$$\Phi'(u;v) = z_v = \int_0^t e^{(t-\tau)} v(\tau) d\tau \tag{31}$$

luego

$$|\Phi'(u;v)(t)| \leq \int_0^t e^{(t-\tau)} d\tau \|v\|_\infty \leq (e-1)\|v\|_\infty$$

y (usando la ODE para z_v)

$$|\frac{d}{dt}\Phi'(u;v)(t)| \leq (e-1)\|v\|_\infty + \|v\|_\infty = e\|v\|_\infty$$

por lo tanto tenemos que

$$\|\Phi'(u;v)\|_{1,\infty} \leq (2e-1)\|v\|_\infty \tag{32}$$

- Consideremos ahora $\varphi : L^\infty(0,1) \to I\!R$ tal que $\varphi(u) = \int_0^1 y^2(t)dt$ donde y es solución de

$$\begin{cases} y' = y + u \\ y(0) = 1 \end{cases}$$

aplicando la definición tenemos que

$$\varphi'(u;v) = \lim_{\alpha\downarrow 0}\int_0^1 \frac{y_{u+\alpha v}^2(t) - y_u^2(t)}{\alpha}dt = \lim_{\alpha\downarrow 0}\int_0^1 \frac{y_{u+\alpha v}(t) - y_u(t)}{\alpha}(y_{u+\alpha v}(t) + y_u(t))dt \tag{33}$$

donde hemos llamado y_u a la solución de le ODE con parámetro u. Vemos que necesitamos calcular la derivada direccional de la aplicación del item anterior, tenemos entonces que

$$\varphi'(u;v) = \int_0^1 2y_u z_v dt. \tag{34}$$

Ahora, ¿será la derivada direccional lineal en v? se puede demostrar que sí, especialmente analizando la EDo que defin a z_v que es un sistema lineal con condiciones homogéneas. Para ver más claramente que la aplicación $\varphi'(u;v)$ es lineal en v quisiéramos poder escribir (34) como

$$\varphi'(u;v) = \int_0^1 \psi v dt.$$

Para poder escribirlo de ese modo utilizamos una función auxiliar $p(t)$ y la siguiente identidad que surge de la regla de integración por partes:

$$\int_0^1 p'(t)z_v(t)dt = [p(t)z_v(t)]_0^1 - \int_0^1 p(t)z_v'(t)dt, \tag{35}$$

como z_v es solución de la ODE anterior, tenemos que

$$\int_0^1 p'(t)z_v(t)dt + \int_0^1 p(t)(z_v(t) + v(t))dt - p(1)z_v(1) = 0. \tag{36}$$

Sumando la identidad (36) en la ecuación (34) tenemos

$$\varphi'(u;v) = \int_0^1 (2y_u + p'(t) + p(t))z_v(t)dt + \int_0^1 p(t)v(t)dt - p(1)z_v(1). \tag{37}$$

Si elegimos p de modo que resuelva la EDO

$$\begin{cases} p' = -p - 2y_u, \\ p(1) = 0, \end{cases}$$

tenemos que la derivada direccional ahora se escribe como

$$\varphi'(u;v) = \int_0^1 p(t)v(t)dt. \tag{38}$$

- Finalmente consideremos $\varphi : L^\infty(0,1) \to I\!R$ tal que $\varphi(u) = \int_0^1 y^2(t) + u^2(t)dt$ donde y es solución de

$$\begin{cases} y' = y + u \\ y(0) = 1 \end{cases}$$

  Usando los resultados anteriores podemos calcular la derivada Gâteaux (Fréchet) obteniedo

$$\varphi'(u;v) = \int_0^1 (2u(t) + p(t))v(t)dt. \tag{39}$$

Demostración. Consideremos la función $\varphi(t) := f(\bar{x} + tv)$ de variable real y a valores reales. Como $\bar{x}$ es un minimizador, φ debe tener un mínimo en 0 para cualquier dirección $v \in X$. Usando la condición necesaria para mínimo en $I\!R$ tenemos entonces que

$$\varphi'(0) = Df(\bar{x})v = 0, \quad \forall v \in X. \tag{40}$$

□

A modo de ejemplo consideremos el problema de control óptimo

$$\min_u \varphi(u) := \tfrac{1}{2} \int_0^1 y^2(t) + u^2(t)dt$$

donde y es solución de

$$\begin{cases} y' = y + u, \\ y(0) = 1. \end{cases}$$

Si aplicamos el teorema tenemos que una condición para que u^* sea mínimo es que la derivada se anule, con lo cual

$$\varphi'(u^*; v) = \int_0^1 (u^*(t) + p(t))v(t)dt = 0, \quad \forall v \in L^\infty. \tag{41}$$

