DEPARTAMENTO DE INGENIERÍA MATEMÁTICA
FACULTAD DE CIENCAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Análisis Convexo y Dualidad

$$f(x) + f^*(x^*) = \langle x^*, x \rangle \Leftrightarrow x^* \in \partial f(x)$$

Apuntes para el curso del programa de
Doctorado en Ciencias de la Ingeniería,
mención Modelación Matemática

Felipe Álvarez

con la colaboración de
Juan Escobar y Juan Peypouquet

Septiembre 2005

Este documento fue financiado a través de los recursos asignados por el DIM para la realización de actividades docentes que le son propias.

Prefacio

El objetivo de estos apuntes es proporcionar los fundamentos del Análisis Convexo y de la Teoría de la Dualidad en espacios de dimensión finita e infinita, junto con abordar algunas aplicaciones en Optimización y Cálculo de Variaciones.

Estos apuntes se basan esencialmente en las notas escritas para el curso "Análisis Convexo y Dualidad" (MA674) del programa de Doctorado en Ciencias de la Ingeniería, mención Modelación Matemática, de la Universidad de Chile.

Mis agradecimientos a Juan Escobar y Juan Peypouquet, quienes participaron activamente en la confección de este apunte y sin cuya valiosa colaboración estas notas no tendrían su forma actual. Juan Escobar transcribió en LaTeX la primera versión del manuscrito, sugiriendo varias ideas y contribuyendo con material adicional, particularmente en la sección que se refiere al principio variacional de Ekeland y al capítulo sobre el esquema primal/dual de penalización en programación convexa. Posteriormente, Juan Peypouquet realizó una revisión muy profunda y exhaustiva del apunte, corrigió errores, mejoró sustancialmente la presentación e incorporó material referente al análisis de recesión. Vaya nuevamente mi reconocimiento al excelente trabajo de ambos.

Todo posible error que el lector pueda encontrar en este apunte es de mi exclusiva responsabilidad. Los comentarios y observaciones son bienvenidos en la siguiente dirección de email: falvarez@dim.uchile.cl

Finalmente, quisiera agradecer el financiamiento proporcionado por el Departamento de Ingeniería Matemática de la Universidad de Chile.

Felipe Álvarez
Santiago, 7 de septiembre de 2005

Índice general

Capítulo 1

Introducción al Análisis Variacional

1.1. Funciones a valores en $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$.

En el análisis de problemas de minimización y maximización es conveniente considerar funciones que toman valores en la *recta real extendida* $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty]$ en lugar de sólo $\mathbf{R} =]-\infty, +\infty[$. Por ejemplo, en un espacio topológico X, consideremos un problema de minimización del tipo

$$\inf\{f_0(x) \mid x \in C\} \tag{1.1}$$

donde $f_0 \colon X \to \mathbf{R}$ es la *función objetivo* a minimizar y $C \subseteq X$ es un *conjunto de restricciones*. Con la topología apropiada, $\overline{\mathbf{R}}$ es un intervalo compacto. En particular, todo subconjunto $A \subseteq \overline{\mathbf{R}}$ admite un ínfimo $\inf A$ (y un supremo $\sup A$). En particular, tomando $A = \{f_0(x) \mid x \in C\}$ tenemos que $\inf\{f_0(x) \mid x \in C\}$ está bien definido en $\overline{\mathbf{R}}$, y denotamos

$$\inf_C f_0 := \inf\{f_0(x) \mid x \in C\}.$$

Para un conjunto A más general, si $\inf A \in A$ escribimos $\min A$ en lugar de $\inf A$ para enfatizar que el ínfimo se alcanza en el conjunto.

En este contexto resulta muy útil introducir la *función indicatriz* del conjunto C, $\delta_C \colon X \to \overline{\mathbf{R}}$, definida por

$$\delta_C(x) = \begin{cases} 0 & x \in C, \\ +\infty & x \notin C. \end{cases}$$

Con la convención

$$\alpha + (+\infty) = (+\infty) + \alpha = +\infty,\ \forall \alpha \in \mathbf{R},$$

es posible definir la función $f_0 + \delta_C$ mediante $(f_0 + \delta_C)(x) := f_0(x) + \delta_C(x)$, y es directo verificar que

$$\inf_C f_0 = \inf_X (f_0 + \delta_C)$$

y más aun, si $C \neq \emptyset$ entonces

$$\inf_C f_0 \in \{f_0(x) \mid x \in C\} \Leftrightarrow \inf_X (f_0 + \delta_C) \in \{f_0(x) + \delta_C(x) \mid x \in X\}.$$

De esta forma, el problema de minimización (1.1) puede formularse de manera equivalente como

$$\inf\{f(x) \mid x \in X\},$$

donde $f \colon X \to \mathbf{R} \cup \{+\infty\}$ está definida por $f = f_0 + \delta_C$. Esto permite considerar las restricciones de manera implícita en la definición de f y dar así un tratamiento unificado a este tipo de problemas.

Similarmente, la consideración de problemas de maximización con restricciones conduce de manera natural a funciones a valores en $\mathbf{R} \cup \{-\infty\}$; los problemas de tipo *minimax*, a funciones a valores en $\overline{\mathbf{R}}$.

Dados un escalar $\lambda \in \mathbf{R}$ y dos funciones $f, g \colon X \to \overline{\mathbf{R}}$, queremos dar un sentido a la expresión $f + \lambda g$, para lo cual necesitamos una aritmética en $\overline{\mathbf{R}}$ que extienda la usual de $\mathbf{R}$. Consideraremos las siguientes convenciones:

1. $(+\infty) + \alpha = \alpha + (+\infty) = +\infty$, $\forall \alpha \in \mathbf{R} \cup \{+\infty\}$.
2. $(-\infty) + \alpha = \alpha + (-\infty) = -\infty$, $\forall \alpha \in \mathbf{R} \cup \{-\infty\}$.
3. $\alpha \cdot (+\infty) = +\infty$, si $\alpha > 0$, $\alpha \cdot (+\infty) = -\infty$, si $\alpha < 0$ (análogo para $-\infty$).

Menos obvias son las expresiones del tipo $0 \cdot (\pm\infty)$ y $(\mp\infty) + (\pm\infty)$; utilizaremos, salvo que se diga explícitamente otra cosa, las siguientes convenciones:

4. $0 \cdot (+\infty) = 0 = 0 \cdot (-\infty) = 0$ y la simétrica para que sea conmutativa.
5. $+\infty + (-\infty) = (-\infty) + (+\infty) = +\infty$, que se llama *inf-adición* pues está orientada a la minimización: violar las restricciones tiene prioridad por sobre un eventual valor $-\infty$ de la función objetivo.

Observación 1.1.1. Estas extensiones para la suma y producto no son continuas. Límites que involucran expresiones de la forma $0 \cdot (\pm\infty)$ o $(\mp\infty) + (\pm\infty)$ pueden estar indeterminados en el sentido que si $\alpha_n \to \alpha$ y $\beta_n \to \beta$ entonces no necesariamente se tiene que $\alpha_n\beta_n \to \alpha\beta$ cuando $\alpha\beta = 0 \cdot (\pm\infty)$, ni tampoco que $\alpha_n + \beta_n \to \alpha + \beta$ cuando $\alpha + \beta = (\pm\infty) + (\mp\infty)$.

Las convenciones anteriores permiten definir

$$(f + \lambda g)(x) := f(x) + \lambda g(x) \text{ y } (fg)(x) := f(x)g(x), \ \forall x \in X,$$

para todo $\lambda \in \mathbf{R}$ y $f, g : X \to \overline{\mathbf{R}}$.

Por otra parte, dado que la minimización es una operación unilateral, es natural que se requieran conceptos unilaterales para su análisis:

Definición 1.1.1. Dada $f \colon X \to \overline{\mathbf{R}}$, su *epigrafo* es el subconjunto de $X \times \mathbf{R}$ dado por

$$\operatorname{epi} f := \{(x, \lambda) \in X \times \mathbf{R} \mid f(x) \le \lambda\},$$

mientras que el *conjunto de nivel inferior (o subnivel) de parámetro* $\gamma \in \mathbf{R}$ es el subconjunto de X dado por

$$\Gamma_\gamma(f) := \{x \in X \mid f(x) \le \gamma\}.$$

El *dominio efectivo* de f es

$$\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\}.$$

Además, definimos

$$\arg\min f := \begin{cases} \{x \in X \mid f(x) = \inf_X f\} & \text{si } \inf_X f < +\infty, \\ \emptyset & \text{si } \inf_X f = +\infty. \end{cases}$$

Notemos que

$$\inf_X f = \inf_{\operatorname{dom} f} f,$$

con la convención

$$\inf_{\emptyset} f = \inf \emptyset = +\infty.$$

La condición $\arg\min f = \emptyset$ si $f \equiv +\infty$ se interpreta diciendo que en ese caso la minimización no tiene soluciones optimales pues ni siquiera hay *puntos factibles* que satisfagan las restricciones implícitas de f (un punto $x \in X$ se dice factible si $f(x) < +\infty$, i.e., si $x \in \operatorname{dom} f$).

Notemos también que se tienen las siguientes propiedades:

(a) $\arg\min f = \bigcap_{\gamma > \inf_X f} \Gamma_\gamma(f)$ (con la convención $\bigcap_{\emptyset} = \emptyset$ para $f \equiv +\infty$).

(b) $\Gamma_\gamma(f) \times \{\gamma\} = \operatorname{epi} f \cap (X \times \{\gamma\})$.

Ejercicio 1.1.1. Demuestre estas propiedades.

Finalmente, otra ventaja de considerar funciones a valores en $\overline{\mathbf{R}}$ es que es una clase estable bajo operaciones del tipo supremo e ínfimo sobre familias arbitrarias; más precisamente, tenemos la siguiente:

Proposición 1.1.1. *Sea $(f_i)_{i \in I}$ una familia arbitraria no vacía ($I \neq \emptyset$) de funciones $f_i \colon X \to \overline{\mathbf{R}}$. Entonces las funciones $\sup_{i \in I} f_i$ e $\inf_{i \in I} f_i$ definidas respectivamente por*

$$(\sup_{i \in I} f_i)(x) := \sup\{f_i(x) \mid i \in I\}$$

e

$$(\inf_{i \in I} f_i)(x) := \inf\{f_i(x) \mid i \in I\}$$

están bien definidas como funciones a valores en $\overline{\mathbf{R}}$. Más aun,

$$\operatorname{epi}(\sup_{i \in I} f_i) = \bigcap_{i \in I} \operatorname{epi} f_i$$

y

$$\operatorname{epi}(\inf_{i \in I} f_i) \supseteq \bigcup_{i \in I} \operatorname{epi} f_i,$$

con igualdad en la última inclusión si $|I| < +\infty$.

Demostración:

$$
\begin{aligned}
\operatorname{epi}(\sup f_i) &= \{(x,\alpha) \in X \times \mathbf{R} \mid \sup f_i(x) \leq \alpha\} \\
&= \{(x,\alpha) \in X \times \mathbf{R} \mid f_i(x) \leq \alpha,\ \forall i \in I\} \\
&= \bigcap_{i\in I} \operatorname{epi}(f).
\end{aligned}
$$

Del mismo modo,

$$
\begin{aligned}
\bigcup_{i\in I} \operatorname{epi}(f_i) &= \{(x,\alpha) \in X \times \mathbf{R} \mid \exists i \in I,\ f_i(x) \leq \alpha\} \\
&\subseteq \{(x,\alpha) \in X \times \mathbf{R} \mid \inf f_i(x) \leq \alpha\} \\
&= \operatorname{epi}(\inf f_i),
\end{aligned}
$$

y se tiene que la inclusión es en realidad una igualdad si $|I| < +\infty$, pues en este caso el ínfimo siempre se alcanza para algún $i \in I$.

□

Nos interesaremos en una subclase de funciones para las cuales el problema de minimización asociado es no trivial, de acuerdo a la siguiente definición.

Definición 1.1.2. Una función $f\colon X \to \overline{\mathbf{R}}$ se dice *propia* si:

(i) $\forall x \in X,\ f(x) > -\infty$,

(ii) $\operatorname{dom} f \neq \phi$. Es decir, $\exists x_0 \in X$ tal que $f(x_0) < +\infty$.

Observación 1.1.2. Notemos que para una función propia eventualmente su ínfimo es $-\infty$. Si f tiene ínfimo mayor estricto que $-\infty$, es decir $\inf_X f > -\infty$, diremos que f está *acotada inferiormente.*

1.2. Semicontinuidad inferior y minimización

Hasta ahora no hemos supuesto ningún tipo de estructura sobre el conjunto subyacente X. En lo que sigue, supondremos que (X, τ) es un espacio topológico.

1.2.1. Semicontinuidad Inferior

Dado $x \in X$, denotemos por $\mathcal{N}_x(\tau)$ la familia de todas las vecindades de x para τ.

Definición 1.2.1. Una función $f\colon X \to \mathbf{R} \cup \{+\infty\}$ se dice *τ-semicontinua inferior (τ-*s.c.i.*)* en x si

$$\forall \lambda < f(x), \exists N_\lambda \in \mathcal{N}_x(\tau) : \forall y \in N_\lambda, f(y) > \lambda.$$

Si lo anterior es válido para todo $x \in X$, decimos simplemente que f es τ-s.c.i. sobre X.

Proposición 1.2.1. *Dada $f\colon X \to \mathbf{R} \cup \{+\infty\}$, las siguientes afirmaciones son equivalentes:*

(i) f es τ-s.c.i. sobre X.

(ii) $\mathrm{epi}(f)$ *es cerrado en* $X \times \mathbf{R}$ *dotado de la topología* $\tau \times \tau_{\mathbf{R}}$*, donde* $\tau_{\mathbf{R}}$ *es la topología usual de* $\mathbf{R}$.

(iii) $\forall \gamma \in \mathbf{R}$, $\Gamma_\gamma(f)$ *es cerrado en* (X, τ).

(iv) $\forall \gamma \in \mathbf{R}$, $\{x \in X \mid f(x) > \gamma\} \in \tau$.

(v) $\forall x \in X,\ f(x) \leq \liminf\limits_{y \to x} f(y) := \sup\limits_{N \in \mathcal{N}_X(\tau)} \inf\limits_{y \in N} f(y)$

Demostración:

- $(i) \Rightarrow (ii)$ Tomemos $(x, \lambda) \notin \mathrm{epi}(f)$, lo que equivale a $\lambda < f(x)$. Sea $\gamma \in \mathbf{R}$ tal que $\lambda < \gamma < f(x)$. De (i), existe $N_\gamma \in \mathcal{N}_x(\tau)$ tal que $\forall y \in N_\gamma,\ f(y) > \gamma$, de modo que $(y, \gamma) \notin \mathrm{epi}(f)$. Se sigue que $(N_\gamma \times]-\infty, \gamma[) \cap \mathrm{epi}(f) = \emptyset$. Como $N_\gamma \times]-\infty, \gamma[\in \mathcal{N}_{(x,\lambda)}(\tau \times \tau_{\mathbf{R}})$, concluimos que $\mathrm{epi}(f)^C$ es abierto.
- $(ii) \Rightarrow (iii)$ Como $\Gamma_\gamma(f) \times \{\gamma\} = \mathrm{epi}(f) \cap (X \times \{\gamma\})$, deducimos que $\Gamma_\gamma(f) \times \{\gamma\}$ es cerrado en $X \times \mathbf{R}$, y de aquí que $\Gamma_\gamma(f)$ es cerrado en X.
- $(iii) \Rightarrow (iv)$ Trivial.
- $(iv) \Rightarrow (v)$ Sea $x \in X$. Dado $\gamma < f(x)$ tenemos que $N = \{y \in X \mid f(y) > \gamma\} \in \mathcal{N}_x(\tau)$ por ser un abierto que contiene a x. De este modo $\gamma \leq \inf\limits_{y \in N} f(y)$, y en particular $\gamma \leq \sup\limits_{N \in \mathcal{N}_x(\tau)} \inf\limits_{y \in N} f(y)$. Como lo anterior es válido para todo $\gamma < f(x)$, se sigue (v).
- $(v) \Rightarrow (i)$ Sea $\lambda < f(x)$. Por (v), $\lambda < \sup\limits_{N \in \mathcal{N}_X(\tau)} \inf\limits_{y \in N} f(y)$ y en consecuencia , existe $N \in \mathcal{N}_x(\tau)$: $\lambda < \inf\limits_{y \in N} f(y)$.

□

Ejercicio 1.2.1. Pruebe que cuando (X, τ) es metrizable, f es τ s.c.i. $\Leftrightarrow \forall x \in X\ x_n \to x,\ f(x) \leq \liminf\limits_{n \to +\infty} f(x_n) := \sup\limits_{n \in \mathbf{N}} \inf\limits_{k \geq n} f(x_k)$.

La clase de funciones s.c.i. satisface varias propiedades de estabilidad, como lo establece el siguiente resultado.

Proposición 1.2.2. *Sea* $\{f_i\}_{i \in I}$ *una familia arbitraria no vacía de funciones* τ*-s.c.i.,* $f_i \colon X \to \mathbf{R} \cup \{+\infty\}$. *Entonces* $\sup\limits_{i \in I} f_i$ *es* τ*-s.c.i. Si además* I *es finito,* $\min\limits_{i \in I} f_i$ *y* $\sum\limits_{i \in I} f_i$ *son ambas* τ*-s.c.i.*

Demostración: Propuesto. □

1.2.2. Inf-compacidad y existencia de minimizadores

Junto con la s.c.i., el segundo ingrediente básico en los problemas de minimización es la propiedad de *inf-compacidad* en el sentido de la siguiente definición.

Definición 1.2.2. Una función $f \colon X \to \mathbf{R} \cup \{+\infty\}$ se dice *τ-inf-compacta* si para todo $\gamma \in \mathbf{R}$, el conjunto $\Gamma_\gamma(f) = \{x \in X \mid f(x) \leq \lambda\}$ es relativamente compacto en X para la topología τ.

Observación 1.2.1. Para una función f que es τ-s.c.i., lo anterior equivale a requerir que $\Gamma_\gamma(f)$ sea compacto.

Definición 1.2.3. Sea $(V, \|\cdot\|)$ un espacio vectorial normado. Una función $f\colon V \to \mathbf{R} \cup \{+\infty\}$ se dice *coerciva* si $\lim\limits_{\|x\|\to\infty} f(x) = +\infty$.

Observación 1.2.2. Sea $(V, \|\cdot\|)$ un espacio vectorial normado y $f\colon V \to \mathbf{R} \cup \{+\infty\}$ una función. Entonces son equivalentes (i) f es coerciva y (ii) $\forall \gamma \in \mathbf{R}$, $\Gamma_\gamma(f)$ es acotada. En particular, si $\dim V < +\infty$ entonces f es inf-compacta ssi f es coerciva. Recordemos el Teorema de Riesz: todo subconjunto acotado en un espacio vectorial normado $(V, \|\cdot\|)$ es relativamente compacto ssi $\dim V < +\infty$. En el caso de la dimensión infinita, las topologías que están asociadas a la coercividad son las débiles. Esto lo discutiremos más adelante.

Teorema 1.2.1 (Weierstrass-Hilbert-Tonelli)**.** *Sea $f\colon X \to \mathbf{R} \cup \{+\infty\}$ una función τ-s.c.i. y τ-inf-compacta. Entonces, $\inf_X f > -\infty$ y existe un punto $x^* \in X$ que minimiza f sobre X, i.e., $f(x^*) = \min\limits_{x\in X} f(x)$.*

Demostración: Veremos dos demostraciones distintas de este teorema.

La primera demostración es conocida como *Método Directo* y fue iniciada por Hilbert y luego desarrollada por Tonelli. Para simplificar, supongamos que τ es metrizable. Consideramos primero una sucesión $(x_n)_{n\in\mathbf{N}}$ minimizante para f, i.e., $\lim\limits_{n\to+\infty} f(x_n) = \inf_X f$. Tal sucesión existe. En efecto, si $\inf_X f > -\infty$ entonces consideramos para $n \geq 1, x_n \in X$ tal que $\inf_X f \leq f(x_n) \leq \inf_X f + \frac{1}{n}$. Si $\inf_X f = -\infty$ entonces sea $x_n \in X$ tal que $f(x_n) \leq -n$ (*a posteriori* veremos que este caso no se tiene). Si $\inf_X f = +\infty$ entonces $f \equiv +\infty$ y basta tomar cualquier $x^* \in X$. En caso contrario, tenemos que $f(x_n) \leq \max\{\inf_X f + \frac{1}{n}, -n\} \leq \max\{\inf_X f + 1, -1\} =: \gamma_0 \in \mathbf{R}$. Notemos que $\gamma_0 > \inf_X f$ y que $x_n \in \Gamma_{\gamma_0}(f)$ para todo $n \geq 1$. Como $\Gamma_{\gamma_0}(f)$ es compacto, por el Teorema de Bolzano-Weiertrass podemos extraer una subsucesión (x_{n_k}) que converge (en la topología τ) a algún punto $x^* \in X$. En particular, $\lim\limits_{k\to\infty} f(x_{n_k}) = \inf_X f$ y, de la τ-s.c.i. de f, $f(x^*) \leq \lim\limits_{k\to\infty} f(x_k) = \inf_X f$. Luego, $\inf_X f > -\infty$ y $f(x^*) = \inf_X f$.

- Queremos demostrar que $\arg\min f \neq \emptyset$. Sabemos que

$$\arg\min f = \bigcap_{\gamma > \inf_X f} \Gamma_\gamma(f) = \bigcap_{\gamma_0 > \gamma > \inf_X f} \Gamma_\gamma(f)$$

para cualquier $\gamma_0 \in \mathbf{R}$ tal que $\gamma_0 > \inf_X f$ (notemos que sin pérdida de generalidad suponemos $\inf_X f < +\infty$), esto último pues $\Gamma_\alpha(f) \subseteq \Gamma_\beta(f)$, si $\alpha \leq \beta$. Como f es τ-s.c.i., $\Gamma_\gamma(f)$ es cerrado y además es compacto por la inf-compacidad de f. En particular, $\{\Gamma_\gamma(f)\}_{\inf_X f < \gamma < \gamma_0}$ es una familia de subconjuntos cerrados y no vací os (¿por qué?) del compacto $\Gamma_{\gamma_0}(f)$. Más aun, esta familia satisface la propieded de intersecciones finitas. En efecto, dados $\gamma_1, ..., \gamma_n$, con $\gamma = \min\{\gamma_1, ..., \gamma_n\} > \inf_X f$ se tiene que

$$\bigcap_{i=1}^{n} \Gamma_{\gamma_i}(f) = \Gamma_\gamma(f) \neq \emptyset.$$

Por compacidad,

$$\bigcap_{\inf_X f < \gamma < \gamma_0} \Gamma_\gamma(f) \neq \emptyset,$$

lo que concluye la demostración.

□

Observación 1.2.3.

- El Teorema 1.2.1, de Weierstrass-Hilbert-Tonelli se conoce también como Teorema de Minimización de Weierstrass y sigue siendo válido si en lugar de la τ-inf-compacidad suponemos que $\exists \gamma_0 > \inf_X f \ : \ \Gamma_{\gamma_0}(f)$ es compacto. Por ejemplo, definamos $f\colon \mathbf{R} \to \mathbf{R}$ por $f(x) = \frac{x^2}{1+x^2}$. Es fácil ver que $\Gamma_\gamma(f)$ es compacto ssi $\gamma < 1$ y $\Gamma_1(f) = \mathbf{R}$ de modo que no es inf-compacta pero sí tiene un minimizador (x* =0).
- Notemos que en la demostración vía el Método Directo hemos supuesto que τ es metrizable para extraer una subsucesión convergente.

Ejercicio 1.2.2. Sea $f\colon X \to \mathbf{R} \cup \{+\infty\}$ una función τ-s.c.i. y $K \subseteq X$ un compacto para τ. Muestre que existe $x^* \in K$ tal que $f(x^*) = \min_K f$. Ind.: Considere $g := f + \delta_K$.

1.2.3. Funciones de recesión e inf-compacidad.

En esta sección presentaremos brevemente las funciones de recesión, muy útiles en análisis convexo. Terminaremos con un resultado que permite caracterizar la inf-compacidad en espacios de dimensión finita y una aplicación a problemas de optimización. En lo que sigue, X es un e.v.n.

Definición 1.2.4. Sea $C \subset X$ un conjunto no vacío. El *cono de recesión*, denotado por C_∞, es el conjunto

$$C_\infty = \left\{ d \in X \;\middle|\; \exists t_k \to \infty,\ \exists x_k \in C \text{ tales que } \lim \frac{x_k}{t_k} = d \right\}.$$

El cono de recesión contiene las direcciones hacia las cuales se dirigen las sucesiones en C que tienden a $+\infty$ en norma. Notemos que el conjunto C_∞ tiene las siguientes propiedades elementales, que el lector debe verificar:

1. C_∞ es un cono cerrado.
2. $\left(\overline{C}\right)_\infty = C_\infty$.
3. Cuando el conjunto C es convexo, también lo es C_∞. Además

$$\begin{aligned} C_\infty &= \{d \in X \mid d + \overline{C} \subset \overline{C}\} \\ &= \{d \in X \mid x + td \in \overline{C} \text{ para todo } t > 0\} \end{aligned}$$

cualquiera que sea el $x \in C$.

Ejemplo 1.2.1. Conos de recesión de algunos conjuntos relevantes:

1. Si C es un cono, entonces $C_\infty = \overline{C}$.
2. Si C es acotado, entonces $C_\infty = \{0\}$. Si la dimensión de X es finita, el recíproco también es cierto.
3. Si C es un hiperplano cerrado, entonces C_∞ es el subespacio vectorial paralelo a C.
4. Suponga que C es un poliedro convexo definido mediante $C = \{ x \in \mathbf{R}^n \mid Ax \leq b \}$, donde A es una matriz de tamaño $m \times n$ y $b \in \mathbf{R}^m$. Entonces $C_\infty = \{ d \in \mathbf{R}^n \mid Ad \leq 0 \}$.

Ejercicio 1.2.3. Consideremos subconjuntos $(C_i)_{i\in I}$ de X. Pruebe que

1. $\left(\bigcap_{i\in I} C_i\right)_\infty \subseteq \bigcap_{i\in I}(C_i)_\infty$ si la intersección es no-vacía. La inclusión es una igualdad si los conjuntos son convexos.

2. $\left(\bigcup_{i\in I} C_i\right)_\infty \supseteq \bigcup_{i\in I}(C_i)_\infty$, con igualdad si I es finito.

Definición 1.2.5. Sea $f : X \to \mathbf{R} \cup \{+\infty\}$ una función propia. La *función de recesión*, denotada por f_∞, se define por

$$f_\infty(d) = \inf\left\{ \liminf_{k\to\infty} \frac{f(t_k d_k)}{t_k} \;\middle|\; t_k \to +\infty, \quad d_k \to d \right\}$$

Ejemplo 1.2.2. Si C es un conjunto no vacío, entonces $(\delta_C)_\infty = \delta_{C_\infty}$

Ejercicio 1.2.4. Sean $f, g : X \to \mathbf{R} \cup \{+\infty\}$ dos funciones propias con $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$. Sea $h = f + g$.

1. Si f y g son s.c.i., entonces h también lo es.

2. Sea $d \in X$ tal que $f_\infty(d)$ y $g_\infty(d)$ no son, respectivamente, $+\infty$ y $-\infty$. Pruebe que

$$h_\infty(d) \geq f_\infty(d) + g_\infty(d).$$

Proposición 1.2.3. *Si f es propia, entonces* $\operatorname{epi}(f_\infty) = (\operatorname{epi} f)_\infty$.

Demostración: Probaremos primero que $(\operatorname{epi} f)_\infty \subseteq \operatorname{epi}(f_\infty)$. Para ello, tomemos $(d,\mu) \in (epi f)_\infty$. De acuerdo con la definición de cono asintótico, existen $t_k \to \infty$ y $(d_k, \mu_k) \in \operatorname{epi} f$ tales que $t_k^{-1}(d_k, \mu_k) \to (d, \mu)$. Como $f(d_k) \leq \mu_k$, tenemos que $t_k^{-1} f(t_k^{-1} d_k \cdot t_k) \leq t_k^{-1}\mu_k$. Pasando al límite tenemos que $f_\infty(d) \leq \mu$, de donde $(d,\mu) \in \operatorname{epi}(f_\infty)$.

Recíprocamente, sea $(d,\mu) \in \operatorname{epi}(f_\infty)$. Por definición, existen sucesiones $t_k \to \infty$ y $d_k \to d$ tales que $f_\infty(d) = \lim_{k\to\infty} t_k^{-1} f(t_k d_k)$. Como $(d,\mu) \in \operatorname{epi}(f_\infty)$, de acuerdo con la definición de límite observamos que para cada $\varepsilon > 0$, tomando k suficientemente grande tenemos que $f(t_k d_k) \leq (\mu - \varepsilon) t_k$. Por lo tanto, el punto $z_k = t_k(d_k, \mu + \varepsilon) \in \operatorname{epi}(f_\infty)$. Como $t_k^{-1} z_k = (d_k, \mu + \varepsilon) \to (d, \mu + \varepsilon)$, concluimos que $(d, \mu + \varepsilon) \in (\operatorname{epi} f)_\infty$. Finalmente, dado que $(\operatorname{epi} f)_\infty$ es cerrado y ε es arbitrario, vemos que $(d,\mu) \in (\operatorname{epi} f)_\infty$. $\square$

Ejercicio 1.2.5. Sea $(f_i)_{i\in I}$ una colección de funciones propias de X en $\mathbf{R} \cup \{+\infty\}$. Pruebe que

$$\left(\sup_{i\in I} f_i\right)_\infty \geq \sup_{i\in I}\{(f_i)_\infty\} \qquad \text{y} \qquad \left(\inf_{i\in I} f_i\right)_\infty \leq \inf_{i\in I}\{(f_i)_\infty\}.$$

Demuestre también que si I es finito, la segunda desigualdad es una igualdad.

Lema 1.2.1. *Si $f : X \to \mathbf{R} \cup \{+\infty\}$ es propia, entonces para todo $\lambda > \inf f$,*

$$[\Gamma_\lambda(f)]_\infty \subseteq \{ d \in X \mid f_\infty(d) \leq 0 \}.$$

Demostración: Sea $d \in [\ \Gamma_\lambda(f)\ ]_\infty$. Entonces existen sucesiones $x_k \in \Gamma_\lambda(f)$ y $t_k \to \infty$ tales que $\lim_{k\to\infty} t_k^{-1}x_k = d$. Escribamos $d_k = t_k^{-1}x_k \to d$. Como $x_k \in \Gamma_\lambda(f)$, tenemos que $t_k^{-1}f(t_kd_k) = t_k^{-1}f(x_k) \leq t_k^{-1}\lambda \to 0$. Por lo tanto, $f_\infty(d) \leq 0$. □

Corolario 1.2.1. *Consideremos un conjunto $(f_i)_{i\in I}$ de funciones propias de X en $\mathbf{R}\cup\{+\infty\}$ y un subconjunto no-vacío S de X. Defina $C = \{\ s \in S \mid f_i(x) \leq 0\ \forall i \in I\ \}$. Demuestre que $C_\infty \subseteq \{\ d \in S_\infty \mid (f_i)_\infty(d) \leq 0\ \forall i \in I\ \}$.*

Demostración: Definimos $C_i = \{x \parallel f_i(x) \leq 0\}$, de manera que $C = S \cap \left(\bigcap_{i\in I} C_i\right)$. Del Ejercicio 1.2.3 tenemos que $C_\infty \subseteq S_\infty \cap \left(\bigcap_{i\in I} (C_i)_\infty\right)$. El resultado se obtiene al aplicar el lema anterior. □

Teorema 1.2.2. *Supongamos que $f : \mathbf{R}^n \to \mathbf{R}\cup\{+\infty\}$ es s.c.i. y propia. Si $f_\infty(d) > 0$ para todo $d \neq 0$, entonces f es inf-compacta.*

Demostración: Supongamos que $f_\infty(d) > 0$ para todo $d \neq 0$. Del Lema 1.2.1 tenemos que para cada $\lambda > \inf f$,

$$0 \in [\ \Gamma_\lambda(f)\ ]_\infty \subseteq \{\ d \in \mathbf{R}^n \mid f_\infty(d) \leq 0\ \} = \{0\}.$$

Como el cono de recesión se reduce al $\{0\}$, concluimos que $\Gamma_\lambda(f)$ es acotado. □

Corolario 1.2.2. *Supongamos que para cada $i = 0, 1, \ldots, m$, la función $f_i : \mathbf{R}^n \to \mathbf{R}\cup\{+\infty\}$ es s.c.i. y propia. Sea $C = \{\ x \in \mathbf{R}^n \mid f_i(x) \leq 0\ \forall i\ \}$. Suponga que $\operatorname{dom} f_0 \cap C \neq \emptyset$ y considere el problema de optimización con restricciones*

$$(\mathcal{P}) \qquad \inf\{\ f_0(x) \mid x \in C\ \}.$$

Escribiendo $f = f_0 + \delta_C$, el problema anterior es equivalente a

$$(\mathcal{P}) \qquad \inf\{\ f(x) \mid x \in \mathbf{R}^n\ \}.$$

Suponga que $(f_0)_\infty(d) > -\infty$ para todo $d \neq 0$. Si las funciones f_i, $i \geq 1$ no tienen ninguna dirección de recesión común; es decir, si

$$(f_i)_\infty(d) \leq 0\ \forall i \ \Rightarrow\ d = 0,$$

entonces el conjunto de soluciones de $(\mathcal{P})$ es no-vacío y compacto.

Demostración: De acuerdo con el Ejercicio 1.2.4 y el Ejemplo 1.2.2, $f_\infty(d) \geq (f_0)_\infty(d) + \delta_{C_\infty}(d)$, lo que implica que $f_\infty(d) \geq (f_0)_\infty(d)$ para todo $d \in C_\infty$. Aplicando el Ejercicio 1.2.1 y usando la hipótesis sobre las direcciones de recesión concluimos que $f_\infty(d) > 0$ para todo $d \neq 0$. Del teorema anterior deducimos inmediatamente que la función objetivo f es inf-compacta. El resultado se obtiene entonces al aplicar el Teorema 1.2.1 de Weierstrass-Hilbert-Tonelli. □

En el Problema 6 del próximo capítulo presentaremos otros resultados aplicados al caso convexo. En [AuT03] se puede encontrar una exposición más completa y detallada sobre conos y funciones de recesión.

1.2.4. Principio variacional de Ekeland en espacios métricos

En los teoremas de existencia de solución de un problema de minimización, la compacidad de los conjuntos de nivel juega un rol clave. Los siguientes resultados muestran que basta la completitud del espacio sobre el cual se minimiza para obtener la existencia de una solución aproximada en un sentido apropiado.

Teorema 1.2.3. *Sean (E, d) un espacio métrico completo y $f\colon E \to \mathbf{R} \cup \{+\infty\}$ una función propia, s.c.i. y acotada inferiormente. Consideremos $x_0 \in \operatorname{dom}(f)$ y $\varepsilon > 0$. Entonces existe $\bar{x} \in E$ tal que*

(i) $f(\bar{x}) + \varepsilon d(x_0, \bar{x}) \leq f(x_0)$,

(ii) $\forall x \neq \bar{x}$, $f(\bar{x}) < f(x) + \varepsilon d(x, \bar{x})$.

Demostración: Sin pérdida de generalidad suponemos que $\varepsilon = 1$ y que $f\colon E \to \mathbf{R}_+ \cup \{+\infty\}$. Consideremos $F : E \to 2^E$ definida por

$$F(x) = \{y \in E \mid f(y) + d(x, y) \leq f(x)\},$$

que toma valores cerrados y satisface las siguientes dos condiciones:

- $y \in F(y)$,
- Si $y \in F(x)$, entonces $F(y) \subseteq F(x)$. Nos referiremos a esta propiedad como *monotonía*.

Definamos ahora $v : \operatorname{dom}(f) \to \mathbf{R}$ por

$$v(y) := \inf_{z \in F(y)} f(z).$$

Dado $y \in F(x)$, se tiene que $d(x, y) \leq f(x) - v(x)$, lo cual implica que

$$\operatorname{diam}(F(x)) \leq 2(f(x) - v(x)).$$

Definamos la sucesión $(x_n)_{n\in\mathbf{N}}$ recursivamente a partir de x_0 por

$$x_{n+1} \in F(x_n), \quad f(x_{n+1}) \leq v(x_n) + 2^{-n}.$$

Luego, de la monotonía de F, $v(x_n) \leq v(x_{n+1})$. Por otro lado, como $v(y) \leq f(y)$, se tiene que

$$v(x_{n+1}) \leq f(x_{n+1}) \leq v(x_n) + 2^{-n} \leq v(x_{n+1}) + 2^{-n}$$

y por lo tanto

$$0 \leq f(x_{n+1}) - v(x_{n+1}) \leq 2^{-n}.$$

En consecuencia, $\operatorname{diam}(F(x_n)) \to 0$ y como $F(x_n)$ es una sucesión decreciente de cerrados en un espacio completo, se sigue que existe $\bar{x} \in E$ tal que

$$\bigcap_{n\in\mathbf{N}} F(x_n) = \{\bar{x}\}.$$

Como $\bar{x} \in F(x_0)$, se tiene (i). Más aun, $\bar{x} \in F(x_n)$ $\forall n$ de donde se sigue que $F(\bar{x}) \subset F(x_n)$ y en consecuencia

$$F(\bar{x}) = \{\bar{x}\}.$$

Dado $x \neq \bar{x}$, se tiene que $x \notin F(\bar{x})$ de donde concluimos (ii) □

Definición 1.2.6. Dada una función propia $f\colon X \to \overline{\mathbf{R}}$ y $\varepsilon > 0$, definimos el conjunto de sus *ε-mínimos* por

$$\varepsilon - \arg\min f = \begin{cases} \{x \in X \mid f(x) \leq \inf f + \varepsilon\} & \text{si } \inf f > -\infty, \\ \{x \in X \mid f(x) \leq -\frac{1}{\varepsilon}\} & \text{si no.} \end{cases}$$

Proposición 1.2.4 (Principio variacional de Ekeland). *Bajo las hipótesis del Teorema 1.2.3, consideremos $\varepsilon, \lambda > 0$ y sea $x_0 \in \varepsilon\lambda - \arg\min f$. Entonces existe $\bar{x} \in E$ tal que*

(i) $f(\bar{x}) \leq f(x_0)$,

(ii) $d(x_0, \bar{x}) \leq \lambda$,

(iii) $\forall x \in E, \quad f(\bar{x}) \leq f(x) + \varepsilon d(x, \bar{x})$.

Demostración: Basta aplicar el Teorema 1.2.3. □

1.3. Espacios vectoriales topológicos

Recordemos que un espacio vectorial real V dotado de una topología τ se dice *espacio vectorial topológico* (e.v.t.) si las operaciones suma

$$\begin{array}{ccc} V \times V & \to & V \\ (v, w) & \mapsto & v + w \end{array}$$

y ponderación por escalar

$$\begin{array}{ccc} \mathbf{R} \times V & \to & V \\ (\lambda, v) & \mapsto & \lambda v \end{array}$$

son continuas para las topologías producto $\tau \times \tau$ sobre $V \times V$ y $\tau_{\mathbf{R}} \times V$ sobre $\mathbf{R} \times V$ respectivamente. Se tiene que las vecindades de cualquier punto se obtienen a partir de las vecindades del origen por traslación. Se dice además que el e.v.t. es *localmente convexo* (l.c.) si el origen admite una base de vecindades convexas.

El ejemplo más simple de e.v.t.l.c. es el de un espacio vectorial normado $(V, \|\cdot\|)$, pues basta tomar la base de vecindades constituida por las bolas con centro en el origen.

Dado un e.v.t. (V, τ) denotaremos por $(V, \tau)^*$, o simplemente por V^*, el espacio vectorial de todas los funcionales lineales (funciones lineales de V a valores en $\mathbf{R}$) que son τ-continuas sobre V. El espacio V^* se conoce como espacio *dual*[1] de (V, τ). Si $v^* \in V^*$ entonces para todo $v \in V$ escribiremos

$$\langle v, v^* \rangle := v^*(v),$$

lo que define una forma bilineal $\langle \cdot, \cdot \rangle : V \times V^* \to \mathbf{R}$ conocida como *producto de dualidad* entre V y V^*. Esta última notación apunta a enfatizar los roles en cierto sentido simétricos jugados por V y V^*. En efecto, dado $v \in V$, el *funcional de evaluación* $\langle v, \cdot \rangle : V^* \to \mathbf{R}$ es una forma lineal sobre V^*.

[1] Al espacio V^* también se le llama *dual topológico* para diferenciarlo del dual algebraico; este último contiene a todas las formas lineales sobre V independiente de cualquier noción topológica.

1.3.1. Separación de convexos: el Teorema de Hahn-Banach

Definición 1.3.1. Un *hiperplano* (o *hiperplano afín*) es un subconjunto $H \subseteq V$ de la forma $\{v \in V \mid f(v) = \alpha\}$ donde $f\colon V \to \mathbf{R}$ es un funcional lineal no nulo y $\alpha \in \mathbf{R}$.

Notación: Escribiremos $[\ell = \alpha] = \{v \in V \mid f(v) = \alpha\}$, $[\ell \leq \alpha] = \{v \in V \mid f(v) \leq \alpha\}$ y se extiende la notación de manera obvia para $<$, $\geq$ y $>$.

Observación 1.3.1. Es posible probar que $[\ell = \alpha]$ es τ-cerrado si, y sólo si, $\ell \in (V,\tau)^*$, en cuyo caso los *semiespacios* $[\ell \leq \alpha]$ y $[\ell \geq \alpha]$ (respectivamente $[\ell < \alpha]$ y $[\ell > \alpha]$) serán cerrados (resp. abiertos) en (V,τ).

Definición 1.3.2. Un conjunto $C \subseteq V$ se dice *convexo* si para todos $x, y \in C$ y para todo $\lambda \in [0,1]$ se tiene que

$$\lambda x + (1-\lambda)y \in C.$$

Una función $f\colon V \to \mathbf{R} \cup \{+\infty\}$ se dice *convexa* si para todos $x, y \in V$ y para todo $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Proposición 1.3.1. *$f\colon V \to \mathbf{R} \cup \{+\infty\}$ es convexa ssi* $\operatorname{epi} f$ *es convexo en $V \times \mathbf{R}$.*

Demostración: Se deja como ejercicio. □

Teorema 1.3.1 (Hahn-Banach)**.** *Sean (V,τ) un e.v.t. y $A, B \subseteq V$ dos subconjuntos <u>convexos</u>, no vacíos y disjuntos (i.e. $A \cap B = \emptyset$).*

(i) Si A es abierto entonces existe un hiperplano cerrado que separa A y B, esto es, existen $\ell \in (V,\tau)^*$ y $\alpha \in \mathbf{R}$ tales que $A \subseteq [\ell \leq \alpha]$ y $B \subseteq [\ell \geq \alpha]$.

(ii) Si (V,τ) es un e.v.t localmente convexo (lo que abreviaremos por e.v.t.l.c.) tendremos lo siguiente: Si A es cerrado y B es compacto, entonces existe un hiperplano cerrado que separa estrictamente A y B, esto es, existen $\ell \in (V,\tau)^*$, $\alpha \in \mathbf{R}$ y $\varepsilon > 0$ tales que $A \subseteq [\ell \leq \alpha - \varepsilon]$ y $B \subseteq [\ell \geq \alpha + \varepsilon]$.

Observación 1.3.2. Si (V,τ) es un e.v.t.l.c. <u>separado</u> (de Hausdorff) entonces el Teorema de Hahn-Banach permite asegurar que existen formas lineales continuas no nulas. Más aun, bajo esta condición, dado cualquier par de puntos $v_1 \neq v_2$ en V, por la parte (ii) del teorema, siempre es posible encontrar $v^* \in (V,\tau)^*$ tal que $v^*(v_1) \neq v^*(v_2)$.

Una consecuencia muy importante del Teorema de Hahn-Banach es la siguiente.

Corolario 1.3.1. *Si (V,τ) es un e.v.t.l.c. y $C \subseteq V$ es un convexo, no vacío y τ-cerrado, entonces*

$$C = \bigcap \{S \mid C \subseteq S \text{ y } S \text{ es semiespacio cerrado}\}.$$

Demostración: Dado $u \notin C$, sean $A := C$ y $B := \{u\}$. Por el Teorema de Hahn-Banach, existe S semiespacio cerrado tal que $C \subseteq S$ y $u \notin S$. □

1.3.2. Topología débil y funciones inferiormente semicontinuas.

Notemos que en principio V^* está desprovisto de estructura topológica. Sin embargo, es natural dotar V^* de la topología de la convergencia puntual sobre V, que resulta ser la topología menos fina (i.e. la más pequeña) que hace que todos los funcionales de evaluación $\langle v, \cdot\rangle$, $v \in V$, sean continuos. Esta topología se denota por $\sigma(V^*, V)$ y se llama *topología débil* de V^* asociada a la dualidad entre V y V^*. Es fácil ver que $(V^*, \sigma(V^*, V))$ resulta ser e.v.t.l.c. separado.

Un resultado útil para el estudio de la dualidad de e.v.t.l.c. es el siguiente teorema.

Teorema 1.3.2 (Banach)**.** *Sean (V, τ) un e.v.t.l.c. separado y V^* su espacio dual. Si $\ell : V^* \to \mathbf{R}$ es una forma lineal $\sigma(V^*, V)$-continua, entonces existe un único $v \in V$ tal que $\forall v^* \in V^*,\ \ell(v^*) = \langle v, v^*\rangle$. Más aun, V y $(V^*, \sigma(V^*, V))^*$ son isomorfos como espacios vectoriales (vía la evaluación $v \mapsto \langle v, \cdot\rangle$).*

Demostración: Dado $v \in V$, la aplicación $V^* \ni v^* \mapsto \langle v, v^*\rangle = v^*(v)$ es lineal y, por definición, $\sigma(V^*, V)$-continua. En este sentido, V se identifica con un subespacio vectorial de $(V^*, \sigma(V^*, V))^*$. Recíprocamente, si $\ell\colon V^* \to \mathbf{R}$ es lineal y $\sigma(V^*, V)$-continuo, entonces queremos probar que existe $v \in V$ tal que $\ell = \langle v, \cdot\rangle$ (la unicidad de v se sigue de la Observación 1.3.2). Sea $U = \{v^* \in V^* \mid \ell(v^*) < 1\}$ que, por continuidad, resulta ser vecindad del origen. Por lo tanto, $\exists\varepsilon > 0$ y un número finito de puntos $v_1, ..., v_n \in V$ tales que $\{v^* \in V^* \mid \langle v_i, v^*\rangle \le \varepsilon,\ \forall i = 1, ..., n\} \subseteq U$. En particular,

$$\bigcap_{i=1}^{n} \mathrm{Ker}\langle v_i, \cdot\rangle \subseteq \mathrm{Ker}\,\ell. \tag{1.2}$$

Definamos la función lineal $F\colon V^* \to \mathbf{R}^n$ mediante

$$F(v^*) = (\langle v_i, v^*\rangle)_{i=1}^n$$

y sea $L\colon F(V^*) \to \mathbf{R}$ con

$$L(y_1, ..., y_n) = \ell(v^*)$$

para cualquier $v^* \in V^*$ tal que $F(v^*) = (y_1, ..., y_n)$. La función L está bien definida gracias a (1.2) y resulta ser una aplicación lineal sobre el subespacio vectorial $F(V^*)$ de $\mathbf{R}^n$, y en consecuencia se puede extender a $\mathbf{R}^n$ y representar como

$$L(y) = \sum_{i=1}^{n} \alpha_i y_i.$$

En particular,

$$\forall v^* \in V^*,\ \ell(v^*) = \sum_{i=1}^{n} \alpha_i \langle v_i, v^*\rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, v^* \right\rangle,$$

de modo tal que basta tomar $v = \sum_{i=1}^{n} \alpha_i v_i$ para concluir. □

Similarmente, es posible dotar a V de la topología débil $\sigma(V, V^*)$ definida como la topología menos fina que hace que todos las formas lineales $\langle \cdot, v^*\rangle$, $v^* \in V^*$, sean continuas. Por definición, $\sigma(V, V^*)$ está contenida en la topología inicial τ, y además se tiene que $(V, \sigma(V, V^*))^*$ es isomorfo a V^*. El espacio $(V, \sigma(V, V^*))$ resulta ser e.v.t.l.c. y es separado cuando (V, τ) es un e.v.t.l.c. separado en virtud del Teorema de Hahn-Banach (ver la Observación 1.3.2).

Otra consecuencia del Teorema de Hahn-Banach, es la siguiente:

Corolario 1.3.2. *Sea (V, τ) un e.v.t.l.c. y sea V^* su espacio dual. Entonces*

(i) Si $C \subseteq V$ es convexo entonces

$$C \text{ es } \tau - \text{cerrado ssi } C \text{ es } \sigma(V, V^*) - \text{cerrado}.$$

(ii) Si $f\colon V \to \mathbf{R} \cup \{+\infty\}$ es convexa entonces

$$f \text{ es } \tau - \text{s.c.i. ssi } f \text{ es } \sigma(V, V^*) - \text{s.c.i.}$$

Demostración: (i) Como $\sigma(V, V^*) \subseteq \tau$ la suficiencia es inmediata. Para la necesidad basta observar que todo semiespacio τ-cerrado es $\sigma(V, V^*)$-cerrado y aplicar el corolario 1.3.1 para concluir que C es $\sigma(V, V^*)$-cerrado.

(ii) Basta observar que f es convexa si, y sólo si, epi(f) es convexo en $V \times \mathbf{R}$, y que f es τ-s.c.i si, y sólo si, epi(f) es cerrado para $\tau \times \tau_{\mathbf{R}}$ en $V \times \mathbf{R}$.

□

1.4. Minimización convexa en espacios de Banach

Sea $(V, \|\cdot\|)$ un espacio vectorial normado y denotemos por V^* su dual topológico. El estudio de funciones coercivas conduce naturalmente a considerar las propiedades topológicas de los subconjuntos acotados de V.

Comencemos recordando el siguiente resultado de Análisis Funcional.

Teorema 1.4.1. *Supongamos que $(V, \|\cdot\|)$ es un espacio de Banach reflexivo. Entonces los subconjuntos acotados son relativamente compactos para la topología débil $\sigma(V, V^*)$. Así, de cualquier sucesión acotada es posible extraer una subsucesión débilmente convergente.*

Demostración: [Bre83]. □

Como corolario del Teorema de Minimización de Weierstrass obtenemos el siguiente resultado.

Corolario 1.4.1. *Sean $(V, \|\cdot\|)$ un espacio de Banach reflexivo y $f\colon V \to \mathbf{R} \cup \{+\infty\}$ una función coerciva y $\sigma(V, V^*)$-s.c.i. Entonces existe $u \in V$ tal que $\forall v \in V, f(u) \leq f(v)$.*

Teorema 1.4.2. *Sea $(V, \|\cdot\|)$ un espacio de Banach reflexivo y $f\colon X \to \mathbf{R} \cup \{+\infty\}$ una función convexa, s.c.i. y coerciva. Entonces existe $u \in V$ tal que $\forall v \in V,\ f(u) \leq f(v)$.*

Demostración: Como f es coerciva, sus conjuntos de nivel inferior son acotados en V y en consecuencia relativamente compactos para la topoligía débil. Luego f es $\sigma(V, V^*)$-inf-compacta y como además es s.c.i. para la topología fuerte, lo es para $\sigma(V, V^*)$ en virtud del corolario 1.3.2. El resultado se obtiene al aplicar el Teorema de Weierstrass. □

Ejercicio 1.4.1. Pruebe el teorema anterior sin usar el Teorema de Weierstrass. Para ello, aplique el método directo con $\tau = \sigma(V, V^*)$.

En el siguiente ejemplo veremos que la reflexividad es esencial para la validez del teorema anterior.

Ejemplo 1.4.1 (Un problema de minimización convexa sin solución óptima). Consideremos $(V, \|\cdot\|) = (C([0,1];\mathbf{R}), \|\cdot\|_\infty)$, donde

$$\|v\|_\infty = \sup\{|v(t)| \mid t \in [0,1]\}.$$

Sea

$$C := \left\{ v \in V \;\middle|\; \int_0^{1/2} v(t)dt - \int_{1/2}^1 v(t)dt = 1 \right\}.$$

Es fácil ver que C es no-vacío, cerrado y convexo (más aun, C es un hiperplano cerrado). Consideremos la función indicatriz δ_C. Como

$$\Gamma_\gamma(\delta_C) = \begin{cases} \emptyset & \gamma < 0 \\ C & \gamma \geq 0 \end{cases}$$

y C es cerrado, entonces δ_C es s.c.i. (con respecto a $\tau_{\|\cdot\|_\infty}$). Como $\text{epi}(\delta_C) = C \times [0, +\infty[$, que es convexo, entonces δ_C es convexa. Consideremos el problema de minimización:

$$d(0,C) = \inf\{\|v\|_\infty \mid v \in C\} = \inf\{\|v\|_\infty + \delta_C(v) \mid v \in X\}.$$

Evidentemente, la función $f(v) = \|v\|_\infty + \delta_C(v)$ es convexa, $\|\cdot\|$-s.c.i. y coerciva. Observemos que si $v \in C$, entonces

$$1 = \int_0^{\frac{1}{2}} v(t)dt - \int_{\frac{1}{2}}^1 v(t)dt \leq \int_0^1 |v(t)|dt \leq \|v\|_\infty.$$

Así, $d(0,C) \geq 1$. Por otra parte, dados $n \geq 1$, $\alpha_n < \frac{1}{2}$, $\beta_n > 1$ definimos $v_n \colon [0,1] \to \mathbf{R}$ por

$$v_n(x) = \begin{cases} \beta_n & x \in [0, \alpha_n] \\ \frac{\beta_n}{\alpha_n - \frac{1}{2}} x + \frac{1}{2} \frac{\beta_n}{\frac{1}{2} - \alpha_n} & x \in]\alpha_n, 1 - \alpha_n[\\ -\beta_n & x \in [1 - \alpha_n, 1] \end{cases}$$

y tomando $\alpha_n = \frac{1}{2} - \frac{1}{n}$, $\beta_n = 1 + \frac{1}{n}$ se tiene que $v_n \in C$ y $\|v_n\|_\infty = \frac{n+1}{n}$ con lo que concluimos que $d(0,C) = 1$.

Supongamos que existe $u \in C$ tal que $\|u\|_\infty = 1$. Como $\left|\int_0^{1/2} u(t)dt\right| \leq \frac{1}{2}$ y $\left|\int_{1/2}^1 u(t)dt\right| \leq \frac{1}{2}$, necesariamente $\left|\int_0^{1/2} u(t)dt\right| = \left|\int_{1/2}^1 u(t)dt\right| = \frac{1}{2}$. Pero entonces $\left|\int_0^{1/2} (1 - u(t))dt\right| = 0$ y como $1 - u(t) \geq 0$ deducimos que $u \equiv 1$ sobre $\left[0, \frac{1}{2}\right]$. Similarmente, $u \equiv -1$ sobre $\left[\frac{1}{2}, 1\right]$ lo que contradice la continuidad de u. Luego, no existe un minimizador para $d(0,C)$. En particular, se deduce que $(C([0,1];\mathbf{R}), \|\cdot\|_\infty)$ no es reflexivo.

Observación 1.4.1. Se puede asegurar la unicidad del minimizador en el Teorema 1.4.2 cuando se tiene que $\inf_V f < +\infty$ si suponemos además que f es *estrictamente convexa*, es decir,

$$\forall u, v \in \text{dom}(f), u \neq v, \ \forall \lambda \in]0,1[, \ f(\lambda u + (1-\lambda)v) < \lambda f(u) + (1-\lambda) f(v).$$

Ejercicio 1.4.2. Demuestre la observación anterior.

1.5. Relajación topológica

1.5.1. La regularizada semicontinua inferior

La efectividad del enfoque topológico presentado radica en la posibilidad de escoger la topología τ de modo de satisfacer las propiedades de inf-compacidad y semicontinuidad inferior. Sin embargo, la dificultad reside en que ambas propiedades son antagonistas en el siguiente sentido: dada $f\colon X \to \mathbf{R} \cup \{+\infty\}$ y dos topologías $\tau_2 \subseteq \tau_1$ sobre X se tiene que

1. Si f es τ_1-inf-compacta entonces f es τ_2-inf-compacta.

2. Si f es τ_2-s.c.i. entonces f es τ_1-s.c.i.

La elección de la topología τ resulta de un balance entre ambas propiedades. Usualmente la inf-compacidad tiene prioridad: se trata de encontrar la topología más fina posible que asegura la inf-compacidad de f, lo que permite por una parte aumentar las posibilidades de que f sea s.c.i. y por otra describir el comportamiento asintótico de las sucesiones minimizantes. Sin embargo, en algunas aplicaciones importantes la semicontinuidad inferior simplemente no se tiene y el problema de minimización asociado no tiene solución, pese a que sí se tiene la inf-compacidad. En tales casos, es interesante entender el comportamiento de las sucesiones minimizantes. Por ejemplo responder a la pregunta ¿Qué se puede decir de los puntos de acumulación?

En lo que sigue, supondremos que (X, τ) es un espacio topológico.

Definición 1.5.1. La *τ-regularizada s.c.i.* o *τ-clausura inferior* de $f\colon X \to \mathbf{R} \cup \{+\infty\}$ es la función definida por

$$\mathrm{cl}_\tau(f) = \overline{f}^\tau := \sup\{g\colon X \to \mathbf{R} \cup \{+\infty\} \mid g \text{ es } \tau\text{-s.c.i}, \ g \leq f\}.$$

Proposición 1.5.1. *Si $f\colon X \to \mathbf{R} \cup \{+\infty\}$, entonces*

(i) $\mathrm{epi}(\mathrm{cl}_\tau(f)) = \mathrm{cl}(\mathrm{epi}(f))$. *En particular,* $\mathrm{cl}_\tau(f)$ *es τ-s.c.i.*

(ii) $\mathrm{cl}_\tau(f)(x) = \liminf_{y \to x} f(y)$.

(iii) *f es τ-s.c.i. ssi $f \leq \mathrm{cl}_\tau(f)$ ssi $f = \mathrm{cl}_\tau(f)$.*

Demostración: (i) Notemos que $A \subseteq X \times \mathbf{R}$ es un epigrafo si, y sólo si,

1. Para todos $(x, \lambda) \in A$ y $\mu > \lambda$ se tiene $(x, \mu) \in A$; y
2. Para todo $x \in X$, el conjunto $\{\lambda \mid (x, \lambda) \in A\}$ es cerrado en $\mathbf{R}$.

Así, es fácil verificar que $A = \mathrm{cl}(\mathrm{epi}(f))$ es un epígrafo, es decir, existe $g\colon X \to \mathbf{R} \cup \{+\infty\}$ tal que $\mathrm{cl}(\mathrm{epi}(f)) = \mathrm{epi}(g)$. Como $\mathrm{epi}(g)$ es cerrado, g es τ-s.c.i. Más aun, dado que $\mathrm{epi}(f) \subseteq \mathrm{epi}(g)$, tenemos que $g \leq f$. Así, g es un minorante τ-s.c.i de f y en consecuencia $g \leq \mathrm{cl}_\tau(f)$. Sea ahora $h\colon X \to \mathbf{R} \cup \{+\infty\}$ otra minorante s.c.i. de f de modo que $\mathrm{epi}(f) \subseteq \mathrm{epi}(h)$, luego $\mathrm{epi}(g) = \mathrm{cl}(\mathrm{epi}(g)) \subseteq \mathrm{epi}(h)$, y así $g \geq h$. Por lo tanto $g = \mathrm{cl}_\tau(f)$.