De donde obtenemos que $u^*(t) = -p(t)$ c.t.p. ¿Pero, quién es p? Para obtenerlo deberemos resolver el sistema

$$\begin{cases} y' = y + u, \\ y(0) = 1, \\ p' = -p - y, \\ p(1) = 0. \end{cases}$$

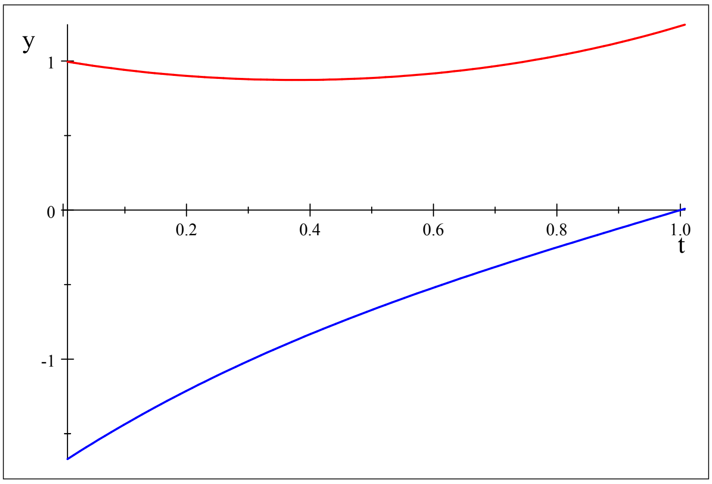
Reemplazando el valor obtenido de u tenemos

$$\begin{cases} y' = y - p, \\ p' = -p - y, \\ y(0) = 1, \\ p(1) = 0. \end{cases} \tag{42}$$

Como el sistema es lineal y la matriz asociada tiene dos autovalores distintos tenemos que la solución general está dada por

$$\begin{array}{l} p(t) = c_1 e^{\sqrt{2}t} + c_2 e^{-\sqrt{2}t}, \\ y(t) = -c_1(\sqrt{2}+1)e^{\sqrt{2}t} + c_2(\sqrt{2}-1)e^{-\sqrt{2}t}, \end{array} \tag{43}$$

y usando las condiciones de borde se puede obtener el resultado que tiene la siguiente representación gráfica, donde y está en rojo y es positiva y u es azul y negativa.



Notemos que también hemos resuelto el problema de cálculo de variaciones

$$\min_y \int_0^1 y^2 + (y' + y)^2 dt \tag{44}$$

sobre las trayectorias y que verifican $y(0) = 1$.

Consideremos ahora un problema variacional de la forma

$$\min_{y \in K} \int_0^1 \ell(y, y') dt \tag{45}$$

donde K es un conjunto de funciones C^1 que verifican una condición inicial. Este problema podemos escribirlo de la forma

$$\min_{u \in C} \int_0^1 \ell(y, u) dt \tag{46}$$

donde y verifica la condición inicial de K y la ODE $y' = u$. Por lo visto anteriormente tenemos que la derivada direccional se escribe

$$\varphi'(u; v) = \int_0^1 (p + \ell_u(y, u)) v dt. \tag{47}$$

donde p verifica la ecuación adjunta $p' = -\ell_y(y, u)$, por lo tanto condición de optimalidad de primer orden es

$$\begin{cases} p + \ell_u(y, u) = 0, \\ p' = -\ell_y(y, u), \\ p(1) = 0. \end{cases} \tag{48}$$

Volviendo a la notación original tenemos que resolver

$$\frac{d}{dt} \ell_{y'}(y, y') = \ell_y(y, y') \tag{49}$$

que es la famosa ecuación de Euler.

Finalmente consideremos el problema de cálculo variacional para un dominio $\Omega \subset I\!R^2$ de frontera Γ.

$$\begin{array}{rl} \text{mín}_u & \Phi(u) = \int_\Omega |\nabla u|^2 dx \\ \text{t.q.} & u|_\Gamma = u_0. \end{array} \tag{50}$$

Calculando la derivada direccional tenemos que

$$\Phi'(u;v) = \int_\Omega \nabla u \nabla v dx = -\int_\Omega \Delta u v dx + \int_\Gamma v \frac{\partial u}{\partial n} d\gamma, \tag{51}$$

donde para la última igualdad utilizamos la fórmula de Green. Ahora, es fácil ver que las direcciones v deben valer cero en la frontera con lo cual la condición de primer orden queda

$$\begin{cases} \Delta u = 0, \\ u|_\Gamma = u_0. \end{cases} \tag{52}$$

Referencias

[1] H. Cartan. *Cours de calcul différentiel*. Hermann, Paris, 1977.

[2] B. Dacorogna. *Diect Methods in the Clculus of VaritionsPertubations analysis of optimizations problems*. Springer Verlag, Berlin, 1989.

[3] A. Shapiro J.F. Bonnans. *Pertubations analysis of optimizations problems*. Springer Verlag, New York, 2000.