(ii) Como $\mathrm{cl}_\tau(f)$ es τ-s.c.i., $\mathrm{cl}_\tau(f)(x) \leq \liminf_{y\to x} \mathrm{cl}_\tau(f)(y) \leq \liminf_{y\to x} f(y)$. Por otra parte, definamos $h(x) := \liminf_{y\to x} f(y) = \sup_{V\in\mathcal{N}_x} \inf_{y\in V} f(y)$, que es una minorante de f y es τ-s.c.i., de modo que $h \leq \mathrm{cl}_\tau(f)$.

(iii) Evidentemente, $\mathrm{cl}_\tau(f) \leq f$ y si f es τ-s.c.i., entonces $f(x) \leq \liminf_{y\to x} f(y) = \mathrm{cl}_\tau f(x)$.

□

Proposición 1.5.2. *Si $f\colon X \to \mathbf{R}\cup\{+\infty\}$, entonces*

$$\mathrm{cl}_\tau(f)(x) = \min\{\liminf_{d} f(x_d) \mid D \textit{ conjunto dirigido, } (x_d)_{d\in D} \textit{ una red}, x_d \to x\}.$$

Si además (X,τ) es metrizable, entonces

$$\mathrm{cl}_\tau(f)(x) = \min\{\liminf_{n\to\infty} f(x_n) \mid (x_n)_{n\in\mathbf{N}} \textit{ es una sucesión, } x_n \to x\}.$$

Demostración: Para simplificar, sólo haremos la demostración en el caso metrizable. Tenemos que

$$\liminf_{y\to x} f(y) = \sup_{\varepsilon>0} \inf_{y\in\mathcal{B}_\tau(x,\varepsilon)} f(y).$$

Sea $x_n \to x$, de modo que $\forall\varepsilon > 0, \exists n_0(\varepsilon) \in \mathbf{N},\ \forall n \geq n_0,\ x_n \in B_\tau(x,\varepsilon)$. En consecuencia, $f(x_n) \geq \inf_{y\in B_\tau(x,\varepsilon)} f(y)$ y más aun $\inf_{n\geq n_0} f(x_n) \geq \inf_{y\in B_\tau(x,\varepsilon)} f(y)$. Luego $\liminf_{n\to\infty} f(x_n) = \sup_{k\in\mathbf{N}} \inf_{n\geq k} f(x_n) \geq \inf_{y\in B_\tau(x,\varepsilon)} f(y)$. Por lo tanto, de la arbitrariedad de ε y $x_n \to x$

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

Por otra parte, para cada $n \geq 1$ existe $x_n \in B_\tau(x,\frac{1}{n})$ tal que

$$\begin{cases} \inf_{y\in B_\tau(x,\frac{1}{n})} f(y) \geq f(x_n) - \frac{1}{n} & \text{si} \quad \inf_{B_\tau(x,\frac{1}{n})} f > -\infty \\ -n \geq f(x_n) & \text{si} \quad \inf_{B_\tau(x,\frac{1}{n})} f = -\infty. \end{cases}$$

En ambos casos

$$\liminf_{y\to x} f(y) \geq \lim_{n\to\infty} \inf_{y\in B_\tau(x,\frac{1}{n})} f(y) \geq \limsup_{n\to\infty} f(x_n) \geq \liminf_{n\to\infty} f(x_n)$$

lo que prueba el resultado. □

Proposición 1.5.3. *Sea $f\colon X \to \mathbf{R}\cup\{+\infty\}$. Entonces $\inf_X f = \inf_X \mathrm{cl}_\tau(f)$ y más generalmente $\inf_U f = \inf_U \mathrm{cl}_\tau(f)$ para todo $U \in \tau$. Además $\arg\min f \subseteq \arg\min \mathrm{cl}_\tau(f)$.*

Demostración: Como $f \geq \mathrm{cl}_\tau(f)$, sólo debemos probar que $\inf_U f \leq \inf_U \mathrm{cl}_\tau(f)$. Dado que $x \in U$, como $U \in \tau$, se tiene que $U \in \mathcal{N}_x(\tau)$ y así $\mathrm{cl}_\tau(f)(x) \geq \inf_U f$. De la arbitrariedad de x se deduce que $\inf_U \mathrm{cl}_\tau(f)(x) \geq \inf_U f$. Finalmente, si $x \in \arg\min f$ entonces

$$\mathrm{cl}_\tau(f)(x) \leq f(x) = \inf_X f = \inf_X \mathrm{cl}_\tau(f)$$

lo que implica que $x \in \arg\min \mathrm{cl}_\tau(f)$. □

Teorema 1.5.1. *Sean $f\colon X \to \mathbf{R}\cup\{+\infty\}$ y $(x_n)_{n\in\mathbf{N}}$ una sucesión minimizante para f. Supongamos que existen $\bar{x} \in X$ y una subsucesión $(x_{n_k})_{k\in\mathbf{N}}$ tales que $x_{n_k} \to \bar{x}$. Entonces $\bar{x}$ minimiza $\operatorname{cl}_\tau(f)$ en X.*

Demostración: Como $\lim\limits_{n\to\infty} f(x_n) = \inf_X f$, deducimos que

$$\operatorname{cl}_\tau(f)(\bar{x}) \leq \lim_{k\to\infty} f(x_{n_k}) = \inf_X f = \inf_X \operatorname{cl}_\tau(f).$$

$\square$

Definición 1.5.2. Decimos que el problema $\min\{\operatorname{cl}_\tau(f)(x) \mid x \in X\}$ es el problema *relajado* del problema de minimización original $\inf\{f(x) \mid x \in X\}$.

Ejercicio 1.5.1. Sea $F\colon (X,\tau) \to \mathbf{R}\cup\{+\infty\}$ y $G\colon (X,\tau) \to \mathbf{R}$ una función continua. Muestre que

$$\overline{(F+G)}^\tau = \overline{F}^\tau + G.$$

Ejemplo 1.5.1. Dado $p \in]1,+\infty[$, consideremos el siguiente problema de minimización:

$$(\mathcal{P}) \quad \inf_{v\in V} \left\{ \frac{1}{p}\int_\Omega |\nabla v(x)|^p dx - \int_\Omega g(x)v(x)dx \;\middle|\; v = 0 \text{ sobre } \Omega \right\},$$

donde $\Omega \subseteq \mathbf{R}^N$ es un abierto acotado y g es una función con propiedades a precisar. La existencia de soluciones dependerá de la elección del espacio V. Una primera idea es considerar $V = C_c^1(\Omega)$, el espacio de las funciones continuas a soporte compacto en Ω, pero esta será insuficiente si nos interesa considerar g irregular.
Sea $F\colon L^p(\Omega) \to \mathbf{R}\cup\{+\infty\}$ definida por

$$F(v) = \begin{cases} \frac{1}{p}\int_\Omega |\nabla v|^p dx & \text{si } v \in C_c^1(\Omega), \\ +\infty & \text{si no.} \end{cases}$$

Nos interesa calcular $\overline{F} = \overline{F}^{L^p}$ (más adelante veremos el porqué). Primero observemos que $v \in \operatorname{dom}(\overline{F})$ si, y sólo si, existe $v_n \to v$ en $L^p(\Omega)$ tal que $\liminf\limits_{n\to+\infty} F(v_n) < +\infty$. Sin pérdida de generalidad podemos suponer que $\lim\limits_{n\to+\infty} F(v_n)$ existe y es finito, y que $\sup_n F(v_n) < +\infty$, lo que equivale a $\sup_n \|\nabla v_n\|_{p,\Omega} < +\infty$. Luego $\sup\|v_n\|_{1,p,\Omega} < +\infty$, donde $\|v_n\|_{1,p,\Omega} = \|v_n\|_{p,\Omega} + \|\nabla v_n\|_{p,\Omega}$. Es decir $(v_n)_{n\in\mathbf{N}} \subset C_c^1(\Omega)$ está uniformemente acotada en $W^{1,p}(\Omega)$. Como éste es un espacio de Banach reflexivo, deducimos que existe $w \in W^{1,p}(\Omega)$ y una subsucesión $(v_{n_k})_{k\in\mathbf{N}}$ tal que $v_{n_k} \rightharpoonup w$ débilmente en $W^{1,p}(\Omega)$. En particular, $v_{n_k} \rightharpoonup w$ en $L^p(\Omega)$ y en consecuencia $w = v$. Así, $v_{n_k} \rightharpoonup v$ débilmente en $W^{1,p}(\Omega)$ y $(v_{n_k})_{k\in\mathbf{N}} \subseteq C_c^1(\Omega) \subseteq W_0^{1,p}(\Omega)$ y este último es subespacio cerrado de $W^{1,p}(\Omega)$, luego es débilmente cerrado. En conclusión $v \in W_0^{1,p}(\Omega)$. Así, $\operatorname{dom}(\overline{F}) \subseteq W_0^{1,p}(\Omega)$.

Recíprocamente, si $v \in W_0^{1,p}(\Omega) = \overline{C_c^1(\Omega)}^{\|\cdot\|_{1,p}}$ entonces existe una sucesión $(v_n)_{n\in N} \subset C_c^1(\Omega)$ tal que $v_n \to v$ en $W_0^{1,p}(\Omega)$ y en particular $\|\nabla v\|_{p,\Omega} = \lim\limits_{n\to\infty} \|\nabla v_n\|_{p,\Omega}$. Por lo tanto, $\operatorname{dom}(\overline{F}) = W_0^{1,p}(\Omega)$ y, más aun, si $v \in W_0^{1,p}(\Omega)$ entonces $\overline{F}(v) \leq \frac{1}{p}\|\nabla v\|_p^p$. Por otra parte, si $v \in W_0^{1,p}(\Omega)$ y $v_n \to v$ en $L^p(\Omega)$ y, razonando sin pérdida de generalidad, podemos suponer que $v_n \rightharpoonup v$ débilmente en

$W^{1,p}(\Omega)$ y en particular $\nabla v_n \rightharpoonup \nabla v$ débilmente en $L^p(\Omega)^N$. Pero $\Phi\colon L^p(\Omega)^N \to \mathbf{R}$ definida por $\Phi(f) = \frac{1}{p}\int_\Omega |f(x)|^p dx$ es convexa continua, y, por lo tanto, es débilmente s.c.i. Luego

$$\frac{1}{p}\|\nabla v\|_p^p \leq \liminf_{n\to\infty} \frac{1}{p}\|\nabla v_n\|_p^p.$$

De la arbitrariedad de las sucesiones se obtiene que $\frac{1}{p}\|\nabla v\|_p^p \leq \overline{F}$. En conclusión

$$\overline{F}(v) = \begin{cases} \frac{1}{p}\|\nabla v\|_p^p & \text{si } v \in W_0^{1,p}(\Omega), \\ +\infty & \text{si no.} \end{cases}$$

Supongamos que $g \in L^q(\Omega)$ con $\frac{1}{p} + \frac{1}{q} = 1$ y definimos $G\colon L^p(\Omega) \to \mathbf{R}$ mediante

$$G(v) := -\int_\Omega g(x)v(x)dx.$$

Luego, si definimos $J(v) := F(v) + G(v)$, del Ejercicio 1.5.1

$$\overline{J}(v) = \begin{cases} \frac{1}{p}\int_\Omega |\nabla v(x)|^p dx - \int_\Omega g(x)v(x)dx & \text{si } v \in W_0^{1,p}(\Omega), \\ +\infty & \text{si no.} \end{cases}$$

Podemos formular el problema original como

$$(\mathcal{P}) \qquad \inf_{v\in L^p(\Omega)} J(v) \leq J(0) = 0 < +\infty.$$

Sea $(v_n)_n \subset L^p(\Omega)$ una sucesión minimizante para $(\mathcal{P})$, es decir, $\lim\limits_{n\to\infty} J(v_n) = \inf_{L^p} J$. Razonando como antes podemos suponer sin pérdida de generalidad que $\sup_n J(v_n) < +\infty$ de modo que $(v_n)_{n\in\mathbf{N}} \subseteq C_c^1(\Omega) \subseteq W_0^{1,p}(\Omega)$ y $\sup_n \|\nabla v_n\|_p < +\infty$. Por la desigualdad de Poincaré, deducimos que $(v_n)_{n\in\mathbf{N}}$ está acotada en $W_0^{1,p}(\Omega)$ y por el Teorema de la Inyección Compacta de Rellich-Kondrachov, se tiene que existen $\bar{v} \in W_0^{1,p}(\Omega)$ y $(v_{n_k})_{k\in\mathbf{N}}$ tales que $v_{n_k} \to \bar{v}$ en $L^p(\Omega)$. Aplicando el Teorema 1.5.1, deducimos que $\bar{v} \in \arg\min \overline{J}$, es decir, $\bar{v}$ es una solución del problema relajado

$$(\overline{\mathcal{P}}) \qquad \inf_{v\in L^p(\Omega)} \overline{J}(v).$$

1.5.2. La Γ-convergencia

A continuación introduciremos la Γ-*convergencia* o *epi-convergencia* y mostraremos sus principales propiedades. Por ejemplo, un resultado de estabilidad para problemas de minimización.
Sea (X, d) un espacio métrico y consideremos una familia $\{F_n\}_{n\in\mathbf{N}}$ de funciones definidas de X en $\overline{\mathbf{R}}$. Para $u \in X$, definimos

$$(\Gamma(d) - \liminf_{n\to+\infty} F_n)(u) := \inf\{\liminf_{n\to+\infty} F_n(u_n) \ : \ u_n \to u\},$$

y

$$(\Gamma(d) - \limsup_{n\to+\infty} F_n)(u) := \inf\{\limsup_{n\to+\infty} F_n(u_n) \ : \ u_n \to u\}.$$

Claramente, se tiene que $\Gamma(d) - \liminf F_n \leq \Gamma(d) - \limsup F_n$ y si coinciden diremos que la sucesión $\{F_n\}_{n\in\mathbf{N}}$ es Γ- *convergente* o *epi-convergente* a F en $u \in X$. En ese caso escribimos $F(u) = (\Gamma(d) - \lim_{n\to\infty} F_n)(u)$. Notemos que dado $u \in X$, la definición es equivalente a:

1. Para cada sucesión $u_n \to u$, se tiene que

$$F(u) \leq \liminf_{n\to+\infty} F_n(u_n) \ ,$$

2. Existe una sucesión $u_n \to u$ tal que

$$F(u) = \lim F_n(u_n).$$

Ejercicio 1.5.2. Muestre que el Γ-límite de una sucesión constante es la clausura inferior. Más precisamente, $\Gamma(d) - \lim F = \mathrm{cl}_d(F)$.

Más generalmente, es posible demostrar que si $F = \Gamma(d) - \lim F_n$ entonces F es s.c.i. y que bajo ciertas condiciones sobre d, la Γ-convergencia define una topología metrizable sobre el espacio de las funciones inferiormente semicontinuas. En general, la Γ convergencia no implica ni es implicada por la convergencia puntual, más aun, es posible encontrar ejemplos de sucesiones Γ-convergentes en X cuyo límite puntual existe en X pero que no coincide con el Γ-límite.

Ejercicio 1.5.3. Sea F_n una sucesión decreciente de funciones de X en $\overline{\mathbf{R}}$. Muestre que el Γ-límite existe y

$$\Gamma(d) - \lim_{n\to+\infty} F_n = \mathrm{cl}_d(\inf_{n\in\mathbf{N}}\{F_n\}).$$

La principal propiedad de la Γ-convergencia se establece en el siguiente teorema y precisa su naturaleza variacional.

Teorema 1.5.2. *Sean $\{F_n\}_n$, F y G funciones de X en $\overline{\mathbf{R}}$ tales que*

$$F = \Gamma - \lim_{n\to+\infty} F_n.$$

y tal que G es continua. Entonces

$$\limsup_{n\to+\infty}(\inf\{F_n + G\}) \leq \inf\{F + G\}.$$

Más aun, si existe una subsucesión $u_{n_k} \in \arg\min\{F_{n_k} + G\}$ convergente a $u \in X$, entonces $u \in \arg\min\{F + G\}$ y $\inf\{F_{n_k} + G\} \to \inf\{F + G\}$.

Demostración: El lector debe verificar que $(F_n + G)_{n\in\mathbf{N}}$ es Γ-convergente a $F + G$. Luego, basta considerar el caso $G \equiv 0$. Supongamos que $\inf F > -\infty$ y para $\varepsilon > 0$ consideremos u_ε un ε-mínimo de F:

$$F(u_\varepsilon) \leq \inf F + \varepsilon.$$

Tomando $u_n \to u_\varepsilon$ tal que $\lim_{n\to+\infty} F_n(u_n) = F(u_\varepsilon)$, se tiene que

$$\limsup_{n\to+\infty}(\inf F_n) \leq \lim_{n\to+\infty} F_n(u_n) \leq \inf F + \varepsilon,$$

y, de la arbitrariedad de $\varepsilon > 0$, concluimos que $\limsup_{n\to+\infty}(\inf F_n) \leq \inf F$. Si $\inf F = -\infty$ el argumento es similar.

Sea $u_k = u_{n_k}$ la subsucesión convergente a u tal que $u_k \in \arg\min(F_{n_k})$. De la Γ-convergencia, se tiene que

$$F(u) \leq \liminf_{k\to+\infty} F_k(u_k),$$

de donde se sigue que

$$F(u) \leq \liminf_{k\to+\infty} \inf F_k$$

lo que implica, junto con la primera parte del teorema, que $u \in \arg\min F$ y que $\lim_{k\to+\infty} \inf F_{n_k} = \inf F$. $\square$

1.6. Problemas

Problema 1. Sea (X, τ) un espacio topológico y dado $x \in X$ considere la función $f(x) = \sup\{g(N) \mid N \in \mathcal{N}_x(\tau)\}$, donde g es una función a valores en $\mathbf{R} \cup \{+\infty\}$ definida sobre τ y $\mathcal{N}_x(\tau)$ denota el conjunto de τ-vecindades de x. Muestre que f es s.c.i.

Problema 2. Sean V un espacio de Banach y $U\colon V \to \mathbf{R}$ una función Gâteaux-diferenciable. Diremos que U satisface la condición (WC) sobre un subespacio $\Omega \subset V$ si para cualquier sucesión $(x_n)_{n\in\mathbf{N}} \subseteq \Omega$ con las siguientes propiedades:

- $(|U(x_n)|)_{n\in\mathbf{N}}$ es acotada,
- $U'(x_n) \neq 0$ para todo n, y
- $U'(x_n) \to 0$ en V^*,

existe $\bar{x} \in X$ tal que $U'(\bar{x}) = 0$ y

$$\liminf_{n\to+\infty} U(x_n) \leq U(\bar{x}) \leq \limsup_{n\to+\infty} U(x_n).$$

(a) Suponga que V es reflexivo y que U es convexa, s.c.i, y $U(x) \to +\infty$ cuando $\|x\| \to +\infty$. Muestre que U satisface la condición (WC) sobre V.

(b) Suponga ahora que U es s.c.i. y acotada inferiormente. Suponga también que la restricción de U' a rectas es continua y que U satisface la condición (WC) sobre V. Pruebe que U alcanza su mínimo sobre V.

Problema 3 (Teorema del Punto Fijo de Caristi)**.** Sea (E, d) un espacio métrico y $G\colon E \to 2^E$. Diremos que $\bar{x}$ es un *punto fijo* de G si $\bar{x} \in G(\bar{x})$.

(a) Supongamos que existe una función $f\colon E \to \mathbf{R} \cup \{+\infty\}$ propia, acotada inferiormente, y s.c.i. tal que para cada $x \in E$, existe un $y \in G(x)$ que satisface $f(y) + d(y, x) \leq f(x)$. Muestre que G posee al menos un punto fijo.

(b) Definamos el *grafo* de G como

$$\mathrm{Grafo}(G) = \{(x, y) \in E \times E \mid y \in G(x)\}.$$

Supongamos que $\mathrm{Grafo}(G)$ es cerrado y que existe $f\colon E \to \mathbf{R}_+ \cup \{+\infty\}$ propia tal que para cada $x \in E$, existe un $y \in G(x)$ que satisface $f(y) + d(y, x) \leq f(x)$. Muestre que G posee al menos un punto fijo.

Problema 4 (Convergencia Variacional de Puntos Sillas)**.** Diremos que una sucesión $\{F_n\colon \mathbf{R}^n \times \mathbf{R}^m \to \overline{\mathbf{R}}, n \in \mathbf{N}\}$ *hipo/epi-converge* a una funcion $F\colon \mathbf{R}^n \times \mathbf{R}^m \to \overline{\mathbf{R}}$ si para cualquier $(x, y) \in \mathbf{R}^n \times \mathbf{R}^m$ se satisface

1. Para cualquier $x_n \to x$ existe $y_n \to y$ tal que $\limsup_{n\to+\infty} F_n(x_n, y_n) \leq F(x, y)$,
2. Para cualquier $y_n \to y$ existe $x_n \to x$ tal que $\liminf_{n\to+\infty} F_n(x_n, y_n) \geq F(x, y)$.

Suponga que (F_n) hipo/epi-converge a F y que existe una secuencia de puntos $(\bar{x}_n, \bar{y}_n)$ tales que para cualquier $(x, y) \in \mathbf{R}^n \times \mathbf{R}^m$ se tiene que

$$F_n(x, \bar{y}_n) \leq F_n(\bar{x}_n, \bar{y}_n) \leq F_n(\bar{x}_n, y).$$

Muestre que si $(\bar{x}_n, \bar{y}_n) \to (\bar{x}, \bar{y})$, entonces

$$F(x, \bar{y}) \leq F(\bar{x}, \bar{y}) \leq F(\bar{x}, y).$$

Capítulo 2

Fundamentos de Análisis Convexo

2.1. Funciones convexas.

Sea V un espacio vectorial real. Recordemos que una función $f\colon V \to \overline{\mathbf{R}}$ se dice convexa si $\operatorname{epi} f$ es un subconjunto convexo de $V \times \mathbf{R}$. De manera equivalente, f es convexa si, y sólo si, $\forall n \geq 2, \forall v_1, ..., v_n \in V, \forall \lambda_1, ..., \lambda_n \in \mathbf{R}_+ \cup \{0\} : \sum\limits_{i=1}^{n} \lambda_i = 1$

$$f\left(\sum_{i=1}^{n} \lambda_i v_i\right) \leq \sum_{i=1}^{n} \lambda_i f(v_i).$$

Proposición 2.1.1. *Sea $f\colon V \to \overline{\mathbf{R}}$ una función convexa.*

(i) *Si $\lambda \geq 0$ entonces λf es convexa.*

(ii) *Si $g\colon V \to \overline{\mathbf{R}}$ es convexa entonces $f + g$ es convexa.*

(iii) *Si $A\colon W \to V$ es lineal afín entonces $f \circ A$ es convexa.*

(iv) *Si $\theta\colon \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ es convexa no-decreciente entonces $\theta \circ f$ es convexa.*

(v) *Si $(f_i)_{i\in I}$ es una familia no vacía de funciones convexas de V en $\overline{\mathbf{R}}$, entonces $f = \sup\limits_{i\in I} f_i$ es convexa.*

(vi) *Supongamos que W es un espacio vectorial y $g : V \times W \to \overline{\mathbf{R}}$ es convexa. Entonces la función $h\colon V \to \overline{\mathbf{R}}$ definida por $h(v) = \inf\limits_{w\in W} g(v, w)$ es convexa.*

(vii) *Si $g : \mathbf{R}^m \to \overline{\mathbf{R}}$ es convexa y $F\colon V \to \mathbf{R}^m$ es de la forma $F(x) = (f_1(x), ..., f_m(x))$ con $f_i\colon V \to \mathbf{R}$ convexas, entonces $g \circ F$ es convexa siempre que $g = g(y_1, ..., y_m)$ sea no-decreciente en y_i para $i = 1, ..., m$.*

Demostración: Se deja como ejercicio para el lector. $\square$

Observemos que (i) y (ii) dicen que el conjunto de las funciones convexas en V es un *cono convexo* (pero no un espacio vectorial pues la diferencia de dos funciones convexas no es necesariamente

convexa). Ejemplos de funciones convexas son las normas y las seminormas en V.
Otra definición a retener es la de concavidad: Decimos que $f\colon V \to \overline{\mathbf{R}}$ es (estrictamente) *cóncava* si $-f$ es (estrictamente) convexa.
En espacio generales, una función convexa, incluso si es lineal, no es necesariamente continua. Sin embargo, el siguiente resultado establece una propiedad interesante acerca de la continuidad y lipschitzianidad de las funciones convexas.

Teorema 2.1.1. *Sean* $(V, \|\cdot\|)$ *un e.v.n. y* $f\colon V \to \mathbf{R} \cup \{+\infty\}$ *una función convexa. Entonces son equivalentes:*

(i) f *está acotada superiormente en una vecindad de* $x_0 \in \mathrm{dom}(f)$.

(ii) f *es localmente Lipschitz en* $x_0 \in \mathrm{dom}(f)$.

(iii) f *es continua en algún* $x_0 \in \mathrm{dom}(f)$.

(iv) f *es localmente Lipschitz en* $\mathrm{int}(\mathrm{dom}(f)) \neq \emptyset$.

(v) f *es continua en* $\mathrm{int}(\mathrm{dom}(f)) \neq \emptyset$.

(vi) $\mathrm{int}(\mathrm{epi}(f)) \neq \emptyset$.

Demostración: Probaremos que $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$ y luego que $(i) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (i)$.

- $(i) \Rightarrow (ii)$ Sin pérdida de generalidad podemos suponer que $x_0 = 0$ (de lo contrario basta considerar $\bar{f}(x) = f(x + x_0)$). Sean $\varepsilon > 0$ y $M \in \mathbf{R}$ tales que para cada x con $\|x\| \leq \varepsilon$ se tiene que $f(x) \leq M$. Dados x_1, x_2 con $\|x_i\| \leq \frac{\varepsilon}{2}$ $i = 1, 2$ y $x_1 \neq x_2$ de modo que $\alpha := \|x_1 - x_2\| > 0$, definimos $y = x_1 + \frac{\varepsilon}{2\alpha}(x_1 - x_2)$. En consecuencia, $\|y - x_1\| = \frac{\varepsilon}{2}$, de donde concluimos que $\|y\| \leq \varepsilon$ y $f(y) \leq M$. Además, $x_1 = \frac{2\alpha}{2\alpha+\varepsilon} y + \frac{\varepsilon}{2\alpha+\varepsilon} x_2$ y de la convexidad de f deducimos

$$f(x_1) \leq \frac{2\alpha}{2\alpha + \varepsilon} f(y) + \frac{\varepsilon}{2\alpha + \varepsilon} f(x_2).$$

  Luego

$$f(x_1) - f(x_2) \leq \frac{2\alpha}{2\alpha + \varepsilon}[f(y) - f(x_2)] \leq \frac{2\alpha}{2\alpha + \varepsilon}[M - f(x_2)]$$

  Pero $0 = \frac{1}{2}x_2 + \frac{1}{2}(-x_2)$ de modo que $f(0) \leq \frac{1}{2}f(x_2) + \frac{1}{2}f(-x_2)$ y se tiene que $-f(x_2) \leq f(-x_2) - 2f(0) \leq M - 2f(0)$. Así,

$$f(x_1) - f(x_2) \leq \frac{4\alpha}{2\alpha + \varepsilon}[M - f(0)] \leq \frac{4\bar{M}}{\varepsilon}\|x_1 - x_2\|,$$

  para algún $\bar{M} > 0$.

  Intercambiando los roles de x_1 y x_2 se deduce que $|f(x_1) - f(x_2)| \leq \frac{4\bar{M}}{\varepsilon}\|x_1 - x_2\|$.

- $(ii) \Rightarrow (iii)$ y $(iii) \Rightarrow (i)$ son inmediatos.

- $(i) \Rightarrow (iv)$. Sea $\bar{x} \in \text{int}(\text{dom}(f))$. Consideremos $\varepsilon > 0$ y definamos $y_\varepsilon = \bar{x} + \varepsilon(\bar{x} - x_0)$ de modo que $\bar{x} = \frac{1}{1+\varepsilon} y_\varepsilon + \frac{\varepsilon}{1+\varepsilon} x_0$. Si $\varepsilon > 0$ es suficientemente pequeño, $y_\varepsilon \in \text{dom}(f)$. Sea U una vecindad abierta de x_0 donde f es acotada superiormente por M y definamos

$$U_\varepsilon := \frac{1}{1+\varepsilon} y_\varepsilon + \frac{\varepsilon}{1+\varepsilon} U$$

que resulta ser una vecindad abierta de $\bar{x}$. Luego, si $z \in U_\varepsilon$ se tiene que $\exists x \in U : \ z = \frac{1}{1+\varepsilon} y + \frac{\varepsilon}{1+\varepsilon} x$ y por la convexidad de f

$$f(z) \leq \frac{1}{1+\varepsilon} f(y) + \frac{\varepsilon}{1+\varepsilon} f(x) \leq \frac{1}{1+\varepsilon} f(y) + \frac{\varepsilon}{1+\varepsilon} M := M_\varepsilon,$$

de modo que f está localmente acotada por M_ε en $\bar{x}$.

- $(iv) \Rightarrow (v)$ También es inmediato.

- $(v) \Rightarrow (vi)$ Sean $x \in \text{int}(\text{dom}(f))$ y $\lambda > f(x)$. Veremos que $(x, \lambda) \in \text{int}(\text{epi}(f))$. En efecto, dado $\gamma \in \mathbf{R}$ tal que $f(x) < \gamma < \lambda$, existe una vecindad abierta U de x tal que $\forall y \in U, \ f(y) < \gamma$. Luego $(x, \lambda) \in U \times]\gamma, +\infty[\subseteq \text{epi}(f)$.

- $(vi) \Rightarrow (i)$ Dados U abierto no vacío y $a < b$ tales que $U \times]a, b[\subseteq \text{epi}(f)$ entonces $\forall x \in U, \ (x, \frac{a+b}{2}) \in \text{epi}(f)$ lo que equivale a decir que $\forall x \in U, \ f(x) \leq \frac{a+b}{2} < +\infty$.

$\square$

Este resultado puede adaptarse al caso $f \colon V \to \overline{\mathbf{R}}$ pero hay que exigir que $f(x_0) \in \mathbf{R}$.

Corolario 2.1.1. *Si $f \colon \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ es convexa y propia entonces f es localmente Lipschitz en* $\text{int}(\text{dom}(f))$. *En particular, si $f \colon \mathbf{R}^n \to \mathbf{R}$ es convexa entonces es continua.*

Demostración: Sea $x_0 \in \text{int}(\text{dom}(f))$ de modo que $\exists \varepsilon > 0 : B(x_0, \varepsilon) \subseteq \text{dom}(f)$. Sean $\hat{e}_1, ..., \hat{e}_n$ los vectores de la base canónica de $\mathbf{R}^n$ y $x_i := x_0 + \varepsilon \hat{e}_i$ para $i = 1, ..., n$. El conjunto

$$S = \left\{ \sum_{i=0}^{n} \lambda_i x_i \,\middle|\, \sum_{i=0}^{n} \lambda_i = 1, \ \lambda_i > 0, \ i \in \{0, 1, ..., n\} \right\}$$

es una vecindad abierta de x_0, y por convexidad

$$f\left(\sum_{i=0}^{n} \lambda_i x_i \right) \leq \max\{ f(x_i) \mid i = 0, 1, ..., n \} < +\infty.$$

Por lo tanto f está acotada superiormente en S. $\square$

Notemos que la convexidad es esencialmente una propiedad unidimensional pues depende del comportamiento en un segmento de recta. Por ejemplo, un conjunto es convexo si, y sólo si, su intersección con cualquier recta lo es. Así, muchas de las propiedades de funciones convexas en $\mathbf{R}^n$ pueden obtenerse de un análisis del caso $n = 1$.

Lema 2.1.1. *Sea I un intevalo. Una función $f : I \to \mathbf{R}$ es convexa si, y sólo si, para todo $x_0 \leq y \leq x_1$ en I se tiene que*

$$\frac{f(y)-f(x_0)}{y-x_0} \leq \frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x_1)-f(y)}{x_1-y}.$$

Luego, dado $x \in I$, se tiene que

$$\Delta_x(y) := \frac{f(y)-f(x)}{y-x}$$

es no-decreciente como función de $y \in I \backslash \{x\}$. Similarmente la convexidad estricta está caracterizada por desigualdades estrictas.

Demostración: Basta observar que $y = \frac{x_1-y}{x_1-x_0}x_0 + \frac{y-x_0}{x_1-x_0}x_1$. □

Teorema 2.1.2. *Si $f : I \subseteq \mathbf{R} \to \mathbf{R}$ es diferenciable en el intervalo abierto I entonces cada una de las siguientes condiciones son necesarias y suficientes para que f sea convexa en I:*

(i) f' *es no-decreciente en* I.

(ii) $f(y) \geq f(x) + f'(x)(y-x), \ \forall x, y \in I$.

(iii) En caso de que f sea de clase C^2, $f''(x) \geq 0, \forall x \in I$.

Similarmente, cada una de las siguientes condiciones son necesarias y suficientes para que f sea estrictamente convexa en I:

(i') f' *es estrictamente creciente en* I.

(ii') $f(y) > f(x) + f'(x)(y-x), \ \forall x, y \in I$ *con* $x \neq y$.

Una condición suficiente pero no necesaria para la convexidad estricta es:

(iii') $f''(x) > 0, \ \forall x \in I$ *(asumiendo que $f \in C^2$).*

Demostración:

- $(convexidad) \Rightarrow (i)$ Directo del lema anterior.
- $(i) \Rightarrow (ii)$ Definamos $g_x(y) := f(x) - f(y) + f'(x)(y-x)$. Notemos que $g_x(x) = 0$ y que $g'_x(y) = -f'(y)+f'(x)$ de modo que $g'_x(y) \geq 0$ si $y \in I \cap]-\infty, x]$ y $g'_x(y) \leq=$ si $y \in I \cap [x, +\infty]$. En consecuencia, g_x tiene un máximo golbal en $y = x$ y el valor es 0.
- $(ii) \Rightarrow$ $(convexidad)$ Sea $l_x(y) = f(x) + f'(x)(y-x)$ una función lineal afín. Tenemos que $\forall x \in I, \ f(y) \geq l_x(y)$ y $f(x) = l_x(x)$. Así, $f(y) = \sup_{y \in I} l_x(y)$ y en consecuencia f es convexa.
- $(ii') \Rightarrow$ $(convexidad\ estricta)$ Dados $x_0 < x_1$ sea $x_\lambda = x_0 + \lambda(x_1 - x_0)$. Para la función afín $l_\lambda(y) = f(x_\lambda) + f`(x_\lambda(y-x_\lambda)$ tenemos que $f(x_0) > l_\lambda(x_0)$ y $f(x_1) > l_\lambda(x_1)$ pero$f(x_\lambda) = l - \lambda(x_\lambda = \lambda l_\lambda(x_1) + (1-\lambda)l_\lambda(x_0)$ luego $f(x_\lambda) < (1-\lambda)f(x_0) + \lambda f(x_1)$.

Queda como ejercicio estudiar las otras implicancias.

□

Ejemplo 2.1.1. Algunas funciones convexas relevantes:

- $f(x) = ax^2 + bx + c$ en $\mathbf{R}$ cuando $a \geq 0$; es estrictamente cuando $a > 0$.
- $f(x) = e^{ax}$ en $\mathbf{R}$; es estrictamente cuando $a \neq 0$.
- $f(x) = x^\alpha$ en $]0, +\infty[$ cuando $\alpha \geq 1$; es estrictamente convexa cuando $\alpha > 1$.
- $f(x) = -x^\alpha$ en $]0, +\infty[$ cuando $0 \leq \alpha \leq 1$; es estrictamente convexa cuando $0 < \alpha < 1$.
- $f(x) = -\ln x$ en $]0, +\infty[$ es estrictamente convexa.

Notemos que $f(x) = x^4$ satisface que $f''(0) = 0$ pero es estrictamente convexa, lo que muestra que (iii') no es necesario.

Corolario 2.1.2. *Sea $U \subseteq \mathbf{R}^n$ un conjunto convexo y abierto. Si $f\colon U \to \mathbf{R}$ es diferenciable, entonces cada una de las siguientes condiciones son necesarias y suficientes para que f sea convexa en U:*

(i) $\forall x, y \in U,\ \langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$.

(ii) $\forall x, y \in U,\ f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.

(iii) $\nabla^2 f(x)$ *es semi-definida positiva* $\forall x \in U$ *(asumiendo que* $f \in C^2$*).*

Para la convexidad estricta es necesario y suficiente tener (i) o (ii) con desigualdad estricta si $x \neq y$. Una condición suficiente pero no necesaria es que $\nabla^2 f(x)$ sea definida positiva.

Demostración: Propuesto. Considerar $g(t) = f(y + tz)$ con $y \in U$ y $z \in \mathbf{R}^n$. □

Ejercicio 2.1.1. Pruebe que las siguientes funciones son convexas:

- $f(x) = \frac{1}{2}\langle x, Ax \rangle + \langle b, x \rangle + c$, con $A \in \mathbf{R}^{n\times n}$ matriz simétrica y semi-definida positiva. Pruebe que es estrictamente convexa si A es definida positiva.
- $f(x) = \ln\left(\sum\limits_{i=1}^{n} e^{x_i}\right)$. Muestre además que esta función no es estrictamente convexa.

2.2. Espacios en dualidad

Diremos que dos e.v.t.l.c. (X, τ), (Y, σ) son *espacios en dualidad* si existe un *producto de dualidad* entre X y Y, esto es, una función bilineal $\langle \cdot, \cdot \rangle\colon X \times Y \to \mathbf{R}$ tal que:

1.1 $\forall x \in X \setminus \{0\}, \exists y \in Y :\ \langle x, y \rangle \neq 0$.

1.2 $\forall y \in Y \setminus \{0\}, \exists x \in X :\ \langle x, y \rangle \neq 0$.

2.1 $\forall y \in Y,\ \langle \cdot, y \rangle \in (X, \tau)^*$ y $\forall \ell \in (X, \tau)^*, \exists! y \in Y :\ \ell = \langle \cdot, y \rangle$.

2.2 $\forall x \in X,\ \langle x, \cdot \rangle \in (Y, \sigma)^*$ y $\forall \ell \in (Y, \sigma)^*, \exists! x \in X :\ \ell = \langle x, \cdot \rangle$.

Decimos que 1.1 y 1.2 son propiedades que hacen a $\langle\cdot,\cdot\rangle$ *separar* puntos. Las topologías τ y σ se dirán *compatibles con la dualidad* $(X, Y, \langle\cdot,\cdot\rangle)$ y necesariamente son separadas: dados $x_1 \neq x_2$ en X se tiene, por 1.1, que existen $\alpha \in \mathbf{R}$ y $y \in Y$ tales que $\langle x_1, y\rangle > \alpha > \langle x_2, y\rangle$. En virtud de (2.1), se tiene que $x_1 \in \{x \mid \langle x, y\rangle > \alpha\} \in \tau$ y $x_2 \in \{x \mid \langle x, y\rangle < \alpha\} \in \tau$. Análogamente se prueba que la topología σ es separada.

Ejemplo 2.2.1 (Espacios en Dualidad)**.** Los siguientes son algunos ejemplos de espacios en dualidad:

- $X = Y$ espacios de Hilbert cuya topología esta inducida por el producto hilbertiano $\langle\cdot,\cdot\rangle$ que a su vez resulta ser el producto de dualidad.
- $(X, \|\cdot\|)$ un e.v.n., $Y = (X, \tau_{\|\cdot\|})^* = X^*$, $\langle\cdot,\cdot\rangle\colon X \times X^* \to \mathbf{R}$ definido por $\langle x, x^*\rangle = x^*(x)$.
  - 1.1 Se tiene por Hahn-Banach.
  - 1.2 Se tiene pues $x^* \neq 0$.
  - 2.1 Se tiene por definición.
  - 2.2 Depende de la topología en X^*:
    - Si $\sigma = \tau_{\|\cdot\|^*}$ entonces 2.2 equivale a que X sea un Banach reflexivo.
    - Si $\sigma = \sigma(X^*, X)$ es la topología débil entonces 2.2 se cumple gracias al Teorema 1.3.2.
- Más generalmente, si X e Y son e.v. y $\langle\cdot,\cdot\rangle\colon X \times Y \to \mathbf{R}$ una función bilineal que separa puntos, entonces $(X, \sigma(X, Y))$ y $(Y, \sigma(Y, X)$ son espacios en dualidad, donde $\sigma(X, Y)$ es la topología más pequeña que hace todos los $\langle\cdot, y\rangle$ continuos.

En las aplicaciones, prácticamente la única situación útil es el caso de $(X, X^*, \langle\cdot,\cdot\rangle)$ con X espacio de Banach y $\langle\cdot,\cdot\rangle$ asociado a las evaluaciones, caso en el que, a modo de resumen, tenemos que:

1. $(X, \tau_{\|\cdot\|})$ es compatible con la dualidad $\langle\cdot,\cdot\rangle$.
2. $(X, \sigma(X, X^*))$ es compatible con la dualidad $\langle\cdot,\cdot\rangle$.
3. $(X^*, \sigma(X^*, X)$ es compatible con la dualidad $\langle\cdot,\cdot\rangle$.
4. $(X^*, \tau_{\|\cdot\|^*})$ es compatible con la dualidad $\langle\cdot,\cdot\rangle$ si, y sólo si, X es reflexivo.

Este caso no es el único considerado en este capítulo porque deseamos enfatizar que el análisis sólo depende de las topologías (débiles) involucradas.

2.3. La conjugada de Fenchel

Sean (X, τ) y (Y, σ) dos e.v.t.l.c. en dualidad $\langle\cdot,\cdot\rangle$ y $f\colon X \to \overline{\mathbf{R}}$ una función. Vimos que en el caso en que la dimensión es finita y la función es diferenciable la convexidad puede caracterizarse vía minorantes lineales afines asociados al gradiente de la función. En el caso general, decimos que dado $y \in Y$ y $\alpha \in \mathbf{R}$ la función lineal $l_\alpha(x) = \langle x, y\rangle - \alpha$ es una *minorante* de f si

$$\forall x \in X,\ \langle x, y\rangle - \alpha \leq f(x),$$

o equivalentemente

$$\alpha \geq \sup_{x\in X}\{\langle x, y\rangle - f(x)\}.$$

Definición 2.3.1 (Fenchel, 1949)**.** La *conjugada de Fenchel* de $f\colon X \to \overline{\mathbf{R}}$ es la funcion $f^*\colon Y \to \overline{\mathbf{R}}$ definida por

$$f^*(y) = \sup_{x\in X}\{\langle x, y\rangle - f(x)\}.$$

Observación 2.3.1.

1. La definición equivale a

$$f^*(y) = \sup_{x\in \mathrm{dom}(f)}\{\langle x, y\rangle - f(x)\}$$

incluso si $\mathrm{dom}(f) = \emptyset$.

2. Si $\exists x_0 \in X$ tal que $f(x_0) = -\infty$ entonces $f^* \equiv +\infty$.
3. f^* es convexa y s.c.i. para cualquier topología sobre Y compatible con la dualidad $\langle\cdot,\cdot\rangle$.
4. Se cumple la *desigualdad de Young-Fenchel*:

$$\forall x \in X, \forall y \in Y,\ f(x) + f^*(y) \geq \langle x, y\rangle.$$

Interpretaciones

- Geométrica. Dado $y \in Y$, $f^*(y)$ es el menor valor α que se debe restar a $\langle\cdot, y\rangle$ para que $\langle\cdot, y\rangle - \alpha$ sea una minorante lineal afín de f.
- Minimización. Notemos que $f^*(0) = -\inf_X f$ y que, más generalmente, $-f^*(y) = \inf_{x\in X}\{f(x) - \langle x, y\rangle\}$ es el valor óptimo de un problema perturbado linealmente por $-\langle\cdot, y\rangle$.
- Económica. Si X es un espacio de bienes , Y es un espacio de precios y f es una función de costos de producción, entonces $\langle x, y\rangle - f(x)$ es el beneficio de producir x cuando los precios están dados por y y por lo tanto $f^*(y)$ es el máximo beneficio asociado al precio y.

Proposición 2.3.1.

1. *Si $f \leq g$ entonces $f^* \geq g^*$.*
2. *Si $(f_i)_{i\in I}$ es una familia de funciones de X en $\overline{\mathbf{R}}$ entonces*

$$\left(\inf_{i\in I} f_i\right)^* = \sup_{i\in I} f_i^*, \quad y \quad \left(\sup_{i\in I} f_i\right)^* \leq \inf_{i\in I} f_i^*.$$

3. *Dado $\lambda > 0$, $(\lambda f)^*(y) = \lambda f^*(\frac{y}{\lambda})$.*
4. *Dado $\alpha \in \mathbf{R}$, $(f + \alpha)^* = f^* - \alpha$.*
5. *Dado $x_0 \in X$, si definimos $f_{x_0}(x) = f(x - x_0)$ entonces $f^*_{x_0}(y) = f^*(y) + \langle x_0, y\rangle$.*

6. *Si* $A\colon X \to X$ *es lineal continua y biyectiva, entonces* $(f \circ A)^* = f^* \circ A^{*-1}$.

7. *Dado* $y_0 \in Y$*, si definimos* $f_{y_0}(x) = f(x) + \langle x, y_0 \rangle$ *entonces* $f^*_{y_0}(y) = f^*(y - y_0)$-

Demostración: Propuesto. □

Ejemplo 2.3.1. Sea $f(x) = \exp(x)$ definida sobre $\mathbf{R}$. Se tiene que

- $f^*(0) = 0$.
- Si $y > 0$ entonces $f^*(y) \geq t(-y) - \exp(t) \to \infty$ cuando $t \to -\infty$ y, por lo tanto, $f^*(y) = +\infty$.
- Si $y < 0$ entonces $f^*(y)$ se obtiene maximizando una función cóncava diferenciable. Luego $f^*(y) = y \ln y - y$.

Así,

$$f^*(y) = \begin{cases} y \ln y - y & \text{si } y \geq 0, \\ +\infty & \text{si } y < 0. \end{cases}$$

Ésta se conoce como la función de entropía de Boltzmann-Shannon.

Ejemplo 2.3.2. Si $\varphi\colon \mathbf{R} \to \overline{\mathbf{R}}$ es una función convexa y $(V, \|\cdot\|)$ es una e.v.n. en dualidad con V^* entonces la conjugada de $f\colon V \to \overline{\mathbf{R}}$ definida por $f(v) = \varphi(\|v\|)$ está dada por

$$f^*(v^*) = \varphi^*(\|v^*\|_*).$$

En efecto, dado $v^* \in V^*$ se tiene que

$$\begin{aligned} f^*(v^*) &= \sup_{v \in V} \{\langle v, v^* \rangle - \varphi(\|v\|)\} \\ &= \sup_{t \geq 0} \sup_{\|v\| = t} \left\{ t \langle \frac{v}{t}, v^* \rangle - \varphi(t) \right\} \\ &= \sup_{t \geq 0} \{t \|v^*\|_* - \varphi(t)\} \\ &= \sup_{t \in \mathbf{R}} \{t \|v^*\|_* - \varphi(t)\} \\ &= \varphi^*(\|v^*\|_*) \end{aligned}$$

donde la última igualdad se sigue del hecho que φ es par. Un caso de particular importancia es cuando consideramos $p \in]1, +\infty[$ y $\varphi(x) = \frac{1}{p}|x|^p$. Se tiene que $\varphi^*(y) = \frac{1}{q}|y|^q$ con $\frac{1}{p} + \frac{1}{q} = 1$ por lo que deducimos que si $f(v) = \frac{1}{p}\|v\|^p$ entonces $f^*(v^*) = \frac{1}{q}\|v^*\|_*^q$. En particular, si $(V, \|\cdot\|) = (L^p(\Omega), \|\cdot\|_p)$ entonces $(V^*, \|\cdot\|_*) = (L^q(\Omega), \|\cdot\|_q)$. En consecuencia si

$$F(f) = \frac{1}{p}\|f\|_p^p$$

entonces

$$F^*(g) = \frac{1}{q}\|g\|_q^q.$$

Ejercicio 2.3.1. Muestre que si $X = Y$ son espacios de Hilbert y $f(x) = \frac{1}{2}\|x\|^2 = \langle x, x\rangle$, entonces $f(y) = f^*(y)$. Pruebe también que $\frac{1}{2}\|x\|^2$ es la única función con esta propiedad.

Ejemplo 2.3.3 (Función Soporte)**.** Sea K un subconjunto de X y

$$\delta_K(x) = \begin{cases} 0 & \text{si } x \in K, \\ +\infty & \text{si } x \notin K \end{cases}$$

su función característica. La función soporte de K está definida por

$$\sigma_K(y) := \delta_K^*(y) = \sup_{x \in K} \langle x, y\rangle.$$

Los siguientes casos son de interés

1. Si $K = \{x_0\}$ entonces $\sigma_{\{x_0\}}(y) = \langle x_0, y\rangle$.

2. Si $(X, \|\cdot\|)$ es un e.v.n. y $K = \{x \in X \mid \ \|x\| \leq 1\}$ entonces $\sigma_K(y) = \|y\|_*$.

3. Si K es un cono entonces $\sigma_K(y) = \delta_{K^0}(y)$ donde

$$K^0 = \{y \in Y \mid \ \forall x \in K, \ \langle x, y\rangle \leq 0\}$$

   es el *cono polar* de K. Si K es un cono convexo cerrado entonces $(K^0)^0 = K$ y además $K = \{x \in X \mid \forall y \in Y, \ \langle x, y\rangle \leq \sigma_K(y)\}$.

4. Si K es un subespacio vectorial entonces $\sigma_K(y) = \delta_{K^\perp}(y)$, donde

$$K^\perp = \{y \in Y \mid \forall x \in K, \ \langle x, y\rangle = 0\}$$

   es el subespacio *ortogonal* a K. Si además K es cerrado, $(K^\perp)^\perp = K$.

Proposición 2.3.2. *Si $K \subseteq X$ es un conjunto entonces σ_K es convexa, s.c.i. y positivamente homogénea. Recíprocamente, cualquier función σ con estas características es la función soporte del conjunto*

$$K = \{x \in X \mid \ \forall y \in Y, \ \langle x, y\rangle \leq \sigma(y)\}$$

Demostración: Propuesto. □

Volvamos al caso general de una dualidad $(X, Y, \langle\cdot,\cdot\rangle)$.

Definición 2.3.2. A continuación definiremos dos espacios muy importantes en análisis convexo.

- $\Gamma(X) = \{f\colon X \to \overline{\mathbf{R}} \mid f \text{ es supremo de lineales afines continuas } \}$

- $\Gamma_0(X) = \Gamma(X) \setminus \{\overline{\omega}, \underline{\omega}\}$ donde $\overline{\omega} \equiv +\infty, \ \underline{\omega} \equiv -\infty$.

Teorema 2.3.1. *$f \in \Gamma_0(X)$ si, y sólo si, f es convexa, s.c.i. y propia.*

Demostración: Ya hemos demostrado la necesidad. Veamos la suficiencia: sea f convexa, s.c.i. y propia, de modo que su epígrafo es convexo, cerrado y no vacío. Consideremos $(x,r) \notin \mathrm{epi}(f)$. De acuerdo con el Teorema de Hahn-Banach, podemos escoger $(y,s) \in Y \times \mathbf{R} \setminus \{(0,0)\}$ y $\alpha \in \mathbf{R}$ tales que

$$\forall (z,\lambda) \in \mathrm{epi}(f),\ \langle x,y\rangle + sr < \alpha \leq \langle z,y\rangle + s\lambda.$$

Notemos que $s \geq 0$ pues de lo contrario, tomando $x_0 \in \mathrm{dom}(f)$ y $\lambda > f(x_0)$ suficientemente grande, contradeciríamos la desigualdad. Basta, entonces, distinguir dos casos:

1. $x \in \mathrm{dom}(f)$. Tomando $\lambda = f(x)$ obtenemos que

$$\langle x,y\rangle + sr < \alpha \leq \langle x,y\rangle + sf(x)$$

y, en consecuancia, $s(f(x)-r) > 0$ lo que implica que $s > 0$ y que el hiperplano separador no es vertical. En particular tenemos que $\forall z \in \mathrm{dom}(f)$ se cumple que

$$\langle x,\frac{y}{s}\rangle + r < \frac{\alpha}{s} \leq \langle z,\frac{y}{s}\rangle + f(z)$$

y definiendo $\ell(z) := \langle z,-\frac{y}{s}\rangle + \frac{\alpha}{s}$ se tiene que $f \geq \ell$ y $f(x) \geq \ell(x) > r$.

2. $x \notin \mathrm{dom}(f)$. Si $s > 0$ podemos razonar como en 1. El problema es que ahora no podemos asegurar que $s \neq 0$. Estudiemos este caso: Si $s = 0$ ó, el hiperplano es vertical y se tiene que

$$\forall (z,\lambda) \in \mathrm{epi}(f),\ \langle x,y\rangle < \alpha \leq \langle z,y\rangle.$$

Razonando como antes (pues $\mathrm{dom}(f) \neq \phi$) se tiene que existe $\ell \in (X,\tau)^*$ tal que

$$\forall z \in X,\ f(z) \geq \ell(z).$$

Así, $\forall k \geq 1$, $\forall z \in X$ se tiene que

$$f(z) \geq \ell(z) \geq \ell(z) + k(\alpha - \langle z,y\rangle)$$

y

$$\ell(x) + k(\alpha - \langle x,y\rangle) \to +\infty$$

cuando $k \to +\infty$.

Hemos demostrado que $\forall x \in X, \forall r < f(x), \exists \ell \in (X,\tau)^* : f \geq \ell,\ f(x) \geq \ell(x) > r$, lo que implica que $f \in \Gamma(X)$ y $f > -\infty$ y $f \neq +\infty$ pues es propia. $\square$

Notemos que dado $f\colon X \to \overline{\mathbf{R}}$ se tiene que $f^* \in \Gamma(X)$.

Definición 2.3.3. Dada $f\colon X \to \overline{\mathbf{R}}$ definimos la *biconjugada* $f^{**}\colon X \to \overline{\mathbf{R}}$ mediante

$$f^{**}(x) = \sup_{y\in Y}\{\langle x,y\rangle - f^*(y)\}.$$

Notemos que $f^{**} \in \Gamma(X)$ y se tiene que $f^{**} \leq f$ por la Desigualdad de Young-Fenchel.

Proposición 2.3.3. *Dada $f\colon X \to \overline{\mathbf{R}}$ se tiene que*

$$f^{**}(x) = \sup\{g(x) \mid g \in \Gamma(X),\ g \leq f\}$$

de modo que f^{**} es la mayor función en $\Gamma(X)$ que minora a f, por lo que habitualmente de llama Γ-regularizada *de f. En particular,*

$$f \in \Gamma(X) \quad ssi \quad f = f^{**}.$$

Demostración: El conjunto $\Gamma(X)$ es estable bajo supremos, luego

$$h = \sup\{g(x) \mid g \in \Gamma(X),\ g \leq f\} \in \Gamma(X)$$

y, por lo discutido anteriormente, $f^{**} \leq h$. Sea $g \in \Gamma(X)$ con $g \leq f$, luego $g^* \geq f^*$ y en consecuencia $g^{**} \leq f^{**}$. De la arbitrariedad de $g \in \Gamma(X)$, basta probar que $g^{**} = g$. Como $g \in \Gamma(X)$ existe una familia $\{\ell_i\}_{i\in I}$ de funciones lineales afines continuas tales que

$$g = \sup \ell_i.$$

De la dualidad, podemos encontrar $(y_i, r_i) \in Y \times \mathbf{R}$ tales que $\ell_i(x) = \langle x, y_i\rangle - r_i$ de modo que

$$\ell_i^*(y) = \begin{cases} r_i & \text{si } y_i = y, \\ +\infty & \text{si no.} \end{cases}$$

Como $g \geq \ell_i$ se tiene que $\forall i \in I, \forall x \in X,\ g^{**}(x) \geq \ell_i^{**}(x) = \langle x, y_i\rangle - r_i = \ell_i(x)$. Luego, $g^{**} \geq \sup_{i\in I} \ell_i = g$ lo que implica que $g^{**} = g$. □

Observación 2.3.2.

1. Si $\bar{f}$ denota la regularizada s.c.i. de $f\colon X \to \overline{\mathbf{R}}$, entonces

$$f^{**} \leq \bar{f} \leq f.$$

   Si f es convexa entonces $\bar{f}$ también lo es (pues $\text{epi}(\bar{f}) = \overline{\text{epi}(f)}$ y la adherencia de un convexo es convexa) pero no necesariamente se tiene que $f^{**} = \bar{f}$. En general, las únicas funciones $f\colon X \to \overline{\mathbf{R}}$ convexas y s.c.i. que alcanzan el valor $-\infty$ son de la forma

$$f(x) = \begin{cases} -\infty & x \in C, \\ +\infty & x \notin C, \end{cases}$$

   con C un convexo cerrado. Si $C \neq X$ se tiene que $f = \bar{f} \geq f^{**} \equiv -\infty$ y $\bar{f} \neq f^{**}$ pese a que f es convexa y s.c.i. A modo de resumen, si f es una función convexa, entonces $\bar{f} = f^{**}$ si, y sólo si, o bien f admite una minorante lineal afín continua, o su regularizada s.c.i. es $\underline{\omega}$. Hay veces en las cuales $\bar{f}$ es convexa pese a que f no lo es (este es el caso en algunos problemas de cálculo de variaciones).

2. Como $\underline{\omega}_X^* \equiv +\infty = \overline{\omega}_Y$ y $\overline{\omega}_X^* \equiv -\infty = \underline{\omega}_Y$, hemos probado que la transformación ${}^*\colon \Gamma_0(X) \to \Gamma_0(Y)$ es uno a uno. De hecho, su inversa es ${}^*\colon \Gamma_0(Y) \to \Gamma_0(X)$. La operación * se conoce como *transformada de Legendre-Fenchel*.

2.4. El subdiferencial

En análisis, la herramienta fundamental para obtener condiciones necesarias de optimalidad es la célebre *Regla de Fermat*, que en el caso diferenciable, dice que la derivada se anula en un punto de mínimo o de máximo local. Esto significa que la recta tangente al grafo de la función diferenciable $f\colon \mathbf{R} \to \mathbf{R}$ que pasa por $(x_0, f(x_0))$ tiene pendiente 0 si x_0 es un punto extremo (mínimo o máximo local). Generalizada apropiadamente, esta regla es válida en diversos contextos, que van desde programación matemática, cálculo de variaciones (Ecuación de Euler-Lagrange) hasta control óptimo. Daremos una versión de esta regla en el caso convexo sin hipótesis de diferenciabilidad (en muchas aplicaciones, la función objetivo no es diferenciable).
Por otra parte, el cálculo diferencial es una herramiente muy flexible en análisis, que necesita de la noción de derivada (o, más generalmente, de gradiente). Cuando la diferenciabilidad en el sentido clásico falla, es natural preguntarse si es posible extender la noción de derivada de modo de recuperar al menos algunas de sus propiedades. Un ejemplo de esta situación es la teoría de distribuciones, donde la propiedad que se pretende preservar es la regla de integración por partes.
Todo lo anterior nos motiva a estudiar una noción de diferenciación en el caso convexo orientada a la resolución de problemas de minimización.
Consideremos espacios (X, τ) y (Y, σ) compatibles con la dualidad $\langle \cdot, \cdot \rangle$ y una función $f \in \Gamma_0(X)$ de modo que

$$f(x_0) = f^{**}(x_0) = \sup_{y \in Y} \{\langle x, y \rangle - f^*(y)\}.$$

Supongamos que el supremo se alcanza en $y_0 \in Y$ y que es finito, es decir

$$f(x_0) = \langle x_0, y_0 \rangle - f^*(y_0) \in \mathbf{R}.$$

Por otra parte, sabemos que $\forall x \in X$,

$$f(x) \geq \langle x, y_0 \rangle - f^*(y_0) = \langle x - x_0, y_0 \rangle - f^*(x_0),$$

lo que equivale a decir que la función lineal afín continua $\ell(x) := \langle x - x_0, y_0 \rangle - f^*(x_0)$ es una minorante de f que coincide con f en x_0. En general, no basta que $f(x_0) \in \mathbf{R}$ para que esto ocurra. En efecto, basta considerar

$$f(t) = \begin{cases} -\sqrt{t} & \text{si } t \geq 0, \\ +\infty & \text{si no} \end{cases}$$

que está en $\Gamma_0(\mathbf{R})$, es finita en 0 y no admite minorantes afines que pasen por (0,0).

Definición 2.4.1. Sean $f\colon X \to \mathbf{R}$ una función convexa[1] y $x_0 \in X$. Diremos que $y \in Y$ es un *subgradiente* de f en x_0 si

$$\forall x \in X,\ f(x_0) + \langle x - x_0, y \rangle \leq f(x).$$

El conjunto de subgradientes se denota $\partial f(x_0)$ y se llama *subdiferencial* de f en x_0. Si $\partial f(x_0) \neq \emptyset$ decimos que f es *subdiferenciable* en x_0.

Ejemplo 2.4.1. Si $f(x) = |x|$, $x \in \mathbf{R}$, entonces $\partial f(0) = [-1, 1]$.

[1] La definición tiene sentido para funciones más generales y algunas propiedades se mantienen. Sin embargo, en este curso consideraremos sólo funciones convexas.

Ejemplo 2.4.2. Algunos casos patológicos:

1. Si $f(x_0) = -\infty$ entonces $\partial f(x_0) = Y$.
2. Si $f(x_0) = +\infty$ entonces
$$\partial f(x_0) = \begin{cases} \emptyset & \text{si } \operatorname{dom}(f) \neq \emptyset, \\ Y & \text{si } f \equiv +\infty. \end{cases}$$

Proposición 2.4.1. *Sean $f\colon X \to \mathbf{R} \cup \{+\infty\}$ propia y $x \in \operatorname{dom}(f)$. Las siguientes afirmaciones son equivlentes:*

(i) $y \in \partial f(x)$.

(ii) $f(x) + f^*(y) \leq \langle x, y\rangle$.

(iii) $f(x) + f^*(y) = \langle x, y\rangle$.

Más aun, si $f \in \Gamma_0(X)$ entonces se cumple la fórmula de reciprocidad de Legendre*:*

$$y \in \partial f(x) \quad \textit{si, y sólo si,} \quad x \in \partial f^*(y).$$

Demostración:

- $(i) \Rightarrow (ii)$ Dado $z \in X$ se tiene que $f(x) + \langle z - x, y\rangle \leq f(z)$ lo que equivale a $f(x) + \langle z, y\rangle - f(z) \leq \langle x, y\rangle$ de donde concluimos que $f(x) + f^*(y) \leq \langle y, x\rangle$.
- $(ii) \Rightarrow (iii)$ Evidente de la Desigualdad de Young-Fenchel.
- $(iii) \Rightarrow (i)$ Lo hicimos cuando supusimos que el supremo del cálculo de f^{**} se alcanzaba en y.

Finalmente, la condición $x \in \partial f^*(y)$ equivale a $f^*(y) + f^{**}(x) = \langle x, y\rangle$. Como $f \in \Gamma_0(X)$ se tiene que $f(x) = f^{**}(x)$ lo que concluye la demostración. □

Sean $\Omega \subseteq \mathbf{R}^n$ un abierto convexo no vacío y $f\colon \Omega \to \mathbf{R}$ una función convexa tal que $\nabla f\colon \Omega \to U \subseteq \mathbf{R}^n$ es un homeomorfismo. Supongamos que existe un par (x, y) tal que $y \in U$ y $x \in \partial f^*(y)$. De la caracterización anterior, x maximiza la función cóncava diferenciable $z \mapsto \langle z, y\rangle - f(z)$ y, en consecuencia, $y - \nabla f(x) = 0$ de donde concluimos que $x = (\nabla f)^{-1}(y) \in \Omega$ y definiendo la *transformada de Legendre* por

$$g(y) := \langle (\nabla f)^{-1}(y), y\rangle - f((\nabla f)^{-1}(y)),$$

se tiene que
$$f^*(y) = g(y).$$

Bajo hipótesis apropiadas de diferenciabilidad, es posible verificar sin la hipótesis de convexidad que $y = \nabla f(x)$ equivale a $x = \nabla g(y)$. Si además suponemos convexidad, hemos probado que la transformada de Legendre coincide con la conjugada de Fenchel.

Proposición 2.4.2. *Si $f\colon X \to \mathbf{R} \cup \{+\infty\}$ es propia entonces $\partial f(x)$ es un convexo cerrado de Y (eventualmente vacío).*

Demostración: Tenemos que

$$\partial f(x) = \{y \in Y \mid f(x) + f^*(y) \leq \langle x, y \rangle\} = \Gamma_{-f(x)}(f^* - \langle x, \cdot \rangle).$$

Aquí Γ denota el conjunto de subnivel inferior. Como $f^* - \langle x, \cdot \rangle$ es convexa y $\sigma(Y, X)$-s.c.i. se sigue el resultado. □

En lo que sigue, $(X, \|\cdot\|)$ es un espacio de Banach en dualidad con X^* y denotaremos indistintamente $x^*(x)$, $\langle x^*, x \rangle$ o $\langle x, x^* \rangle$.

Teorema 2.4.1. *Sean $f\colon X \to \overline{\mathbf{R}}$ convexa y $x_0 \in X$ tales que $f(x_0) \in \mathbf{R}$ y f es continua en x_0. Entonces $\partial f(x_0) \neq \emptyset$ y es compacto para la topología débil-* $\sigma(X^*, X)$.*

Demostración: Como f es convexa y continua en x_0, epi(f) es un convexo de interior no vacío. En particular, $(x_0, f(x_0))$ no pertenece al convexo int(epi(f)) y, por el Teorema de Hahn-Banach, podemos separar (no estrictamente) tal punto de tal convexo mediante un hiperplano cerrado. Es fácil ver que la "pendiente"de este hiperplano resulta ser un subgradiente de f en x_0. En particular f no puede tomar el valor $-\infty$ y como dom$(f) \neq \emptyset$ se tiene que además f es propia.
Como $\partial f(x_0)$ es convexo cerrado en X^*, para la compacidad basta verificar que es acotado para $\|\cdot\|_*$. Por ser f localmente Lipschitz en x_0, existen $\varepsilon > 0$ y $L > 0$ tales que

$$\|x - x_0\| < \varepsilon \Rightarrow |f(x) - f(x_0)| \leq L\|x - x_0\|.$$

Luego, si $x^* \in \partial f(x_0)$ tenemos que para todo $x \in X$ con $\|x - x_0\| < \varepsilon$ se tiene que

$$f(x_0) + \langle x^*, x - x_0 \rangle \leq f(x) \leq f(x_0) + L\|x - x_0\|$$

de donde concluimos que $\forall x \in B(x_0, \varepsilon)$, $\langle x^*, x - x_0 \rangle \leq L\|x - x_0\|$ y que, por lo tanto, $\|x^*\|_* \leq L$. □

Teorema 2.4.2 (Regla de Fermat I)**.** *Sea $f\colon X \to \mathbf{R} \cup \{+\infty\}$ una función convexa y propia. Consideremos*

$$(\mathcal{P}) \quad \inf_X f.$$

Entonces, x_0 es una solución de $(\mathcal{P})$ (es decir, $x_0 \in \arg\min_X f$) si, y sólo si, $0 \in \partial f(x_0)$. Si además $f \in \Gamma_0(X)$, entonces

$$\arg\min_X f = \partial f^*(0)$$

el cual es un convexo cerrado y acotado si f^* es finita y continua en 0.

Demostración: La condición $f(x_0) + f^*(0) = 0$ equivale a $\inf_X f = -f^*(0) = f(x_0)$. □

Proposición 2.4.3. *Sea $f\colon X \to \mathbf{R} \cup \{+\infty\}$ convexa y propia. Sean $x_0 \in$ dom(f) y $d \in X$. Entonces la derivada direccional en x_0 según d está definida en $\overline{\mathbf{R}}$ mediante*

$$f'(x_0; d) = \lim_{t \to 0+} \frac{f(x_0 + td) - f(x_0)}{t} = \inf_{t > 0} \frac{f(x_0 + td) - f(x_0)}{t}.$$

Demostración: Sea $q(t) = \frac{f(x_0+td)-f(x_0)}{t}$. Consideremos $0 < t \leq s$ y veamos que $q(t) \leq q(s)$. Notemos que

$$x_0 + td = \frac{t}{s}(x_0 + sd) + \left(1 - \frac{t}{s}\right) x_0,$$

de modo que por convexidad

$$f(x_0 + td) \leq \frac{t}{s} f(x_0 + sd) + \left(1 - \frac{t}{s}\right) f(x_0)$$

y se tiene que

$$\frac{f(x_0 + td) - f(x_0)}{t} \leq \frac{f(x_0 + sd) - f(x_0)}{s}.$$

□

Proposición 2.4.4. *Bajo las condiciones de la proposición anterior, $f'(x_0;\cdot)$ es una función sublineal y se tiene que*

$$\partial f(x_0) = \{x^* \in X^* \mid \forall d \in X,\ \langle x^*, d\rangle \leq f'(x_0; d)\} = \partial[f'(x_0;\cdot)](0).$$

Demostración: Verifiquemos la sublinealidad:

(i) $f'(x_0; 0) = 0$.

(ii) Sea $\lambda > 0$, luego

$$f'(x_0; \lambda d) = \lim_{t\to 0+} \frac{f(x_0 + t\lambda d) - f(x_0)}{t} = \lambda \lim_{t\to 0+} \frac{f(x_0 + (t\lambda)d) - f(x_0)}{\lambda t} = \lambda f'(x_0; d).$$

(iii) Dados $d_1, d_2 \in X$ verifiquemos que $f'(x_0; d_1 + d_2) \leq f'(x_0; d_1) + f'(x_0; d_2)$. Para ello, basta probar la convexidad de $f'(x_0;\cdot)$ pues $f'(x_0; d_1 + d_2) = 2f'(x_0; \frac{1}{2}d_1 + \frac{1}{2}d_2)$. Sea $\alpha \in]0, 1[$, luego

$$\begin{aligned} f'(x_0; \alpha d_1 + (1-\alpha)d_2) &= \lim_{t\to 0+} \frac{f(x_0 + t\alpha d_1 + t(1-\alpha)d_2) - f(x_0)}{t} \\ &\leq \inf_{t>0} \frac{\alpha}{t}(f(x_0 + td_1) - f(x_0)) + \frac{1-\alpha}{t}(f(x_0 + td_2) - f(x_0)) \\ &= \alpha f'(x_0; d_1) + (1-\alpha) f'(x_0; d_2). \end{aligned}$$

Finalmente,

$$\begin{aligned} x^* \in \partial f(x_0) &\Leftrightarrow f(x_0) + \langle x^*, y - x_0\rangle \leq f(y),\ \forall y \in X \\ &\Leftrightarrow f(x_0) + \langle x^*, td\rangle \leq f(x_0 + td),\ \forall d \in X,\ \forall t > 0 \\ &\Leftrightarrow \langle x^*, d\rangle \leq f'(x_0; d),\ \forall d \in X. \end{aligned}$$

□

Proposición 2.4.5. *Sean $(X, \|\cdot\|)$ un e.v.n. y $f \in \Gamma_0(X)$. Tomemos $x_0 \in X$ y supongamos que f es finita y continua en x_0. Entonces*

$$f'(x_0; d) = \sup_{x^* \in \partial f(x_0)} \langle x^*, d\rangle = \sigma_{\partial f(x_0)}(d).$$

Demostración: Ya vimos que $f'(x_0; d) \geq \sup_{x^* \in \partial f(x_0)} \langle x^*, d \rangle$. Para probar la igualdad observemos que $f'(x_0; \cdot)$ es convexa y continua. En efecto, tenemos que $\partial f(x_0) \neq \emptyset$ y en consecuencia $f'(x_0; d) \in \mathbf{R}$, $\forall d \in X$. Luego basta probar que $f'(x_0; \cdot)$ es continua en 0, lo que resulta de

$$f'(x_0; d) \leq f(x_0 + d) - f(x_0) \leq M$$

para algún $M > 0$ y para todo d en alguna bola centrada en 0. De lo anterior, $f'(x_0; \cdot) = [f'(x_0; \cdot)]^{**}$. Además

$$\begin{aligned} [f'(x_0; \cdot)]^*(x^*) &= \sup_{d \in X} \{\langle x^*, d \rangle - f'(x_0; d)\} \\ &= \begin{cases} 0 & \text{si } x^* \in \partial f(x_0), \\ +\infty & \text{si no} \end{cases} \\ &= \delta_{\partial f(x_0)}(x^*) \end{aligned}$$

Finalmente,

$$f'(x_0; d) = \sup_{x^* \in X^*} \{\langle x^*, d \rangle - [f'(x_0; \cdot)]^*(x^*)\} = \delta^*_{\partial f(x_0)}(d) = \sigma_{\partial f(x_0)}(d).$$

□

Corolario 2.4.1. *Sean $(X, \|\cdot\|)$ un e.v.n. y $f \in \Gamma_0(X)$. Si $x_0 \in \mathrm{dom}(f)$ es tal que f es continua en x_0 y $\partial f(x_0) = \{x_0^*\}$ entonces f es Gâteaux-diferenciable en x_0 y $\nabla f(x_0) = x_0^*$.*

Demostración: Se tiene que

$$f'(x_0; d) = \sup_{x^* \in \partial f(x_0)} \langle x^*, d \rangle = \langle x_0^*, d \rangle$$

es una función lineal y continua. □

Corolario 2.4.2. *Si $f \in \Gamma_0(\mathbf{R}^n)$ y $x_0 \in \mathrm{dom}(f)$ es tal que f es continua en x_0 y $\partial f(x_0) = \{x^*\}$, entonces f es Fréchet-diferenciable en x_0.*

Demostración: Basta probar que

$$\limsup_{y \to x_0} \frac{f(y) - f(x_0) - \langle x^*, y - x_0 \rangle}{\|y - x_0\|} \leq 0.$$

Sea $L \in \overline{\mathbf{R}}$ este límite superior y tomemos $y_k \to x_0$, tal que

$$\lim_{k \to \infty} \frac{f(y_k) - f(x_0) - \langle x^*, y_k - x_0 \rangle}{\|y_k - x_0\|} = L.$$

Por compacidad podemos suponer que $d_k := \frac{y_k - x_0}{\|y_k - x_0\|} \to d$ con $\|d\| = 1$. Como f es localmente Lipschitz en torno a x_0, tenemos que

$$\frac{f(x_0 + \|y_k - x_0\| d_k) - f(x_0 + \|y_k - x_0\| d)}{\|y_k - x_0\|} \to 0$$

y en consecuencia

$$L = \lim_{k \to \infty} \frac{f(x_0 + \|y_k - x_0\| d) - f(x_0) - \langle x^*, y_k - x_0 \rangle}{\|y_k - x_0\|} = f'(x_0; d) - \langle x^*, d \rangle = 0.$$

□

Observación 2.4.1. Notemos que si $f \in \Gamma_0(X)$ es Gâteaux-diferenciable en $x_0 \in \mathrm{dom}(f)$, entonces $f'(x_0; d) = \langle \nabla f(x_0), d \rangle$. Así,

$$\partial f(x_0) = \{x^* \in X^* \mid \langle x^*, d\rangle \leq \langle \nabla f(x_0), d\rangle,\ \forall d \in X\} = \{\nabla f(x_0)\}.$$

Proposición 2.4.6. *Sea $U \subseteq X$ un abierto convexo y $f\colon U \to \mathbf{R}$ una función Gâteaux-diferenciable (en U). Son equivalentes:*

(i) f *es convexa sobre* U.

(ii) $\forall x, y \in U,\ f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle$.

(iii) $\forall x, y \in U,\ \langle \nabla f(x) - \nabla f(y), x - y\rangle \geq 0$.

Demostración: Propuesto. □

Observación 2.4.2. Notemos que, en general, si $f\colon X \to \mathbf{R} \cup \{+\infty\}$ es propia, entonces $\partial f\colon X \to 2^X$ tiene la siguiente propiedad: si x, y, x^*, y^* son tales que $x^* \in \partial f(x)$ e $y^* \in \partial f(y)$, entonces

$$\langle x^* - y^*, x - y\rangle \geq 0.$$

Las funciones del tipo $A\colon X \to 2^X$ suelen llamarse *multiaplicaciones* y cuando una multiaplicación tiene la propiedad anterior, decimos que es *monótona*.

Para terminar la sección estudiaremos algunas propiedades elementales del cálculo subdiferencial.

Ejercicio 2.4.1.

- Sean $f\colon X \to \overline{\mathbf{R}}$ y $\lambda > 0$. Entonces $\forall x \in X,\ \partial(\lambda f)(x) = \lambda \partial f(x)$.
- Sean $f_1, f_2\colon X \to \overline{\mathbf{R}}$. Entonces $\forall x \in X,\ \partial f_1(x) + \partial f_2(x) \subseteq \partial(f_1 + f_2)(x)$.

Proposición 2.4.7 (Teorema de Moreau-Rockafellar)**.** *Si $f_1, f_2 \in \Gamma_0(X)$ y f_1 es continua en algún $x_0 \in \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$, entonces*

$$\forall x \in X,\ \partial f_1(x) + \partial f_2(x) = \partial(f_1 + f_2)(x).$$

Demostración: En virtud del ejercicio anterior, probaremos sólo la contención ($\supseteq$). Sean x, x^* tales que $x^* \in \partial(f_1 + f_2)(x)$. Tenemos que

$$\forall y \in X,\ f_1(y) + f_2(y) \geq \langle x^*, y - x\rangle + f_1(x) + f_2(x).$$

Introduzcamos los siguientes conjuntos convexos

$$C_1 := \{(y, \lambda) \in X \times \mathbf{R} \mid f_1(y) - \langle x^*, y - x\rangle \leq \lambda\},$$
$$C_2 := \{(y, \lambda) \in X \times \mathbf{R} \mid f_2(x) - f_2(y) \geq \lambda\}.$$

Además, $(y, \lambda) \in C_1 \cap C_2$ equivale a $f_1(y) + f_2(y) = \langle x^*, y - x\rangle + f_1(x) + f_2(x)$. Pero $C_1 = \mathrm{epi}(g)$ con $g = f_1 - \langle x^*, \cdot\rangle \in \Gamma_0(X)$ continua en x_0. Luego, $\mathrm{int}(C_1) = \{(y, \lambda) \in X \times \mathbf{R} \mid g(y) < \lambda\} \neq \emptyset$ y además $\mathrm{int}(C_1) \cap C_2 = \emptyset$. Podemos separar $\mathrm{int}(C_1)$ de C_2 mediante un hiperplano cerrado que además es no vertical (verificar), y en consecuencia es el grafo de una función lineal afín $\ell(y) = \langle -x_2^*, y\rangle + \alpha$ para $x_2^* \in X^*$ y $\alpha \in \mathbf{R}$. Podemos escribir la separación como

$$\forall y \in X,\ f_2(x) - f_2(y) \leq \langle -x_2^*, y\rangle + \alpha \leq f_1(y) - \langle x^*, y - x\rangle - f(x).$$

Tomando $y = x$ deducimos que $\alpha = \langle x_2^*, x\rangle$ y en consecuencia

$$\forall y \in X: \; f_1(x) + \langle -x_2^* + x^*, y - x\rangle \leq f_1(y)$$

lo que implica que $-x_2^* + x^* \in \partial f_1(x)$ y análogamente

$$\forall y \in X: \; f_2(x) + \langle x_2^*, y - x\rangle \leq f_2(y)$$

y concluimos que $x_2^* \in \partial f_2(x)$. Definiendo $x_1^* = x^* - x_2^*$ se tiene que $x_1^* \in \partial f_1(x)$ y $x_1^* + x_2^* = x^*$ lo que concluye la demostración. □

Consideremos una función $f \in \Gamma_0(X)$ y un convexo cerrado no-vacío C. Supongamos que existe $x_0 \in \mathrm{int}(C)$ tal que f es finita en x_0. Consideremos el problema de optimización

$$(\mathcal{P}) \qquad \inf_{x \in C} f(x),$$

que equivale a

$$\inf_{x \in X} \{f(x) + \delta_C(x)\}.$$

De la Regla de Fermat se sigue que x^* es solución de $(\mathcal{P})$ si y sólo si $0 \in \partial(f + \delta_C)(x^*)$ lo que, de acuerdo con el Teorema de Moreau-Rockafellar, equivale a

$$0 \in \partial f(x^*) + \partial \delta_C(x^*).$$

y definiendo el *Cono Normal a C en $x \in X$* por $N_C(x) := \partial \delta_C(x)$ podemos reescribir la condición de optimalidad como

$$0 \in \partial f(x^*) + N_C(x^*).$$

Además, dado $x \in C$, se tiene que

$$\begin{aligned} N_C(x) &= \{y \in Y \mid \forall u \in X, \; \delta_C(x) + \langle u - x, y\rangle \leq \delta_C(u)\} \\ &= \{y \in Y \mid \forall u \in C, \; \langle u - x, y\rangle \leq 0\}. \end{aligned}$$

Luego, $x^* \in C$ es solución de $(\mathcal{P})$ si, y sólo si, existe $\psi \in \partial f(x^*)$ tal que

$$\forall u \in C, \; \langle u - x^*, \psi\rangle \geq 0,$$

y hemos establecido el siguiente:

Teorema 2.4.3 (Regla de Fermat II). *Sea $f \in \Gamma_0(X)$ y C un convexo cerrado no vacío y supongamos que existe $x_0 \in \mathrm{int}(C)$ tal que f es finita en x_0. Entonces, $x^* \in C$ es solución de*

$$\inf_C f$$

si, y sólo si, $0 \in \partial f(x^*) + N_C(x^*)$ lo que equivale a que exista $p \in \partial f(x^*)$ tal que

$$(\mathcal{C}) \quad \forall u \in C, \quad \langle u - x^*, p\rangle \geq 0.$$

Observación 2.4.3. La condición de optimalidad anterior es llamada *Desigualdad Variacional de C y ∂f*. Sin embargo, el concepto de desigualdad variacional es un poco más general y aparece en diversos contextos.

Ejercicio 2.4.2. Obtenga la condición de optimalidad para el problema

$$\inf_{x\in X}\{f(x)+\delta_C(x)\}$$

cuando se satisfacen las hipótesis de la proposisción anterior y f es diferenciable.

Proposición 2.4.8. *Sean X, Y dos espacios de Banach en dualidad con X^*, Y^* respectivamente. Si $A\colon X\to Y$ es una función lineal continua y $f\in\Gamma_0(Y)$ entonces $f\circ A\in\Gamma_0(X)$ y si f es continua en algún $y_0\in\operatorname{dom}(f)$ entonces*

$$\partial(f\circ A)(x)=A^*\partial f(Ax),$$

donde A^* denota el adjunto de A.

Demostración: Sea $y^*\in\partial f(Ax)$. Por definición,

$$\forall y\in Y,\ f(Ax)+\langle y^*, y-Ax\rangle\leq f(y)$$

y, en particular,

$$\forall x\in X,\ f(Ax)+\langle y^*, A(z-x)\rangle\leq f(Az).$$

Esto equivale a

$$\forall z\in X,\ (f\circ A)(x)+\langle A^*y^*, z-x\rangle\leq (f\circ A)(z)$$

y se tiene que $A^*y^*\in\partial(f\circ A)(x)$. Así, $A^*\partial f(Ax)\subseteq\partial(f\circ A)(x)$.
Recíprocamente, sea $x^*\in\partial(f\circ A)(x)$, luego

$$\forall z\in X,\ f(Ax)+\langle x^*, z-x\rangle\leq f(Az).$$

Definimos

$$S:=\{(Az, f(Ax)+\langle x^*, z-x\rangle\in Y\times\mathbf{R}\mid z\in X\}$$

que resulta ser un hiperplano cerrado de $Y\times\mathbf{R}$. Notemos que $(y,\lambda)\in S\cap\operatorname{epi}(f)$ si, y sólo si,

$$\exists z\in X:\ Az=y,\ f(Ax)+\langle x^*, z-x\rangle=f(Az)=\lambda.$$

Como f es convexa y continua en $y_0\in Y$, tenemos que $\operatorname{int}(\operatorname{epi}(f))\neq\emptyset$ y obviamente $S\cap\operatorname{int}(\operatorname{epi}(f))=\emptyset$. En consecuencia, podemos separar S de $\operatorname{int}(\operatorname{epi}(f))$ mediante un hiperplano cerrado que resulta ser no vertical y, por lo tanto, el grafo de una función lineal afín $\ell(y)=\langle y^*, y\rangle+\alpha$ con $y^*\in Y$ y $\alpha\in\mathbf{R}$. La separación implica que

$$\forall z\in X,\ f(Ax)+\langle x^*, z-x\rangle\leq\langle y^*, Az\rangle+\alpha\leq f(Az)$$

de donde $\alpha=f(Ax)-\langle y^*, Ax\rangle$ y así $\ell(y)=\langle y^*, y-Ax\rangle+f(Ax)$. Por lo tanto

$$\forall y\in Y,\ f(Ax)+\langle y^*, y-Ax\rangle\leq f(y)$$

y

$$\forall z\in X,\ \langle x^*, z-x\rangle\leq\langle A^*y^*, z-x\rangle$$

condiciones que implican que $y^*\in\partial f(Ax)$ y $x^*=A^*y^*\in A^*\partial f(Ax)$. Luego $\partial(f\circ A)(x)\subseteq A^*\partial f(Ax)$.

$\square$

2.5. Problemas

Problema 5. Sea X un e.v.t. Dado $A \subseteq X$ definimos la *Envoltura Convexa* de A como

$$\operatorname{co}(A) = \bigcap \{\ C \mid C \text{ es un convexo}, A \subseteq C\}$$

y la *Envoltura Convexa Cerrada* de A como $\overline{\operatorname{co}} A = \operatorname{cl}(\mathrm{A})$.

(a) Pruebe que

$$\operatorname{co}(A) = \left\{ \sum_{i=1}^{n} \lambda_i v_i \;\middle|\; n \in \mathbf{N},\ v_i \in A,\ \lambda_i \geq 0,\ \sum_{i=1}^{n} \lambda_i = 1 \right\}$$

(b) (*Teorema de Carathéodory*) Demuestre que si X es de dimensión finita igual a n, entonces

$$\operatorname{co}(A) = \left\{ \sum_{i=1}^{n+1} \lambda_i v_i \;\middle|\; v_i \in A,\ \lambda_i \geq 0,\ \sum_{i=1}^{n+1} \lambda_i = 1 \right\}.$$

(c) Pruebe que si $C \subseteq X$ es un convexo, entonces $\operatorname{cl}(C)$ también lo es. Deduzca que

$$\overline{\operatorname{co}}(A) = \bigcap \{\ C \mid C \text{ es un convexo cerrado, } A \subseteq C\}.$$

(d) Pruebe que si A es abierto, entonces $\operatorname{co}(A)$ también lo es. Muestre que si $C \subseteq X$ es convexo, entonces $\operatorname{int}(C)$ también lo es.

(e) Muestre que si $A_1, \ldots, A_n \subseteq X$ son convexos y compactos, entonces $\operatorname{co}(A_1 \cup \cdots \cup A_n)$ es compacto.

(f) Sea $(V, \|\cdot\|)$ un e.v.n. y supongamos que A es totalmente acotado. Pruebe que $\operatorname{co}(A)$ es totalmente acotado. Deduzca que si $(V, \|\cdot\|)$ es de Banach y $K \subseteq V$ es compacto, entonces para todo $A \subseteq K$, $\bar{\operatorname{co}} A$ es compacto.

Problema 6. En lo que sigue, X es un e.v.n. y todas las funciones están en $\Gamma_0(X)$. Siguiendo la Definición **??**, denotamos por f_∞ la función de recesión de f.

1. Para todo $d \in X$ y cualquier $x \in \operatorname{dom}(f)$, se tiene que

$$f_\infty(d) = \lim_{t \to \infty} \frac{f(x+td) - f(x)}{t} = \sup_{t>0} \frac{f(x+td) - f(x)}{t}.$$

2. Pruebe que $\left(\sup\limits_{i \in I} f_i \right)_\infty = \sup\limits_{i \in I} \{(f_i)_\infty\}$.

3. Consideremos un subconjunto convexo, cerrado y no-vacío S de X. Definimos $C = \{\ s \in S \mid f_i(x) \leq 0\ \forall i \in I\ \}$. Demuestre que $C_\infty = \{\ d \in S_\infty \mid (f_i)_\infty(d) \leq 0\ \forall i \in I\ \}$.

4. Para todo $\lambda > \inf f$, se tiene que $[\ \Gamma_\lambda(f)\ ]_\infty = \{\ d \in X \mid f_\infty(d) \leq 0\ \}$.

5. Si f es inf-compacta, entonces $f_\infty(d) > 0$ para todo $d \neq 0$ (compare con el Teorema 1.2.2).

Problema 7. Sea $(V, \|\cdot\|)$ un e.v.n. en dualidad con V^*, y $f \in \Gamma_0(V)$. Pruebe que las siguientes afirmaciones son equivalentes.

(i) f es coerciva;

(ii) Existe $\alpha > 0$ y $\beta \in \mathbf{R}$ tales que para todo $v \in V$, $f(v) \geq \alpha\|v\| + \beta$;

(iii) $0 \in \mathrm{int}(\mathrm{dom}(f^*))$;

(iv) $\liminf\limits_{\|x\|\to+\infty} \frac{f(x)}{\|x\|} > 0$.

La equivalencia entre (i) y (iii) se conoce como *Teorema de Moreau*.

Problema 8. Sea $(X, Y, \langle\cdot,\cdot\rangle)$ una dualidad entre e.v.t.l.c. Dadas dos funciones convexas $f, g\colon X \to \overline{\mathbf{R}}$, se define la *inf-convolución* de f y g mediante

$$(f_* g)(x) := \inf\{f(x_1) + g(x_2) \mid x_1 + x_2 = x\}.$$

(a) Pruebe que $f_* g$ es convexa, con $\mathrm{dom}(f_* g) = \mathrm{dom}(f) + \mathrm{dom}(g)$.

(b) Sea $y \in Y$, muestre que $(f_* g)^*(y) = f^*(y) + g^*(y)$.

(c) Pruebe que si $\bar{x}_1 \in \mathrm{dom}(f)$ y $\bar{x}_2 \in \mathrm{dom}(g)$ son tales que $(f_* g)(\bar{x}_1 + \bar{x}_2) = f(\bar{x}_1) + g(\bar{x}_2)$, entonces $\partial(f_* g)(\bar{x}_1 + \bar{x}_2) = \partial f(\bar{x}_1) \cap \partial g(\bar{x}_2)$.

(d) (*Efecto Regularizante*) Suponga que $\bar{x}_i$ son los considerados en la parte anterior. Asumiendo que $f_* g$ es subdiferenciable en $\bar{x} = \bar{x}_1 + \bar{x}_2$, muestre que $f_* g$ es Gâteaux-diferenciable en $\bar{x}$ si g lo es en $\bar{x}_2$ con
$$\nabla(f_* g)(\bar{x}) = \nabla g(\bar{x}_2).$$
Muestre que si además $(X, \|\cdot\|)$ es un e.v.n. y g es Fréchet-diferenciable en $\bar{x}_2$, entonces $f_* g$ es Fréchet-diferenciable en $\bar{x}$.

Problema 9. Sea $(H, \langle\cdot,\cdot\rangle)$ un espacio de Hilbert en dualidad con H^* (que identificamos con H).

(a) Sea $f(x) = \frac{1}{2}\|x\|^2$. Verifique que $f^*(x^*) = \frac{1}{2}\|x^*\|^2$ de modo que $f = f^*$. Más aún, demuestre que f es la única función en $\Gamma_0(H)$ con esta propiedad.

(b) Sean $f \in \Gamma_0(H)$ y $S \subseteq H$ un s.e.v. cerrado. Muestre que si existe $x_0 \in S$ tal que $f(x_0) < +\infty$, entonces $f + \delta_S \in \Gamma_0(H)$ y se tiene

$$(f + \delta_S)^* = (f \circ P_S)^* \circ P_S,$$

donde $P_S\colon H \to H$ es la proyección ortogonal sobre S.

(c) Supongamos que $f(x) = \frac{1}{2}\langle Ax, x\rangle$, con $A\colon H \to H$ un operador lineal continuo autoadjunto y semi-definido positivo. Pruebe que

$$(f + \delta_S)^*(x^*) = \begin{cases} \frac{1}{2}\langle x^*, x\rangle & \text{si } x^* \in \mathrm{Im}A + S^\perp, \text{ con } (P_S \circ A \circ P_S)(x) = x^*, \\ +\infty & \text{si no.} \end{cases}$$

Problema 10. Sea X un espacio de Banach y sea f una función localmente Lipschitz en $x \in X$. Definimos la *Derivada Direccional Generalizada* de f en x en la dirección v mediante

$$f^0(x;v) = \lim \sup_{y \to x, t \to 0+} \frac{f(y+tv) - f(y)}{t}.$$

(a) Muestre que $v \mapsto f^0(x;v)$ es finita, positivamente homogénea, y subaditiva en X. Pruebe además que
$$|f^0(x;v)| \leq K\|v\|.$$

(b) Muestre que $f^0(x;v)$ es s.c.s como función de (x,v) y que, como función sólo de v, es Lipschitz-continua de constante K.

(c) Definimos el conjunto de *Subgradientes Generalizados* mediante
$$\hat{\partial} f(x) = \{y \in X^* \mid f^0(x;v) \geq \langle y, v \rangle, \quad \forall v \in X\}.$$
Muestre que $\hat{\partial} f(x)$ es no-vacío y convexo. Pruebe que
$$\forall y \in \hat{\partial} f(x), \quad \|y\|_* \leq K.$$

(d) Pruebe que si f es además Gâteaux-diferenciable, entonces $\nabla f(x) \in \hat{\partial} f(x)$. Muestre que $\hat{\partial} f(x)$ puede contener otros puntos además de $\nabla f(x)$.

(e) Pruebe que si f es convexa, entonces $\hat{\partial} f(x)$ coincide con el subdiferencial estudiado a lo largo del capítulo (que se conoce usualmente como el subdiferencial de Análisis Convexo.)

Capítulo 3

Dualidad en Optimización Convexa

3.1. Problemas Perturbados

Consideremos el problema de minimización convexa

$$(\mathcal{P}) \quad \alpha := \inf_X f$$

donde $f \in \Gamma_0(X)$. Diremos que $(\mathcal{P})$ es el *Problema Primal* y que una función $\varphi \in \Gamma_0(X \times Y)$, con X, Y espacios de Banach en dualidad con X^*, Y^* respectivamente, es una *función de perturbación* para $(\mathcal{P})$ si $\varphi(x, 0) = f(x)$, $\forall x \in X$. A φ le asociamos la *función marginal* o *función valor* definida por

$$v(y) = \inf_{x \in X} \varphi(x, y).$$

Notemos que $v(0) = \alpha$.

Ejemplo 3.1.1 (Programación Lineal I)**.** Sea

$$(\mathcal{P}) \quad \min\{c^T x \mid Ax \leq b\},$$

con $c \in \mathbf{R}^n$, $b \in \mathbf{R}^m$ y $A \in \mathbf{M}_{n \times n}(\mathbf{R})$. Tomando $X = \mathbf{R}^n$ y

$$f(x) = c^T x + \delta_{\mathbf{R}^m_-}(Ax - b) = \begin{cases} c^T x & \text{si } Ax \leq b, \\ +\infty & \text{si no} \end{cases}$$

se tiene que $f \in \Gamma_0(X)$. Introducimos la función de perturbación $\varphi \colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R} \cup \{+\infty\}$ dada por

$$\varphi(x, y) = c^T x + \delta_{\mathbf{R}^m_-}(Ax - b + y).$$

Es fácil ver que $\varphi \in \Gamma_0(\mathbf{R}^n \times \mathbf{R}^m)$. Observemos que $\varphi^*\colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R} \cup \{+\infty\}$ se calcula mediante

$$\begin{aligned}
\varphi^*(x^*, y^*) &= \sup\left\{(x^* - c)^T x + y^{*T} y \mid x \in \mathbf{R}^n,\ Ax + b - y \le 0\right\} \\
&= \begin{cases} +\infty & \text{si } \exists i \in \{1, ..., m\} \text{ tal que } y_i^* < 0, \\ \sup_{x\in\mathbf{R}^n} (x^* - c)^T x + y^{*T}(b - Ax) & \text{si } y^* \ge 0 \end{cases} \\
&= \begin{cases} +\infty & \text{si } \exists i \in \{1, ..., m\} \text{ tal que } y_i^* < 0, \\ \sup_{x\in\mathbf{R}^n} \{(x^* - c - A^T y^*)^T x\} + y^{*T} b & \text{si } y^* \ge 0 \end{cases} \\
&= \begin{cases} b^T y^* & \text{si } x^* = c + A^T y^*, y^* \ge 0, \\ +\infty & \text{si no.} \end{cases}
\end{aligned}$$

En particular, el problema de optimización

$$(\mathcal{D}) \quad \beta := \inf_{y^* \in \mathbf{R}^m} \varphi^*(0, y^*) = \min_{A^T y^* + c = 0,\ y^* \ge 0} b^T y^*,$$

corresponde al Dual clásico de $(\mathcal{P})$ en la teoría de programación lineal.

Volvamos al caso general. Por analogía, llamamos *Problema Dual* de $(\mathcal{P})$ relativo a la perturbación $\varphi\colon X \times Y \to \mathbf{R} \cup \{+\infty\}$ al problema de minimización

$$(\mathcal{D}) \quad \beta := \inf_{y^* \in Y^*} \varphi^*(0, y^*)$$

el cual tiene asociado de manera natural la función de perturbación $\varphi^* \in \Gamma_0(X^* \times Y^*)$ y la correspondiente función valor

$$w(x^*) := \inf_{y^* \in Y^*} \varphi^*(x^*, y^*).$$

Observemos que si repetimos el procedimiento anterior obtenemos el *Problema Bidual*

$$(\mathcal{DD}) \inf_{x \in X} \varphi^{**}(x^*, 0)$$

el cual coincide con el primal $(\mathcal{P})$ si suponemos que $\varphi \in \Gamma_0(X \times Y)$ pues en este caso $\varphi^{**} = \varphi$.

Ejercicio 3.1.1. Muestre que $v\colon Y \to \overline{\mathbf{R}}$ y $w\colon X^* \to \overline{\mathbf{R}}$ son funciones convexas.

Observación 3.1.1. Para lo anterior, basta que φ sea convexa. Por otra parte, si $\varphi \in \Gamma_0(X \times Y)$ no necesariamente se tiene que $v \in \Gamma_0(Y)$.

Ejercicio 3.1.2. Considere el *programa convexo*

$$(\mathcal{P}) \qquad \inf_{x \in \mathbf{R}^n} \{f(x) \mid g_i(x) \le 0, i = 1, \ldots, p\}$$

donde f y $(g_i)_i$ son funciones convexas de $\mathbf{R}^n$ en $\mathbf{R}$. Calcule el problema dual de $(\mathcal{P})$ perturbando el objetivo primal de manera análoga al ejemplo de Programación Lineal I.

Lema 3.1.1. *Para cada $y^* \in Y^*$ se tiene que $v^*(y^*) = \varphi^*(0, y^*)$. Análogamente, dado $x \in X$, se tiene que $w^*(x) = \varphi(x, 0)$.*

Demostración: Basta verificar la fórmula para v^*. Dado $y^* \in Y^*$ tenemos que

$$\begin{aligned} v^*(y^*) &= \sup_{y\in Y}\{\langle y, y^*\rangle - v(y)\} \\ &= \sup_{y\in Y}\{\langle y, y^*\rangle - \inf_{x\in X} \varphi(x,y)\} \\ &= \sup_{x\in X, y\in Y}\{\langle x, 0\rangle + \langle y, y^*\rangle - \varphi(x,y)\} \\ &= \varphi^*(0, y^*). \end{aligned}$$

□

Corolario 3.1.1. *Se tiene que* $-v^{**}(0) = \beta = \inf(\mathcal{D})$ *y en particular* $\inf(\mathcal{P}) + \inf(\mathcal{D}) \geq 0$.

Demostración: Observemos primero que

$$-v^{**}(0) = -\sup_{y^*\in Y^*} -v^*(y^*) = \inf_{y^*\in Y^*} v^*(y^*) = \inf_{y^*\in Y^*} \varphi^*(0, y^*) = \beta.$$

Como $v \geq v^{**}$ se sigue que $\alpha + \beta = v(0) - v^{**}(0) \geq 0$. □

Proposición 3.1.1. *Una condición necesaria y suficiente para que*

$$\inf(\mathcal{P}) + \inf(\mathcal{D}) = 0$$

es que v sea s.c.i. y finita en 0. En ese caso decimos que no hay salto de dualidad.

Demostración: Veamos la suficiencia. Sea $\overline{v}$ la regularizada s.c.i. de v. Tenemos que $v^{**} \leq \overline{v} \leq v$ con $\overline{v}$ convexa y s.c.i. Además, $\overline{v}(0) = v(0) \in \mathbf{R}$ de modo que $\overline{v}$ es propia, es decir $\overline{v} \in \Gamma_0(Y)$ y en consecuencia $\overline{v} = v^{**}$ y en particular $-\beta = v^{**}(0) = v(0) = \alpha \in \mathbf{R}$.
Mostremos ahora la necesidad. Como $\alpha + \beta = 0$ necesariamente α y β son finitos y en consecuencia $v(0) \in \mathbf{R}$. Además, $v^{**} \leq \overline{v} \leq v$ con v^{**} s.c.i. y tenemos que $v(0) = v^{**}(0) \leq \overline{v}(0) \leq v(0)$ de modo que v es s.c.i. en 0. □

Observación 3.1.2.

1. Como $\alpha + \beta = v(0) + w(0) = 0$ entonces $\alpha + \beta = 0$ si, y sólo si, w es s.c.i. y finita en 0 .

2. Recordemos que, de acuerdo con la Observación 2.3.2, si $f\colon Y \to \overline{\mathbf{R}}$ es convexa, entonces

$$\overline{f} = f^{**} \Leftrightarrow \overline{f} = -\infty \text{ o } f \text{ admite una minorante lineal afín continua.}$$

   En particular, si $f\colon Y \to \mathbf{R} \cup \{+\infty\}$ es convexa y acotada inferiormente, entonces $\overline{f} = f^{**}$. Si $v(0) = -\infty$ entonces $v^{**}(0) = v(0) = -\infty$, luego $\alpha + \beta = +\infty - \infty = +\infty$ y el salto de dualidad es infinito.

3. La equivalencia anterior sólo requiere que v sea convexa, lo cual es cierto si φ es convexa (aunque no esté en $\Gamma_0(X \times Y)$). Sin embargo, esta condición por sí sola no basta para la equivalencia cuando v se reemplaza por w (ver parte 1.).

Proposición 3.1.2. *Se tiene que* $S(\mathcal{D}) = \partial v^{**}(0)$*. En particular, una condición necesaria y suficiente para que*

$$\inf(\mathcal{P}) + \inf(\mathcal{D}) = 0,\ S(\mathcal{D}) \neq \emptyset$$

es

$$v(0) \in \mathbf{R}, \quad \partial v(0) \neq \emptyset.$$

Demostración: Notemos que

$$\begin{aligned} y^* \in S(\mathcal{D}) &\Leftrightarrow \forall z^* \in Y^*,\ \varphi^*(0, y^*) \leq \varphi^*(0, z^*) \\ &\Leftrightarrow \forall z^* \in Y^*,\ -v^*(y) \geq \langle 0, z^* \rangle - v^*(z^*) \\ &\Leftrightarrow -v^*(y^*) = v^{**}(0) \\ &\Leftrightarrow y^* \in \partial v^{**}(0). \end{aligned}$$

Probemos ahora la equivalencia. Para la necesidad, notemos que $v(0) \in \mathbf{R}$ y v s.c.i. en 0 implica que $v(0) = v^{**}(0)$. Tenemos que ℓ es una minorante lineal-afín continua de v si y sólo si lo es de v^{**} y, más aun, si $\ell(0) = v(0)$, entonces $\ell(0) = v^{**}(0)$ de modo que siempre se tiene que $\partial v(0) = \partial v^{**}(0)$. Si además $v(0) = v^{**}(0)$, se deduce que $\partial v(0) = \partial v^{**}(0) = S(\mathcal{D}) \neq \emptyset$.

Mostremos la suficiencia. Si $\partial v(0) \neq \emptyset$ y $v(0) \in \mathbf{R}$, entonces existe ℓ minorante lineal-afín continua de v tal que $\ell(0) = v(0)$. Pero $\ell \leq v^{**} \leq v$ de modo que $v^{**}(0) = v(0)$ y en consecuencia $\partial v(0) = \partial v^{**}(0) = S(\mathcal{D})$. Así, $S(\mathcal{D}) \neq \emptyset$ y además $v(0) = \overline{v}(0) = v^{**}(0)$ de modo que v es s.c.i. y por lo tanto $\inf(\mathcal{P}) + \inf(\mathcal{D}) = 0$. □

Corolario 3.1.2. *Si* $v(0) \in \mathbf{R}$ *y* $\partial v(0) \neq \emptyset$ *entonces* $S(\mathcal{D}) = \partial v(0) \neq \emptyset$ *y además* $\inf(\mathcal{P}) + \inf(\mathcal{D}) = 0$.
Del mismo modo, si $w(0) \in \mathbf{R}$ *y* $\partial w(0) \neq \emptyset$ *entonces* $S(\mathcal{P}) = \partial w(0) \neq \emptyset$ *y no hay salto de dualidad.*

Interpretación: Si $y^* \in S(\mathcal{D})$ entonces $v(y) \geq v(0) + \langle y^*, y \rangle$, que es una información de primer orden sobre la función marginal primal. Si $S(\mathcal{D}) = \{y^*\}$ entonces $\nabla v(0) = y^*$. En particular, si v es continua y finita en 0, entonces $\partial v(0) \neq \emptyset$. Un caso patológico se tiene cuando $v(0) = -\infty$. En tal caso, $\partial v(0) = Y^*$, pero $\beta = w(0) = +\infty$ de modo que el dual es infactible.

Teorema 3.1.1 (Teorema de Dualidad)**.** *Supongamos que* $\varphi \in \Gamma_0(X \times Y)$ *y que existe* $x_0 \in X$ *tal que* $\varphi(x_0, \cdot)$ *es finita y continua en 0 (i.e.* $\varphi(x_0, \cdot)$ *es acotada superiormente en una vecindad de 0). Entonces, o bien* $v(0) = -\infty$ *y el dual es infactible, o bien* $v(0) \in \mathbf{R}$*,* $\alpha + \beta = 0$ *y* $S(\mathcal{D}) = \partial v(0) \neq \emptyset$*, con* v *continua en 0. En particular, si* $\varphi(x_0, \cdot)$ *es continua en 0 respecto a* $\|\cdot\|$ *con* $(Y, \|\cdot\|)$ *un espacio de Banach, entonces* $S(\mathcal{D}) \neq \emptyset$ *es acotado.*

Demostración: Obviamente $v(\cdot) \leq \varphi(x_0, \cdot)$ implica que v es acotada en una vecindad de 0. Si además $v(0) \in \mathbf{R}$, entonces v es continua en 0 y luego $\partial v(0) \neq \emptyset$. □

Observación 3.1.3. Si X es e.v.t.l.c. se cumple la propiedad "$f\colon X \to \overline{\mathbf{R}}$ convexa y continua en $x_0 \in X$ si, y sólo si, f es acotada superiormente de manera uniforme en una vecindad de x_0."

Comentario: Notemos que $y^* \in \partial v(0)$ es equivalente a $0 \in \partial v^*(y^*)$ y esta condición es la que da la Regla de Fermat para el problema $\inf_{Y^*} v^*$, que no es otra cosa que el problema dual. Así, al asegurar que v es continua en 0 con respecto a $\|\cdot\|$ en Y tenemos que el problema dual admite solución y además $S(\mathcal{D}) = \partial v(0) \neq \emptyset$ es acotado.

Proposición 3.1.3 (Condición de Extremalidad). *Si $(\mathcal{P})$ y $(\mathcal{D})$ admiten soluciones y no hay salto de dualidad, entonces $\forall \bar{x} \in S(\mathcal{P}), \forall \bar{y}^* \in S(\mathcal{D})$ se tiene que*

$$\varphi(\bar{x}, 0) + \varphi^*(0, \bar{y}^*) = 0$$

o, de manera equivalente, $(0, \bar{y}^*) \in \partial\varphi(\bar{x}, 0)$. Recíprocamente, si $(\bar{x}, \bar{y}^*) \in X \times Y^*$ es un par que satisface esta relación, entonces $\bar{x} \in S(\mathcal{P})$, $\bar{y}^* \in S(\mathcal{D})$ y se tiene que $\inf(\mathcal{P}) + \inf(\mathcal{D}) = 0$.

Demostración: Tenemos que

$$\varphi(\bar{x}, 0) = \inf(\mathcal{P}) = -\inf(\mathcal{D}) = -\varphi^*(0, y^*),$$

lo que prueba la condición de extremalidad. Recíprocamente,

$$\inf(\mathcal{P}) \leq \varphi(\bar{x}, 0) = -\varphi^*(0, \bar{y}^*) \leq -\inf(\mathcal{D}),$$

lo que implica que no hay salto de dualidad y $\bar{x} \in S(\mathcal{P})$, y $\bar{y}^* \in S(\mathcal{D})$. □

Corolario 3.1.3. *Supongamos que $(X, \|\cdot\|)$ es un espacio de Banach reflexivo, que $\varphi \in \Gamma_0(X \times Y)$ y $x_0 \in X$ son tales que $\varphi(x_0, \cdot)$ es finita y continua en 0, y que*

$$\lim_{\|x\| \to +\infty} \varphi(x, 0) = +\infty.$$

Entonces, $S(\mathcal{P})$ es no-vacío y acotado, $S(\mathcal{D})$ es no-vacío (y acotado si $(Y, \|\cdot\|)$ es de Banach) , $\inf(\mathcal{P}) = -\inf(\mathcal{D}) = 0$ y las soluciones de $(\mathcal{D})$ y $(\mathcal{P})$ satisfacen las condiciones de extremalidad.

Ejemplo 3.1.2 (Programación Lineal II). Teníamos que

$$(\mathcal{P}) \quad \min_{x \in \mathbf{R}^n} \{c^T x \mid Ax \leq b\},$$

y tomando

$$\varphi(x, \lambda) = c^T x + \delta_{\mathbf{R}^m_-}(Ax + \lambda - b) \in \Gamma_0(\mathbf{R}^n \times \mathbf{R}^m)$$

obtuvimos que

$$(\mathcal{D}) \quad \min_{\lambda \in \mathbf{R}^m} \{b^T \lambda \mid A^T \lambda + c = 0,\ \lambda \geq 0\}.$$

Evidentemente $\inf(\mathcal{D}) \geq \inf(\mathcal{P})$. Supongamos que existe $x_0 \in \mathbf{R}^n$ que satisface la *Condición de Slater*:

$$Ax_0 < b.$$

Entonces, para $\epsilon > 0$ suficientemente pequeño, se tiene que si $\|\lambda\| \leq \epsilon$, entonces $\varphi(x_0, \lambda) = c^T x$ de modo que $\varphi(x_0, \cdot)$ es finita y continua en una vecindad de 0. Así, v es finita y continua en una vecindad de 0 y, en consecuencia, $S(\mathcal{D}) = \partial v(0)$ es no-vacío y compacto. En ese caso $\alpha = \inf(\mathcal{P}) = -\inf(\mathcal{D}) = \min(\mathcal{D}) \in \mathbf{R}$ y el mínimo se alcanza; en efecto, se tiene que $\{x \in \mathbf{R}^n \mid Ax \leq b\} \neq \emptyset$ y $S(\mathcal{P}) = \{x \in \mathbf{R}^n \mid Ax \leq b\} \cap \{x \in \mathbf{R}^n \mid c^T x = \alpha\}$ y si $S(\mathcal{P}) = \emptyset$ entonces existiría $\alpha' < \alpha$ tal que $[Ax \leq b] \cap [c^T x \leq \alpha'] \neq \emptyset$ lo que es una contradicción.

Recapitulando, si $(\mathcal{P})$ o $(\mathcal{D})$ tiene valor óptimo finito, entonces el otro también y las soluciones $\bar{x} \in S(\mathcal{P}), \bar{\lambda} \in S(\mathcal{D})$ satisfacen

$$c^T \bar{x} + b^T \lambda = 0, \quad A\bar{x} \leq b, \quad A\bar{\lambda} + c = 0, \quad \bar{\lambda} \geq 0.$$

Notemos que $S(\mathcal{P})$ es acotado si, y sólo si,

$$\forall d \in \mathbf{R}^n,\ [c^T d = 0,\ Ad \leq 0 \Rightarrow d = 0].$$

3.2. Dualidad Lagrangeana

Consideremos el siguiente problema de programación matemática

$$(\mathcal{P}) \quad \alpha := \inf\{f(x) \mid g_i(x) \leq 0, i = 1, ..., p, \quad h_j(x) = 0, j = 1, ..., q\}$$

donde $f, g_i, h_j \colon X \to \overline{\mathbf{R}}$. Sea

$$C = \{x \in X \mid g_i(x) \leq 0, i = 1, ..., p, \quad h_j(x) = 0, j = 1, ..., q\},$$

de modo que $(\mathcal{P})$ equivale a

$$(\mathcal{P}) \quad \inf[f + \delta_C].$$

En lo que sigue suponemos que $\alpha \in \mathbf{R}$.

Lema 3.2.1. *Para cada $x \in X$ se tiene que*

$$\delta_C(x) = \sup_{\lambda \in \mathbf{R}^p_+,\ \mu \in \mathbf{R}^q} \{\langle \lambda, g(x) \rangle + \langle \mu, h(x) \rangle$$

donde $g(x) = (g_i(x))_{i=1}^p$, $h(x) = (h_j(x))_{j=1}^q$ y $\mathbf{R}^q_+ = \{\lambda \in \mathbf{R}^q \mid \lambda_i \geq 0, i = 1, ..., q\}$.

Demostración: Obviamente,

$$\sup_{\lambda_i \geq 0} \lambda_i g_i(x) = \begin{cases} 0 & \text{si } g_i(x) \leq 0, \\ +\infty & \text{si } g_i(x) > 0, \end{cases}$$

y

$$\sup_{\mu_j \geq 0} \mu_j h_j(x) = \begin{cases} 0 & \text{si } h_j(x) = 0, \\ +\infty & \text{si } h_j(x) \neq 0, \end{cases}$$

de modo que

$$\delta_C(x) = \sup_{\lambda \in \mathbf{R}^p_+,\ \mu \in \mathbf{R}^q} \left\{ \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \mu_j h_j(x) \right\}.$$

$\square$

La *Función Lagrangeana* o *Lagrangeano* de $(\mathcal{P})$ es la función $L \colon X \times \mathbf{R}^p \times \mathbf{R}^q \to \overline{\mathbf{R}}$ definida mediante

$$L(x, \lambda, \mu) = \begin{cases} f(x) + \langle \lambda, g(x) \rangle + \langle \mu, h(x) \rangle & \text{si } \lambda \in \mathbf{R}^p_+, \\ -\infty & \text{si no.} \end{cases}$$

Tenemos que el problema $(\mathcal{P})$ es equivalente a

$$(\mathcal{P}) \quad \inf_{x \in X} \sup_{(\lambda,\mu) \in \mathbf{R}^p \times \mathbf{R}^q)} L(x, \lambda, \mu).$$

Definimos el problema dual al intercambiar "inf" con "sup" en la formulación Primal

$$(\mathcal{D}) \quad \gamma := \sup_{(\lambda,\mu) \in \mathbf{R}^p \times \mathbf{R}^q)} \inf_{x \in X} L(x, \lambda, \mu).$$

En general, $\alpha \neq \gamma$. Sin embargo, como $L(x,\bar{\lambda},\bar{\mu}) \leq \sup_{(\lambda,\mu)\in \mathbf{R}^p\times\mathbf{R}^q)} L(x,\lambda,\mu)$, se tiene que

$$L(x,\bar{\lambda},\bar{\mu}) \leq \inf_{x\in X} \sup_{(\lambda,\mu)\in \mathbf{R}^p\times\mathbf{R}^q)} L(x,\lambda,\mu) = \alpha,$$

de donde $\gamma \leq \alpha$. Si eventualmente $\alpha = \gamma$, y el problema dual $(\mathcal{D})$ tiene una solución $(\hat{\lambda},\hat{\mu})$ entonces necesariamente

$$\inf_{x\in X} L(x,\hat{\lambda},\hat{\mu}) = \gamma = \alpha$$

y como dado $x \in X$ tenemos que

$$\sup_{(\lambda,\mu)\in \mathbf{R}^p\times\mathbf{R}^q} L(x,\lambda,\mu) \geq L(x,\hat{\lambda},\hat{\mu}) \geq \alpha = \inf_{x\in X} \sup_{(\lambda,\mu)\in \mathbf{R}^p\times\mathbf{R}^q} L(x,\lambda,\mu),$$

y deducimos que si $\hat{x} \in S(\mathcal{P})$ entonces $\hat{x}$ es solución de

$$(\hat{\mathcal{P}}) \quad \alpha = \inf_{x\in X} L(x,\hat{\lambda},\hat{\mu}),$$

problema que es interesante pues es un problema de minimización sin restricciones. Un par $(\hat{\lambda},\hat{\mu}) \in \mathbf{R}^p_+ \times \mathbf{R}^q$ con estas propiedades se llama *Multiplicador de Lagrange* de $(\mathcal{P})$.

Proposición 3.2.1. *Las siguientes son equivalentes.*

(i) $\hat{x}$ es una solución de $(\mathcal{P})$ y $(\hat{\lambda},\hat{\mu})$ es un multiplicador de Lagrange de $(\mathcal{P})$ con $\alpha = \gamma$.

(ii) $(\hat{x},\hat{\lambda},\hat{\mu})$ es un punto silla del Lagrangeano de $(\mathcal{P})$, es decir,

$$\forall x \in X, \forall(\lambda,\mu) \in \mathbf{R}^p \times \mathbf{R}^q, \quad L(\hat{x},\lambda,\mu) \leq L(\hat{x},\hat{\lambda},\hat{\mu}) \leq L(x,\hat{\lambda},\hat{\mu}).$$

Ambas afirmaciones implican la condición de complementariedad

$$\hat{\lambda}_i g_i(\hat{x}) = 0,\ \ i = 1,...,p.$$

Demostración:

- $(i) \Rightarrow (ii)$ Tenemos que $\hat{x}$ es solución de $(\mathcal{P})$ de modo que

  $$\forall x \in X, \quad L(\hat{x},\hat{\lambda},\hat{\mu}) \leq L(x,\hat{\lambda},\hat{\mu}).$$

  Por otra parte,

  $$L(\hat{x},\lambda,\mu) \leq \sup_{\lambda,\mu} L(\hat{x},\lambda,\mu) = f(\hat{x}) + \delta_C(\hat{x}) = \alpha = L(\hat{x},\hat{\lambda},\hat{\mu}).$$

- $(ii) \Rightarrow (i)$ Tenemos que

  $$\gamma \leq \alpha \leq \sup_{\lambda,\mu} L(\hat{x},\lambda,\mu) = L(\hat{x},\hat{\lambda},\hat{\mu}) = \inf_{x\in X} L(x,\lambda,\mu) \leq \gamma,$$

  de modo que $\gamma = \alpha$. Evidentemente, $f(\hat{x})+\delta_C(\hat{x}) = \sup_{\lambda,\mu} L(\hat{x},\lambda,\mu) = \alpha$, de modo que $\hat{x} \in S(\mathcal{P})$, y además

  $$\gamma = \sup_{\lambda,\mu} \inf_{x\in X} L(x,\lambda,\mu) \leq \sup_{\lambda,\mu} L(\hat{x},\lambda,\mu) = L(\hat{x},\hat{\lambda},\hat{\mu}) \leq \gamma,$$

lo que implica que $(\lambda, \mu) \in S(\mathcal{D})$. Más aun, como necesariamente $\hat{x} \in C$, tenemos que

$$f(\hat{x}) = L(\hat{x}, \hat{\lambda}, \hat{\mu}) = f(\hat{x}) + \langle \lambda, g(\hat{x}) \rangle + \langle \mu, h(\hat{x}) \rangle = f(\hat{x}) + \langle \lambda, g(\hat{x}) \rangle,$$

de modo que $\langle \lambda, g(\hat{x}) \rangle = 0$. Como $g(\hat{x}) \leq 0$ y $\lambda \geq 0$ concluimos que $\lambda_i g_i(\hat{x}) = 0,\ i = 1, \ldots, p$. □

Observación 3.2.1. La existencia del multiplicador de Lagrange no siempre se tiene, incluso para datos $f, (g_i)_i, (h_j)_j$ muy regulares. Volveremos a este problema más adelante.

Pero, ¿cuál es la relación con la dualidad vía perturbaciones? Sea $Y^* := \mathbf{R}^p \times \mathbf{R}^q$ y denotemos $y^* = (\lambda, \mu)$. Tenemos

$$(\mathcal{P}) \quad \alpha = \inf_{x \in X} \sup_{y^* \in Y^*} L(x, y^*)$$

y

$$(\mathcal{D}) \quad \gamma = \sup_{y^* \in Y^*} \inf_{x \in X} L(x, y^*).$$

Definamos $\varphi \colon X \times Y \to \mathbf{R} \cup \{+\infty\}$ mediante

$$\varphi(x, y) := \sup_{y^* \in Y^*} \{\langle y^*, y \rangle + L(x, y^*)\} = (-L(x, \cdot))^*(y).$$

Entonces, $(\mathcal{P})$ es equivalente a

$$(\mathcal{P}) \quad \alpha = \inf_{x \in X} \varphi(x, 0).$$

Por otra parte, si $y^* \mapsto -L(x, y^*)$ es convexa, s.c.i. y propia, entonces

$$-L(x, y^*) = \sup_{y \in Y} \{\langle y^*, y \rangle - \varphi(x, y)\}.$$

Luego,

$$\inf_{x \in X} L(x, y^*) = -\sup_{x \in X} \sup_{y \in Y} \{\langle y^*, y \rangle - \varphi(x, y)\} = -\varphi^*(0, y^*),$$

lo que conduce a que $(\mathcal{D})$ es equivalente a

$$(\mathcal{D}) \quad \gamma = \sup_{y^* \in Y^*} -\varphi^*(0, y^*) = -\inf_{y^* \in Y^*} \varphi^*(0, y^*) = -\beta.$$

Por otra parte, dada una función de perturbación $\varphi \colon X \times Y \to \mathbf{R} \cup \{+\infty\}$ se define la función Lagrangeana $L \colon X \times Y^* \to \overline{\mathbf{R}}$ mediante

$$-L(x, y^*) := \sup_{y \in Y} \{\langle y^*, y \rangle - \varphi(x, y)\} = [\varphi(x, \cdot)]^*(y^*).$$

Lema 3.2.2.

- *$\forall x \in X,\ y^* \mapsto L(x, y^*)$ es cóncava, s.c.s. de Y^* en $\overline{\mathbf{R}}$.*
- *Si φ es convexa, entonces $\forall y^* \in Y^*,\ x \mapsto L(x, y^*)$ es convexa de X en $\overline{\mathbf{R}}$, pero no necesariamente es s.c.i., incluso cuando $\varphi \in \Gamma_0(X \times Y)$.*

Demostración: Por definición

$$-L(x,\cdot) = [\varphi(x,\cdot)]^*(\cdot) \in \Gamma(Y^*).$$

Por otra parte,

$$L(x,y^*) = \inf_{y\in Y}\{\varphi(x,y) - \langle y^*,y\rangle\} = \inf_{y\in Y}\Phi(x,y,y^*)$$

con Φ convexa, luego $L(\cdot,y^*)$ es convexa. $\square$

Análogamente,

$$\begin{aligned}\varphi^*(x^*,y^*) &= \sup_{(x,y)\in X\times Y}\{\langle x^*,x\rangle + \langle y^*,y\rangle - \varphi(x,y)\}\\ &= \sup_{x\in X}\{\langle x^*,x\rangle + \sup_{y\in Y}\langle y^*,y\rangle - \varphi(x.y)\}\\ &= \sup_{x\in X}\{\langle x^*,x\rangle - L(x,y^*)\},\end{aligned}$$

de donde

$$\varphi^*(0,y^*) = -\inf_{x\in X} L(x,y^*).$$

Así el problema dual se puede escribir en términos de L como

$$(\mathcal{D})\quad \beta = \inf_{y^*\in Y^*}\varphi^*(0,y^*) = -\sup_{y^*\in Y^*}\inf_{x\in X} L(x,y^*).$$

Similarmente, si $\varphi\in\Gamma_0(X\times Y)$, entonces $\varphi(x,\cdot)\in\Gamma_0(Y)$ y en consecuencia $[\varphi(x,\cdot)]^{**}(y) = \varphi(x,y)$, de donde

$$\varphi(x,y) = \sup_{y^*\in Y^*}\{\langle y^*,y\rangle + L(x,y^*)\},$$

en particular $\varphi(x,0) = \sup_{y^*\in Y^*} L(x,y^*)$ y tenemos que

$$(\mathcal{P})\quad \alpha = \inf_{x\in X}\varphi(x,0) = \inf_{x\in X}\sup_{y^*\in Y^*} L(x,y^*).$$

Proposición 3.2.2. *Si $\varphi\in\Gamma_0(X\times Y)$ entonces las siguientes son equivalentes*

(i) $\hat{x}$ es solución de $(\mathcal{P})$, $\hat{y}^*$ es solución de $(\mathcal{D})$ y $\alpha+\beta=0$.

(ii) $(\hat{x},\hat{y}^*)\in X\times Y^*$ es un punto silla de L (con $L(\hat{x},\hat{y}^*)\in\mathbf{R}$), es decir,

$$\forall x\in X,\ \forall y^*\in Y^*,\quad L(\hat{x},y^*)\le L(\hat{x},\hat{y}^*)\le L(x,\hat{y}^*).$$

Demostración: Propuesto. $\square$

3.3. Teoremas de Dualidad de Fenchel-Rockafellar y de Attouch-Brezis

Comenzaremos esta sección con algunos resultados preliminares.

Proposición 3.3.1. *Sea X un e.v.n. y $C \subseteq X$ un convexo de interior no-vacío. Entonces*

$$\operatorname{int}(C) = \operatorname{int}(\overline{C}).$$

Demostración: Obviamente, $\operatorname{int}(C) \subseteq \operatorname{int}(\overline{C})$. Como $\operatorname{int}(C) \neq \emptyset$, deducimos que $\operatorname{int}(\overline{C}) \neq \emptyset$. Sean $\bar{x} \in \operatorname{int}(\overline{C})$ y $x_0 \in \operatorname{int}(C)$. Dado $\epsilon > 0$, definimos $x_1 = \bar{x} + \epsilon(\bar{x} - x_0)$ y si ϵ es suficientemente pequeño, entonces $x_1 \in \overline{C}$. Como $\bar{x} \in]x_0, x_1[$, basta probar que $]x_0, x_1[\subseteq \operatorname{int}(C)$ para deducir que $\bar{x} \in \operatorname{int}(C)$ y de aquí que $\operatorname{int}(C) = \operatorname{int}(\overline{C})$. Como $x_0 \in \operatorname{int}(C)$, podemos tomar $r > 0$ tal que $B(x_0, r) \subseteq C$. Sean $\lambda \in]0, 1[$ y $x_\lambda := \lambda x_0 + (1 - \lambda)x_1$.

1. Reflexión de $B(x_0, r)$ a través de x_λ.
   Tenemos que

$$x_1 = \frac{1}{1-\lambda} x_\lambda - \frac{\lambda}{1-\lambda} x_0 \in \frac{1}{1-\lambda} x_\lambda - \frac{\lambda}{1-\lambda} B(x_0, r) = B\left(x_1, \frac{\lambda}{1-\lambda} r\right)$$

   y como $x_1 \in \overline{C}$, deducimos que

$$\exists x_2 \in C \cap \left[\frac{1}{1-\lambda} x_\lambda - \frac{\lambda}{1-\lambda} B(x_0, r)\right].$$

2. Envoltura convexa.
   Como $B(x_0, r) \subseteq C$ y $x_2 \in C$ con C convexo, en particular se tiene que

$$\lambda B(x_0, r) + (1 - \lambda)x_2 \subseteq C.$$

   Pero $x_2 = \frac{1}{1-\lambda} x_\lambda - \frac{\lambda}{1-\lambda} y$ para algún $y \in B(x_0, r)$ lo que implica que $x_\lambda = \lambda y + (1 - \lambda)x_2 \in C$ y concluye la demostración.

□

Observación 3.3.1. La hipótesis $\operatorname{int}(C) \neq \emptyset$ es esencial. Por ejemplo, si C es un hiperplano denso, entonces $\operatorname{int}(C) = \emptyset$ pero $\overline{C} = X$ de modo que $\operatorname{int}(\overline{C}) = X$.

Ejercicio 3.3.1. Si X es un e.v.n. y $C \subseteq X$ es un convexo con $\operatorname{int}(C) \neq \emptyset$, entonces

$$\overline{C} = \overline{\operatorname{int}(C)}.$$

El interés del siguiente resultado es que no requiere saber a priori que $\operatorname{int}(C) \neq \emptyset$.

Lema 3.3.1 (Lema de Robinson)**.** *Sean $(X, \|\cdot\|)$ un espacio de Banach, Y un e.v.n. y $C \subseteq X \times Y$ un convexo cerrado. Denotemos por C_X y C_Y sus proyecciones sobre X e Y respectivamente. Si C_X es acotado, entonces* $\operatorname{int}(C_Y) = \operatorname{int}(\overline{C_Y})$.

Demostración: La inclusion $\mathrm{int}(C_Y) \subseteq \mathrm{int}(\overline{C_Y})$ es evidente. Sean $y \in \mathrm{int}(\overline{C_Y})$ y $\epsilon > 0$ tales que $B(y,\epsilon) \subseteq \overline{C_Y}$. Basta demostrar que existe $x \in X$ tal que $(x,y) \in C$, pues en este caso $y \in C_Y$ y se concluye que $\mathrm{int}(\overline{C_Y}) \subseteq \mathrm{int}\, C_Y$. Para demostrar que tal $x \in X$ existe, construiremos una sucesión $(x_n, y_n) \in C$ con $y_n \to y$ y $(x_n)_n$ de Cauchy. Tomemos $(x_0, y_0) \in C$ (si $C = \emptyset$ la propiedad es trivial) y consideremos el siguiente algoritmo:
Mientras $y_n \neq y$

1. Defínase $\alpha_n := \frac{\epsilon}{2\|y_n - y\|}$ y $w = y + \alpha_n(y - y_n) \in B(y,\epsilon) \subseteq \overline{C_Y}$.
2. Sea $(u,v) \in C$ tal que $\|w - v\| \leq \frac{1}{2}\|y - y_n\|$.
3. Defínase $(x_{n+1}, y_{n+1}) := \frac{\alpha_n}{1+\alpha_n}(x_n, y_n) + \frac{1}{1+\alpha_n}(u,v) \in C$.

Fin. Si el algoritmo se detiene en el paso N, tomamos $x = x_N$. Supongamos que el algoritmo no se detiene, en cuyo caso tenemos que $((x_n, y_n))_n \subseteq C$ tal que

$$\begin{aligned}
\|y_{n+1} - y\| &= \left\| \frac{\alpha_n}{1+\alpha_n} y_n + \frac{1}{1+\alpha_n} v - y \right\| \\
&= \frac{1}{1+\alpha_n} \|\alpha_n(y_n - y) - y + v\| \\
&= \frac{1}{1+\alpha_n} \|w - v\| \\
&\leq \frac{1}{2}\|y - y_n\|,
\end{aligned}$$

lo que implica que $\|y_{n+1} - y\| \leq \frac{1}{2}\|y_n - y\|$ y en consecuencia $y_n \to y$. Por otra parte

$$\begin{aligned}
\|x_{n+1} - x_n\| &= \left\| \frac{\alpha_n}{1+\alpha_n} x_n + \frac{1}{1+\alpha_n} u - x_n \right\| \\
&= \frac{1}{1+\alpha_n} \|u - x_n\| \\
&\leq \frac{\mathrm{diam}(C_X)}{\alpha_n} = \frac{2\,\mathrm{diam}(C_X)}{\epsilon} \|y_n - y\|
\end{aligned}$$

lo que implica que (x_n) es de Cauchy y, por ser X un espacio de Banach y C un cerrado, existe $x \in C$ tal que $x_n \to x$ y $(x,y) \in C$.

$\square$

Lema 3.3.2. *Si X es un espacio de Banach y $C \subseteq X$ es un convexo cerrado y aborbente, entonces $0 \in \mathrm{int}(C)$.*

Demostración: Como C es absorbente, tenemos que $X = \bigcup_{k \in \mathbf{N}} kC$ con

$$kC = \{kx \mid x \in C\}$$

cerrado. Del Lema de Baire, deducimos que existe $k_0 \in \mathbf{N}$ tal que $\mathrm{int}(k_0 C) \neq \emptyset$ lo que implica que $\mathrm{int}(C) \neq \emptyset$. Más aun, podemos suponer que $C = -C$ (si no, podemos tomar el convexo absorbente $C \cap (-C)$), en cuyo caso debe tenerse que $0 \in \mathrm{int}(C)$. $\square$

Teorema 3.3.1 (Attouch-Brezis). *Sean X, Y dos espacios de Banach y $\varphi \in \Gamma_0(X \times Y)$. Sea $v(y) = \inf_{x\in X} \varphi(x, y)$ y supongamos que $v(0) \in \mathbf{R}$ y que el convexo*

$$\operatorname{dom}(v) = \bigcup_{x\in X} \operatorname{dom}(\varphi(x, \cdot))$$

es absorbente en Y. Entonces, v es continua en 0 . En particular,

$$v(0) = -\min_{y^*\in Y^*} \varphi^*(0, y^*),$$

es decir, el dual tiene solución.

Demostración: Sea $k > v(0)$ y definamos

$$U = \{y \in Y \mid v(y) \leq k\}.$$

Basta probar que $0 \in \operatorname{int}(U)$. Sea $x_0 \in X$ tal que $\varphi(x_0, 0) < k$ y definamos

$$C = \{(x, y) \in X \times Y \mid \varphi(x, y) \leq k,\ \|x\| \leq 1 + \|x_0\|\},$$

que es un convexo cerrado cuya proyección sobre X, C_X, es acotada y $C_Y \subseteq U$. Basta probar que C_Y es una vecindad del origen, pero, gracias al Lema de Robinson, $\operatorname{int}(C_Y) = \operatorname{int}(\overline{C_Y})$, por lo que basta probar que $0 \in \operatorname{int}(\overline{C_Y})$. Veamos que C_Y es absorbente en el Banach Y. Sea $y \in Y$. Por hipótesis, existen $\lambda > 0$ y $x \in X$ tales que $\varphi(x, \lambda y) < +\infty$. Dado $t \in]0, 1[$, tenemos que $\varphi((1-t)x_0 + tx, t\lambda y) \leq (1-t)\varphi(x_0, 0) + t\varphi(x, \lambda y)$ y si t es suficientemente pequeño,

$$\varphi((1-t)x_0 + tx, t\lambda y) \leq k, \quad \|(1-t)x_0 + tx\| \leq 1 + \|x_0\|.$$

Luego, existe $t > 0$ tal que $((1-t)x_0 + tx_0, t\lambda y) \in C$ y se sigue que C_Y es absorbente. Por lo tanto, $\overline{C_Y}$ es absorbente y $0 \in \operatorname{int}(\overline{C_Y})$. $\square$

Corolario 3.3.1 (Teorema de Dualidad de Fenchel-Rockafellar). *Sean X, Y espacios de Banach, $f \in \Gamma_0(X)$, $g \in \Gamma_0(Y)$ y $A \colon X \to Y$ lineal continua. Consideremos los problemas de minimización*

$$(\mathcal{P}) \quad \inf_{x\in X} f(x) + g(Ax)$$

y

$$(\mathcal{D}) \quad \inf_{y^*\in Y^*} f^*(-A^*y^*) + g^*(y^*).$$

Supongamos que se tiene que $\inf(\mathcal{P}) \in \mathbf{R}$ *y la hipótesis de Calificación Primal:*

$$0 \in \operatorname{int}(\operatorname{dom}(g) - A\operatorname{dom}(f)),$$

entonces $S(\mathcal{D}) \neq \emptyset$.
Similarmente, si se tiene que $\inf(\mathcal{D}) \in \mathbf{R}$ *y la hipótesis de Calificación Dual:*

$$0 \in \operatorname{int}(\operatorname{dom}(f^*) - A^*\operatorname{dom}(g^*)),$$

entonces $S(\mathcal{P}) \neq \emptyset$.
En cualquier caso, $\inf(\mathcal{P}) = -\inf(\mathcal{D})$. *Además, son equivalentes*

(i) $\bar{x} \in S(\mathcal{P})$, $\bar{y^*} \in S(\mathcal{D})$.

(ii) Las Relaciones de Extremalidad:

$$f(\bar{x}) + f^*(-A^*\bar{y^*}) = \langle \bar{x}, -A^*\bar{y^*} \rangle, \quad g(A\bar{x}) + g^*(\bar{y^*}) = \langle A\bar{x}, \bar{y^*} \rangle.$$

(iii) La Inclusión de Euler-Lagrange:

$$-A^*\bar{y^*} \in \partial f(\bar{x}), \quad \bar{y^*} \in \partial g(A\bar{x}).$$

Demostración: Sea $\varphi(x,y) = f(x) + g(Ax+y)$. Tenemos que

$$\begin{aligned}
\varphi^*(x^*, y^*) &= \sup_{(x,y)\in X\times Y} \{\langle x^*, x \rangle + \langle y^*, y \rangle - f(x) - g(Ax+y)\} \\
&= \sup_{(x,z)\in X\times Y} \{\langle x^*, x \rangle + \langle y^*, z - Ax \rangle - f(x) - g(z)\} \\
&= f^*(x^* - A^*y^*) + g^*(y^*).
\end{aligned}$$

Así, $(\mathcal{P})$ equivale a $\inf \varphi(\cdot, 0)$ y $(\mathcal{D})$ equivale a $\inf \varphi^*(0, \cdot)$. Tenemos que $y \in \mathrm{dom}(\varphi(x, \cdot))$ si, y sólo si, $f(x) < +\infty$ y $Ax + y \in \mathrm{dom}(g)$, de modo que

$$\bigcup_{x\in X} \mathrm{dom}(\varphi(x, \cdot)) = \mathrm{dom}(g) - A\,\mathrm{dom}(f)$$

y podemos aplicar el Teorema de Attouch-Brezis. En este caso, la condición de Extremalidad se escribe

$$\varphi(\bar{x}, 0) + \varphi^*(0, \bar{y^*}) = 0,$$

que es equivalente a

$$f(\bar{x}) + f^*(-A^*\bar{y^*}) + \langle \bar{x}, A^*\bar{y^*} \rangle - \langle A\bar{x}, \bar{y^*} \rangle + g(A\bar{x}) + g^*(y^*) = 0,$$

lo que a su vez es equivalente a

$$f(\bar{x}) + f^*(-A^*\bar{y^*}) = \langle \bar{x}, -A^*\bar{y^*} \rangle, \; g(A\bar{x}) + g^*(y^*) = \langle A\bar{x}, \bar{y^*} \rangle.$$

que son las relaciones de Extremalidad. Claramente estas relaciones son equivalentes a la inclusión de Euler-Lagrange. $\square$

3.4. Teoremas de Fritz John y Kuhn-Tucker

Sea X un espacio de Banach y consideremos $f, g_i \colon X \to \overline{\mathbf{R}}$, $i = 1, \ldots, p$ convexas y propias, h_j, $j = 1, \ldots, q$ lineales afines y continuas. El programa convexo asociado a $f, (g_i)_i, (h_j)_j$ es el problema de optimización

$$(\mathcal{P}) \quad \min_{x\in X} \{f(x) \mid g_i(x) \leq 0, i = 1, \ldots, p, \quad h_j(x) = 0, j = 1, \ldots, q\},$$

que tiene asociada la función lagrangeana $L\colon X \times \mathbf{R}^p \times \mathbf{R}^q \to \overline{\mathbf{R}}$ dada por

$$L(x,\lambda,\mu) = \begin{cases} f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \mu_j h_j(x) & \text{si } \lambda \geq 0, \\ -\infty & \text{si no,} \end{cases}$$

de modo que

$$(\mathcal{P}) \quad v(\mathcal{P}) := \inf_{x\in X} \sup_{(\lambda,\mu)\in\mathbf{R}^p\times\mathbf{R}^q} L(x,\lambda,\mu).$$

Suponemos que $v(\mathcal{P})$ es finito. El dual asociado a L es

$$(\mathcal{D}) \quad \sup_{(\lambda,\mu)\in\mathbf{R}^p\times\mathbf{R}^q} \inf_{x\in X} L(x,\lambda,\mu),$$

cuyas soluciones se llaman multiplicadores de Lagrange, de tenerse $\inf(\mathcal{P}) = \sup(\mathcal{D})$.

Teorema 3.4.1 (Fritz-John Débil)**.** *Definamos $D := \mathrm{dom}(f) \cap (\bigcap_{i=1}^p \mathrm{dom}(g))$ y $H := \bigcap_{j=1}^q \ker(h_j)$. Si $v(\mathcal{P}) \in \mathbf{R}$, f y g_i, $i = 1,\ldots,p$ tienen un punto de continuidad en común y además $0 \in \mathrm{int}(D\setminus H)$, entonces existen reales no-negativos $\hat{\lambda}_0,\ldots,\hat{\lambda}_p$, no todos nulos y reales $\hat{\mu}_1,\ldots,\hat{\mu}_q \in \mathbf{R}$ tales que*

$$\forall x \in D,\ \hat{\lambda}_0 v(\mathcal{P}) \leq \hat{\lambda}_0 f(x) + \sum_{i=1}^p \hat{\lambda}_i g_i(x) + \sum_{j=1}^q \hat{\mu}_j h_j(x).$$

Observación 3.4.1. En este teorema $(\hat{\lambda},\hat{\mu})$ no es necesariamente un multiplicador de Lagrange. Pero si, por ejemplo, $D = X$, f, g_i son finitas en todas partes y $\hat{\lambda_0} = 1$, entonces se tendrá que

$$\forall x \in X,\ v(\mathcal{P}) \leq L(x,\hat{\lambda},\hat{\mu})$$

lo que es equivalente a

$$v(\mathcal{P}) \leq \inf_{x\in X} L(x,\hat{\lambda},\hat{\mu}) \leq v(\mathcal{D})$$

e implica que $v(\mathcal{P}) = v(\mathcal{D})$ y así $(\hat{\lambda},\hat{\mu}) \in S(\mathcal{D})$ será un multiplicador de Lagrange de $(\mathcal{P})$.

Para la demostración del Teorema, necesitaremos algunas variantes del Teorema de Hahn-Banach.

Teorema 3.4.2. *Sea $C \subseteq X$ un convexo de interior no-vacío. Si $0 \notin C$, entonces existe $x^* \in X^*\setminus\{0\}$ tal que*

$$\forall x \in X,\ \langle x^*, x\rangle \geq 0.$$

Demostración: Podemos separar $\mathrm{int}(C)$ de 0 mediante un hiperplano cerrado, de modo que existe $x^* \in X^* \setminus \{0\}$ tal que $\mathrm{int}(C) \subseteq [\langle x^*,\cdot\rangle \geq 0]$. Luego

$$\overline{C} = \overline{\mathrm{int}(C)} \subseteq [\langle x^*,\cdot\rangle \geq 0].$$

$\square$

Teorema 3.4.3 (Teorema de Separación)**.** *Sean $A, B \subseteq X$ dos convexos no-vacíos. Si $\mathrm{int}(A-B) \neq \emptyset$ y $A\cap B = \emptyset$, entonces A y B pueden separarse mediante un hiperplano. Si $0 \notin \overline{(A-B)}$ y $A\cap B = \emptyset$, entonces pueden separarse estrictamente mediante un hiperplano.*

Demostración: Basta observar que A y B pueden separarse (resp. estrictamente) mediante un hiperplano ssi $A-B$ puede separse (resp. estrictamente) de 0 mediante un hiperplano. $\square$

Teorema 3.4.4. *Sea $F\colon X \to \overline{\mathbf{R}}$ convexa y propia con $\operatorname{int}(\operatorname{epi}(F)) \neq \emptyset$. Sea $H \subseteq X$ un subespacio vetorial de X y $L\colon H \to \mathbf{R}$ una función lineal tal que*

$$\forall x \in X,\ F(x) \geq L(x).$$

Si $0 \in \operatorname{int}(\operatorname{dom}(F) - H)$ entonces existe $x_0^* \in X^*$ tal que

$$\forall x \in H,\ L(x) = \langle x_0^*, x\rangle;$$

$$\forall x \in X,\ \langle x_0^*, x\rangle \leq F(x).$$

Demostración: Sean los convexos $A := \{(x,\lambda) \in X \times \mathbf{R} \mid F(x) < \lambda\}$ y $B := \{(x,\lambda) \in X \times \mathbf{R} \mid x \in H, \lambda = L(x)\}$. Se tiene que $\operatorname{int}(A) = \operatorname{int}(\operatorname{epi}(F)) \neq \emptyset$ y $A \cap B = \emptyset$. Como $\operatorname{int}(A) \subseteq \operatorname{int}(A-B)$, deducimos que A y B pueden separarse mediante un hiperplano cerrado, es decir, existe $(x^*, s) \in X^* \times \mathbf{R}$ tal que

$$\forall y \in H, \forall (x,\lambda) \in A, \quad \langle x^*, y\rangle + sL(y) \leq \langle x^*, x\rangle + s\lambda.$$

Como H es un s.e.v. se tiene que $\forall y \in H, \langle x^*, y\rangle + sL(y) = 0$ y haciendo $\lambda \to \infty$ deducimos que $s \geq 0$. Si $s = 0$, tendríamos en particular que $\forall y \in H, \forall x \in \operatorname{dom}(F),\ \langle x^*, x-y\rangle \geq 0$ y como $0 \in \operatorname{int}(\operatorname{dom}(F) - H)$ se sigue qiue $x^* = 0$ y luego que $(x^*, s) = (0,0)$, contradicción. Así, $s > 0$. Definiendo $x_0^* = -\frac{x^*}{s}$ obtenemos que $\forall y \in H,\ L(y) = \langle x_0^*, y\rangle$ y $\forall (x,\lambda) \in A,\ \lambda \geq \langle x_0^*, x\rangle$. Para $x \in \operatorname{dom}(F)$, haciendo $\lambda \to F(x)$ se deduce el resultado. $\square$

Observación 3.4.2. Cuando $F(x) = \|x\|$ o bien F es un funcional de Minkowski continuo en X (de modo que $\operatorname{dom}(F) = X$), este teorema se conoce como el Teorema de Hahn-Banach analítico.

Para demostrar el Teorena de Fritz-John necesitaremos el siguiente teorema general sobre funciones convexas.

Teorema 3.4.5. *Sean $f_i\colon X \to \overline{\mathbf{R}}$, $i = 1,\ldots,m$ funciones convexas tales que para cada $x \in X$, $\max_{1\leq i\leq m} f_i(x) \geq 0$. Entonces existen reales no-negativos $\nu_1,\ldots,\nu_m$ y no todos nulos tales que $\forall x \in \bigcap_{i=1}^m \operatorname{dom}(f_i),\ \nu_1 f_1(x) + \cdots + \nu_m f_m(x) \geq 0$.*
Si además existe $\hat{x} \in X$ tal que $\max_{1\leq i\leq m} f_i(\hat{x}) = 0$, entonces $\nu_1,\ldots,\nu_m$ satisfacen $\nu_i f_i(\hat{x}) = 0,\ i = 1,\ldots m$.

Demostración: Supongamos que $\bigcap_{i=1}^m \operatorname{dom}(f_i) \neq \emptyset$ (el caso $\bigcap_{i=1}^m \operatorname{dom}(f_i) = \emptyset$ es trivial) y definamos el conjunto

$$A := \{y = (y_1,\ldots,y_m) \in \mathbf{R}^m \mid \exists x \in X, f_1(x) < y_1,\ldots,f_m(x) < y_m\}.$$

Es fácil ver que A es un convexo de interior no-vacío. Como $0 \notin A$, podemos separa A y $\{0\}$ mediante un hiperplano, es decir, existe $\nu = (\nu_1,\ldots,\nu_m) \in \mathbf{R}^m \setminus \{0\}$ tal que

$$\forall y \in A,\ \nu^T y = \sum_{i=1}^m \nu_i y_i \geq 0.$$

Para cada $i \in \{1, \ldots, m\}$ podemos hacer $y_i \to +\infty$ y luego $\nu_i \geq 0, \forall i$. Más aun, para todo $x \in \bigcap_{i=1}^m \text{dom}(f_i)$ con $f_i(x) > -\infty, i = 1, \ldots, m$ y $\forall \epsilon > 0$ tenemos

$$(f_1(x) + \epsilon, \ldots, f_m(x) + \epsilon) \in A.$$

De la arbitrariedad de $\epsilon > 0$, se deduce la desigualdad. Si eventualmente existe $x_0 \in \bigcap_{i=1}^m \text{dom}(f_i)$ con $f_{i_0}(x) = -infty$ para algún i_0, entonces podemos hacer $y_{i_0} \to -\infty$ y necesariamente $\nu_{i_0} = 0$. En este caso $f_{i_0}(x)+\epsilon$ puede reemplazarse por cualquier real y se deduce la desigualdad. Finalmente, si $\max_{1 \leq i \leq m} f_i(\hat{x}) = 0$, entonces $f_i(\hat{x}) \leq 0, i = 1, \ldots, m$ y $\nu_i f_i(\hat{x}) = 0$. □

Demostración: [del Teorema de Fritz-John sin h_j.] Tenemos que para todo $x \in X$, $\max_{1 \leq i \leq p}\{f(x) - v(\mathcal{P}), g_i(x)\} \geq 0$. Del teorema anterior se deduce la parte 1 del resultado. Más aun, si $\hat{x} \in S(\mathcal{P})$, entonces $\max_{1 \leq i \leq p}\{f(\hat{x}) - v(\mathcal{P}), g_i(\hat{x})\} = 0$ de modo que $\hat{\lambda_0}(f(\hat{x}) - v(\mathcal{P})) = 0$ y $\hat{\lambda_i} g_i(\hat{x}) = 0, \forall i = 1, \ldots, m$.

□

Proposición 3.4.1. *Sean $x_1^*, \ldots, x_q^* \in X^*$. Supongamos que $\bar{x}^* \in X^*$ satisface:*

$$(\forall x \in X, \forall j = 1, \ldots, q,\ \langle x_j^*, x \rangle = 0) \Rightarrow \langle \bar{x}^*, x \rangle = 0$$

o, de manera equivalente,

$$\bigcap_{j=1}^q \ker\langle x_j^*, \cdot \rangle \subseteq \ker\langle \bar{x}^*, \cdot \rangle.$$

Entonces existen $\mu_1, \ldots, \mu_q \in \mathbf{R}$ tal que $\bar{x}^* = \mu_1 x_1^* + \cdots + \mu_q x_q^*$.

Demostración: Sea $V^* := \langle \{x_1^*, \ldots, x_q^*\} \rangle$. Tenemos que V^* es convexo y débilmente cerrado para la topología débil-*. Si $\bar{x}^* \notin V^*$, entonces por Hahn-Banach existe $x \in X$ tal que

$$\forall x^* \in V^*,\ \langle \bar{x}^*, x \rangle > \langle x^*, x \rangle$$

lo que implica que $\langle \bar{x}^*, x \rangle > 0$ y $\langle x^*, x \rangle = 0, \forall x^* \in V^*$, contradicción. □

Demostración: [del Teorema de Fritz-John.] Tenemos que $h_j(x) = \langle x_j^*, x \rangle - \alpha_j$ con $x_j^* \in X^*$ y $\alpha_j \in \mathbf{R}$. Como $v(\mathcal{P}) \in \mathbf{R}$, necesariamente

$$H = \{x \in X \mid \langle x_j^*, x \rangle - \alpha_j = 0, j = 1, \ldots, q\}.$$

Sea $\bar{x} \in H$, entonces

$$H = \{x \in X \mid \langle x_j^*, x - \bar{x} \rangle = 0, j = 1, \ldots, q\}.$$

Sin pérdida de generalidad, podemos suponer que $\bar{x} = 0$, de modo que H es un subespacio cerrado de X y $h_j(x) = \langle x_j^*, x \rangle$. Además, para $x \in H$ se satisface que

$$F(x) := \max_{1 \leq i \leq p}\{f(x) - v(\mathcal{P}), g_i(x)\} \geq 0.$$

Por otra parte, existe un punto de continuidad común a f y $g_i, i = 1, \ldots, p$, luego F es continua en dicho punto y el interior de $\text{epi}(F)$ es no-vacío. Como

$$\text{dom}(F) = D = \text{dom}(f) \cap \left(\bigcap_{i=1}^p \text{dom}(g_i) \right)$$

y hemos supuesto que $0 \in \operatorname{int}(D - H)$, deducimos que existe $\bar{x}^* \in X^*$ tal que

$$\forall x \in H, \langle \bar{x}^*, x\rangle = 0 \Rightarrow \bar{x}^* = \sum_{j=1}^{q} \bar{\mu}_j x_j^*, \text{ para algunos } \bar{\mu}_j \in \mathbf{R}$$

y

$$\forall x \in X, F(x) - \langle \bar{x}^*, x\rangle \geq 0 \Leftrightarrow \forall x \in X, \max_{1\leq i\leq p}\{f(x) - v(\mathcal{P}) - \langle \bar{x}^*, x\rangle, g_i(x) - \langle \bar{x}^*, x\rangle\} \geq 0.$$

Del Teorema 3.4.5, deducimos que existen $\hat{\lambda}_0, \ldots, \hat{\lambda}_p \geq 0$ no todos nulos tales que

$$\forall x \in D,\ \hat{\lambda}_0(f(x) - v(\mathcal{P})) + \hat{\lambda}_1 g_1(x) + \cdots + \hat{\lambda}_p g_p(x) - (\hat{\lambda}_0 + \ldots \hat{\lambda}_p)\langle \bar{x}^*, x\rangle \geq 0.$$

Definiendo $\hat{\mu}_j := -(\hat{\lambda}_0 + \ldots \hat{\lambda}_p)\bar{\mu}_j$, se deduce el resultado. □

Una propiedad deseable en la práctica, es que $\hat{\lambda}_0$ sea estrictamente positivo. En los siguientes teoremas, veremos que una condición clásica en Optimización Convexa nos permite asegurar esta propiedad.

Teorema 3.4.6 (Kuhn-Tucker Débil)**.** *Bajo las hipótesis del Teorema de Fritz John débil, supongamos que se tiene la siguiente condición de calificación de restricciones, llamada* Condición de Slater

$$(\mathcal{S}) \quad \exists x_0 \in \operatorname{dom}(f) \colon g_i(x_0) < 0, i = 1, \ldots, p,\ h_j(x_0) = 0, j = 1, \ldots, q.$$

Entonces existe un par $(\hat{\lambda}, \hat{\mu}) \in \mathbf{R}^p_+ \times \mathbf{R}^q \setminus \{(0,0)\}$ *tal que*

$$\forall x \in D, \quad v(\mathcal{P}) \leq L(x, \hat{\lambda}, \hat{\mu}).$$

Demostración: Si $\hat{\lambda}_0 = 0$, basta tomar $x = x_0$ en la desigualdad de Fritz-John débil para deducir que $\sum_{i=1}^{p} \hat{\lambda}_i g_i(x_0) \geq 0$ con $\hat{\lambda}_i \geq 0$ y $g_i(x_0) < 0$. Luego $\hat{\lambda}_i = 0,\ i = 1, \ldots, p$ lo que contradice $\hat{\lambda}_0, \ldots, \hat{\lambda}_p$ no son todos nulos. □

En general, las condiciones en términos de subdiferenciales son útiles en muchas aplicaciones. Terminamos esta sección estudiando algunos resultados que, en esencia, son casos particulares de la Regla de Fermat y caracterizan el cono normal en términos de los subdiferenciales de las funciones (g_i).
La siguiente proposición caracteriza el subdiferencial del máximo puntual de funciones convexas.

Proposición 3.4.2. *Sean* $f_1, \ldots f_m \colon X \to \overline{\mathbf{R}}$ *funciones convexas y definamos*

$$F(x) := \max_{1\leq i\leq p} f_i(x).$$

Entonces,

$$\forall x \in X, \quad \partial F(x) = \bigcup_{\lambda \in \Lambda_F(x)} \partial\left(\sum_{i=1}^{m} \lambda_i f i + \delta_D\right)(x),$$

donde $D = \operatorname{dom}(F) = \bigcap_{i=1}^{m} \operatorname{dom}(f_i)$ *y*

$$\Lambda_F(x) = \{\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbf{R}^m_+ \mid \forall i = 1, \ldots, m\ \lambda_i(F(x) - f_i(x)) = 0\}.$$

Demostración: Sea $x^* \in \bigcup_{\lambda\in\Lambda_F(x)} \partial \left(\sum_{i=1}^m \lambda_i f_i(x) + \delta_D\right)(x)$. Entonces existen $\bar{\lambda}_1, \ldots, \bar{\lambda}_m \geq 0$, $\sum_{i=1}^m \bar{\lambda}_i = 1$ con $\bar{\lambda}_i(F(x) - f_i(x)) = 0,\ i = 1, \ldots, m$ tales que

$$\forall y \in D,\ \sum_{i=1}^m \bar{\lambda}_i f_i(x) + \langle x^*, y - x\rangle \leq \sum_{i=1}^m \bar{\lambda}_i f_i(y) \geq F(y),$$

lo que implica que

$$F(x) + \langle x^*, y - x\rangle \leq F(y)$$

y luego $x^* \in \partial F(x)$. Recíprocamente, sea $x^* \in \partial F(x)$ de modo que

$$\forall y \in X,\ F(x) + \langle x^*, y - x\rangle \leq F(y).$$

Definamos $G(y) := F(y) - F(x) - \langle x^*, y - x\rangle$ y $g_i(y) := f_i(y) - F(x) - \langle x^*, y - x\rangle$. Tenemos que existen $\bar{\lambda}_1, \ldots, \bar{\lambda}_m \geq 0$ no todos nulos, tales que

$$\forall y \in D,\ \sum_{i=1}^m \bar{\lambda}_i g_i(y) \geq 0,\ \text{y } \bar{\lambda}_i g_i(x) = 0, i = 1, \ldots, m.$$

Como $\sum_{i=1}^m \bar{\lambda}_i > 0$, podemos suponer que $\sum_{i=1}^m \bar{\lambda}_i = 1$. Tenemos que $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_m) \in \Lambda_F(x)$ y que $\forall y \in D$,

$$\sum_{i=1}^m \bar{\lambda}_i f_i(y) \geq \sum_{i=1}^m \bar{\lambda}_i f_i(x) + \langle x^*, y - x\rangle.$$

De aquí, $x^* \in \bigcup_{\lambda\in\Lambda_F(x)} \partial \left(\sum_{i=1}^m \bar{\lambda}_i f_i + \delta_D\right)(x)$. □

Observación 3.4.3. Si x es un punto de continuidad común para $f_1, \ldots, f_m$, entonces

$$\partial F(x) = \operatorname{co}\left(\bigcup_{i\in I_F(x)} \partial f_i(x)\right)$$

donde $I_F(x) := \{i \mid f_i(x) = F(x)\}$.

Ejercicio 3.4.1. Pruebe la observación anterior.

Diremos que el programa convexo original $(\mathcal{P})$ es propio, s.c.i. y finito si tanto f como g_i son convexas, propias y s.c.i. y, además, $v(\mathcal{P}) \in \mathbf{R}$.

Teorema 3.4.7 (Fritz-John Fuerte)**.** *Sea $(\mathcal{P})$ un programa convexo, propio , s.c.i. y finito tal que se satisfacen las hipótesis del Teorema de Fritz John débil. Entonces una condición necesaria y suficiente para que $\hat{x} \in S(\mathcal{P})$ es que existan $\hat{\lambda}_0, \ldots, \hat{\lambda}_p \geq 0$, $\sum_{i=1}^p \hat{\lambda}_i = 1$ y reales $\hat{\mu}_1, \ldots, \hat{\mu}_q$ tales que*

$$0 \in \partial\left(\hat{\lambda_0} f + \sum_{i=1}^p \hat{\lambda}_i g_i + \delta_D\right)(\hat{x}) + \sum_{i=1}^q \mu_j \partial h_j(\hat{x})$$

y $\hat{\lambda}_i g_i(\hat{x}) = 0, i = 1, \ldots, p$.

Demostración: Sea $F(x) = \max_{1\leq i\leq p}\{f(x)-v(\mathcal{P}), g_i(x)\}$. El programa convexo $(\mathcal{P})$ puede escribirse de manera equivalente como

$$(\mathcal{P}) \quad \inf\{F + \delta_H\}.$$

Tenemos que $\hat{x} \in S(\mathcal{P})$ si, y sólo si, $0 \in \partial(F + \delta_H)(\hat{x})$. Si $x_0 \in D$ es punto de continuidad de F tal que $x_0 \in H$, tenemos que $\partial(F + \delta_H)(\hat{x}) = \partial F(\hat{x}) + \partial\delta_H(\hat{x})$. Más generalmente, se prueba que la condición $0 \in \operatorname{int}(D - H)$ permite tener la misma igualdad.
La proposición anterior nos permite concluir que

$$\partial F(\hat{x}) = \bigcup_{\lambda\in\Lambda_F(\hat{x})} \left(\lambda(f - v(\mathcal{P})) + \sum_{i=1}^{p} \lambda_i g_i + \delta_D\right)(\hat{x}).$$

Si $h_j(x) = \langle x_j^*, x - \hat{x}\rangle$, entonces $\partial\delta_h(\hat{x}) = \{x_j^*\}$ de donde se sigue que

$$\partial\delta_H(\hat{x}) = \bigcup_{\mu\in\mathbf{R}^q} \left(\sum_{j=1}^{q} \mu_j \partial h_j(\hat{x})\right)$$

de donde se obtiene el resultado □

Teorema 3.4.8 (Kuhn-Tucker Fuerte). *Supongamos que se satisfacen las hipótesis de Fritz-John Fuerte y que además se cumple la condicion de calificación de Slater $\mathcal{S}$, entonces una condición necesaria y suficiente para que $\hat{x} \in S(\mathcal{P})$ es que existan $\hat{\lambda}_1, \ldots, \hat{\lambda}_p \geq 0$ y reales $\hat{\mu}_1, \ldots, \hat{\mu}_q \in \mathbf{R}$ tales que*

$$0 \in \partial\left(f + \sum_{i=1}^{p} \hat{\lambda}_i g_i + \delta_D\right)(\hat{x}) + \sum_{i=1}^{q} \mu_j \partial h_j(\hat{x})$$

y $\hat{\lambda}_i g_i(\hat{x}) = 0, i = 1, \ldots, p$. En particular, si $\hat{x}$ es un punto de continuidad común a f y $g_i, i = 1, \ldots, p$, entonces $\hat{x} \in S(\mathcal{P})$ si, y sólo si,

$$0 \in \partial f(\hat{x}) + \sum_{i=1}^{p} \hat{\lambda}_i \partial g_i(\hat{x}) + \sum_{i=1}^{q} \mu_j \partial h_j(\hat{x}))$$

y $\hat{\lambda}_i g_i(\hat{x}) = 0, i = 1, \ldots, p$.

Demostración: Basta observar que bajo $\mathcal{S}$ podemos obtener que $\hat{\lambda}_0 > 0$ y, luego, podemos normalizar. □

3.5. Problemas

Problema 11 (Regularización de problemas min-max). Considere el problema min-max

$$(P) \qquad \min_{x\in\mathbf{R}^n} \max\{f_1(x), \ldots, f_m(x)\}$$

con $f_i : \mathbf{R}^n \to \mathbf{R}$ convexa y $S(P)$ no vacío y acotado. Para $r > 0$ considere el problema regularizado

$$(P_r) \qquad \min_{x\in\mathbf{R}^n} M_r(f_1(x), \ldots, f_m(x)).$$

donde $M_r(y) = rM(y/r)$ con $M(y_1, \ldots, y_m) = \min_{v\in\mathbf{R}}[v + \sum_{i=1}^{m} \theta(y_i - v)]$ y $\theta : \mathbf{R} \to [0, \infty)$ es una función estrictamente convexa, creciente, diferenciable y tal que $\theta_\infty(1) = +\infty$.

(a) Muestre que M es convexa y diferenciable.

(b) Pruebe que $-r\,\theta^*(1) \leq M_r(y) - \max y_i \leq -r\,m\,\theta^*(1/m)$.

(c) Deduzca que $M_r(f_1(x), \ldots, f_m(x))$ converge a $\max\{f_1(x), \ldots, f_m(x)\}$.

(d) Muestre que (P_r) admite al menos una solución x_r la cual permanece acotada cuando r tiende hacia 0. Concluir que $v(r) \to v$ y que todo punto de acumulación de x_r es solución de (P).

(e) Explicite el dual (D_r) de (P_r) asociado a la función de perturbación $\varphi(x, y) = M_r(f_1(x) + y_1, \ldots, f_m(x) + y_m)$.

(f) Pruebe que $v(P_r) + v(D_r) = 0$ y que (D_r) admite una única solución.

Problema 12 (Dualidad en problemas de aproximación cuadrática en espacios de Hilbert)**.** Sean X e Y dos espacios de Hilbert reales, con producto interno denotado indistintamente por $\langle \cdot, \cdot \rangle$. Dado un conjunto convexo y no vacío $C \subset X$ y un operador lineal, continuo y sobreyectivo $A : X \to Y$, considere el problema de minimización

$$(P) \qquad \min_{x \in C} \tfrac{1}{2}\|Ax\|^2$$

(a) Pruebe que si $C + \ker A$ es cerrado entonces el conjunto de soluciones óptimas $S(P)$ es no vacío.

(b) Pruebe que $\bar{x} \in C$ es solución óptima de (P) si, y sólo si, $\bar{y} := A^* A \bar{x}$ satisface la inecuación variacional: $\langle \bar{x} - x, \bar{y} \rangle \leq 0$, $\forall x \in C$.

(c) Suponga que $C = \{x \in X \mid \langle a^i, x \rangle \leq b_i,\ i = 1, \ldots, m\}$. Verifique que (P) admite solución óptima.

(d) Asuma que existe $x_0 \in X$ tal que $\langle a^i, x_0 \rangle < b_i$, $i = 1, \ldots, m$. Pruebe que $\bar{x} \in S(P)$ si, y sólo si, $\exists \lambda \in \mathbf{R}^m_+$ tal que $A^* A \bar{x} = \sum_{i=1}^m \lambda_i a^i$ y además $\lambda_i(\langle a^i, x \rangle - b_i) = 0$, $i = 1, \ldots m$.

Problema 13 (Aproximación de máxima entropía)**.** Sea $\bar{u} \in L^1 \equiv L^1([0, 1], \mathbf{R})$ una función desconocida tal que $\bar{u}(x) \geq 0$ c.t.p. en $[0, 1]$. Deseamos estimar $\bar{u}$ en base a sus $(n + 1)$ primeros momentos $\int_0^1 x^i \bar{u}(x)dx = m_i > 0$, $i = 0, 1, \ldots, n$. Para ello consideramos el problema

$$(P) \qquad \min_{u \in L^1} \left\{ \int_0^1 E(u(x))dx : \int_0^1 x^i u(x)dx = m_i,\ i = 0, \ldots, n \right\}$$

donde $E : \mathbf{R} \to \mathbf{R} \cup \{\infty\}$ es (menos) la entropía de Boltzmann-Shannon definida por $E(u) = u \ln u$ para $u \geq 0$ y $E(u) = \infty$ para $u < 0$. El problema (P) equivale a $\min_{u \in L^1}\{\Phi(u) : Au = m\}$ con $\Phi : L^1 \to \mathbf{R} \cup \{+\infty\}$ y $A : L^1 \to \mathbf{R}^{n+1}$ definidas por

$$\begin{aligned} \Phi(u) &= \begin{cases} \int_0^1 E(u(x))dx & \textit{si } E \circ u \in L^1 \\ +\infty & \textit{si no,} \end{cases} \\ (Au)_i &= \int_0^1 x^i u(x)dx, \quad i = 0, 1, \ldots, n. \end{aligned}$$

(a) Pruebe que $\Phi \in \Gamma_0(H)$. Indicación: utilizar el Lema de Fatou.

(b) Explicite el dual que se obtiene al perturbar la restricción de (P) mediante $Au = m - y$. Para ello pruebe que $A^* : \mathbf{R}^{n+1} \to L^\infty$ está dado por $(A^*\lambda)(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_n x^n$ mientras que para todo $u^* \in L^\infty$ se tiene $\Phi^*(u^*) = \int_0^1 \exp(u^*(x) - 1)dx$.

(c) Usando el teorema de dualidad pruebe que (P) tiene una única solución.

(d) Pruebe que si λ es una solución dual entonces la solución de (P) es $u(x) = \exp(-1 - \sum_{i=1}^n \lambda_i x^i)$.

Problema 14 (Un problema de Ingeniería)**.** Sea $(X, \|\cdot\|)$ un e.v.n. y $(X^*, \|\cdot\|_*)$ su dual topológico. Sean $x_1, \ldots, x_n \in Y$ y $a_1, \ldots, a_n \in \mathbf{R}$ tales que el sistema $\langle x^*, x_i \rangle = a_i$, $i = 1, \ldots, m$ admite al menos una solución $x^* \in X^*$. Consideremos el problema de encontrar una solución de norma mínima.

(a) Pruebe la identidad

$$\min_{x^* \in X^*} \{\|x^*\|_* : \langle x^*, x_i \rangle = a_i, i = 1 \ldots n\} = \max_{y^* \in \mathbf{R}^n} \{\textstyle\sum_{i=1}^n a_i y_i^* : \|\sum_{i=1}^n y_i^* x_i\| \leq 1\}$$

y que ambos óptimos son alcanzados. Para ello considere la función de perturbación $\varphi : X^* \times \mathbf{R}^n \to \overline{\mathbf{R}}$

$$\varphi(x^*, y) = \begin{cases} \|x^*\|_* & \text{si } \langle x^*, x_i \rangle = a_i + y_i, i = 1 \ldots n \\ +\infty & \text{en caso contrario.} \end{cases}$$

(b) Pruebe que $x^* \in X^*$ e $y^* \in \mathbf{R}^n$ son soluciones de los problemas anteriores si, y sólo si, $\langle x^*, x_i \rangle = a_i$, $\|\sum y_i^* x_i\| \leq 1$, y se tiene $\langle x^*, \sum_{i=1}^n y_i^* x_i \rangle = \|x^*\|_* \|\sum_{i=1}^n y_i^* x_i\|$.

(c) La velocidad angular ω de un motor eléctrico alimentado por una corriente variable $u(t)$ se rige por la ecuación $\dot{\omega}(t) + \omega(t) = u(t)$. Se requiere llevar el motor desde la posición inicial $\theta(0) = \omega(0) = 0$, hasta $\theta(1) = \omega(1) = 1$ (θ=posición angular, $\omega = \dot{\theta}$), minimizando la norma infinito de $u(t)$.

(c.1) Verifique que el problema puede plantearse como

$$\min_{u \in L^\infty[0,1]} \left\{ \|u\|_\infty : \int_0^1 e^{t-1} u(t)\, dt = 1, \int_0^1 (1 - e^{t-1}) u(t)\, dt = 1 \right\}.$$

(c.2) Deduzca que el control óptimo es de la forma $u(t) = M \cdot \operatorname{sgn}(x_1 e^{t-1} + x_2[1 - e^{t-1}])$ para constantes adecuadas x_1, x_2, M. Probar que $u(t)$ tiene a lo más un cambio de signo y calcular $u(t)$.

Problema 15 (Dualidad de Toland-Singer)**.** En lo que sigue, denotamos por $\dot{+}$ la inf-adición en $\overline{\mathbf{R}}$ (en particular, $\pm\infty \dot{+} (\mp\infty) = +\infty$) y por $\dot{-}$ la inf-sustracción definida por $\alpha \dot{-} \beta := \alpha \dot{+} (-\beta)$. La sup-adición $\underset{\cdot}{+}$ y la sup-sustracción $\underset{\cdot}{-}$ se definen de manera análoga. Dadas dos funciones $f, g : X \to \mathbf{R} \cup \{+\infty\}$, donde $(X, Y, \langle \cdot, \cdot \rangle)$ es una dualidad entre e.v.t.l.c., se define $h := f \dot{-} g$.

(a) Pruebe que se satisface la siguiente relación para las conjugadas de Legendre-Fenchel:

$$\forall y \in Y,\ h^*(y) \geq \sup_{z \in \operatorname{dom} g^*} \{f^*(y + z) - g^*(z)\} \tag{3.1}$$

(b) Suponga que $g \in \Gamma_0(X)$. Pruebe que entonces se tiene igualdad en (3.1), y deduzca la fórmula de dualidad de Toland-Singer:

$$\inf_X \{f \dot{-} g\} = \inf_Y \{g^* \dot{-} f^*\}.$$

(c) Suponga ahora que $(X, \|\cdot\|)$ es un e.v.n. en dualidad con $Y = X^*$. Dados dos convexos cerrados $C, D \subset X$, se define el *exceso de Hausdorff* de C sobre D mediante $e(C,D) := \sup_{x \in C} d(x,D) \in [0,+\infty]$ con $d(x,D) = \inf_{y \in D} \|y - x\|$. La *distancia de Hausdorff* entre C y D se define como $h(C,D) = \max\{e(C,D), e(D,C)\} \in [0,+\infty]$. Usando (b), pruebe que

$$e(C,D) = \sup_{x^* \in B^*} \{\sigma_C(x^*) \dot{-} \sigma_D(x^*)\},$$

donde $B^* := \{x^* \in X^* \mid \|x^*\|_* \leq 1\}$. Ind.: sea $g(x) = d(x,D)$, pruebe que $g^* = \sigma_D + \delta_{B^*}$. Si además C y D son acotados, muestre que se tiene la *fórmula de Hőrmander*:

$$h(C,D) = \sup_{x^* \in B^*} |\sigma_C(x^*) - \sigma_D(x^*)|.$$

Problema 16 (Un caso particular de la dualidad de Clarke-Ekeland)**.** Sea $(H, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert y $f \in \Gamma_0(H)$. Sea $A : H \to H$ un operador lineal, continuo y autoadjunto. Se definen las funciones $\Phi, \Psi : H \to \mathbf{R} \cup \{+\infty\}$ mediante $\Phi(x) = \frac{1}{2}\langle x, Ax \rangle + f(x)$ y $\Psi(x) = \frac{1}{2}\langle x, Ax \rangle + f^*(-Ax)$. Note que Φ y Ψ no son necesariamente convexas.

(a) Pruebe que si $\bar{x} \in H$ es un *punto crítico* para Φ, lo que por definición significa que $-A\bar{x} \in \partial f(\bar{x})$, entonces todo $\bar{y} \in \ker A + \bar{x}$ es un punto crítico para Ψ, es decir $-A\bar{y} \in \partial[f^* \circ (-A)](\bar{y})$.

(b) Demuestre que si $0 \in \mathrm{int}(\mathrm{dom}\, f^* - \mathrm{Im} A)$ entonces para todo punto crítico $\bar{y}$ de Ψ existe $\bar{x} \in \ker A + \bar{y}$ que es punto crítico de Φ, y más aun $\Phi(\bar{x}) + \Psi(\bar{y}) = 0$.

Problema 17. Sea S_n el espacio de Hilbert formado por las matrices simétricas de orden n, con el producto interno $\langle\!\langle A, B \rangle\!\rangle := \mathrm{tr}(AB) = \sum_{i,j=1}^n A_{ij} B_{ij}$. Sea S_n^+ el conjunto de matrices simétricas definidas positivas y sea $\Theta : S_n \to \mathbf{R} \cup \{+\infty\}$ definida por

$$\Theta(A) = \begin{cases} \ln[\det(A^{-1})] & \text{si } A \in S_n^+ \\ +\infty & \text{en caso contrario.} \end{cases}$$

Notemos que $\mathrm{dom}(\Theta) = S_n^+$ es abierto en S_n, $\Theta(\cdot)$ es $\mathcal{C}^\infty$ sobre S_n^+ y continua en todo S_n.

(a) Pruebe que $\Theta(A) + \mathrm{tr}(A) \geq n$ para todo $A \in S_n$, con igualdad sólo en el caso en que todos los valores propios de A son iguales a 1. Indicación: considerar el problema $\min\{\sum_{i=1}^n \lambda_i - \ln \lambda_i : \lambda_i > 0\}$.

(b) Deduzca que $\Theta(A + H) > \Theta(A) - \mathrm{tr}(A^{-1}H)$ para todo $A \in S_n^+, H \in S_n \setminus \{0\}$. Concluya que $\Theta \in \Gamma_0(S_n)$, con $\Theta(\cdot)$ estrictamente convexa en S_n^+ y $\nabla\Theta(A) = -A^{-1}$ para $A \in S_n^+$.

(c) Demuestre (si lo desea comience por el caso $n = 1$) que

$$\Theta^*(A^*) = \begin{cases} \Theta(-A^*) - n & \text{si } -A^* \in S_n^+ \\ +\infty & \text{si no.} \end{cases}$$

Salvo constante multiplicativa, el número $\sqrt{\det(A^{-1})}$ es el volumen del elipsoide $\{x \in \mathbf{R}^n : x^t Ax \leq 1\}$. Así, el problema de encontrar el elipsoide de volumen mínimo (centrado en el origen) que contiene un conjunto de puntos dados $\{x_1, \ldots, x_k\} \subset \mathbf{R}^n$ se puede formular como

$$(P) \qquad \min_{A \in S_n} \{\Theta(A) : x_i^t A x_i \leq 1, i = 1 \ldots k\}.$$

Notemos que $x_i^t A x_i = \langle\langle A, A_i \rangle\rangle$ con $A_i := x_i x_i^t$, de modo que (P) es un problema con restricciones de desigualdad lineales. Para evitar soluciones degeneradas supondremos que $\{x_1, \ldots, x_k\}$ genera todo $\mathbf{R}^n$. Perturbemos (P) mediante $\varphi : S_n \times \mathbf{R}^k \to \mathbf{R} \cup \{\infty\}$

$$\varphi(A, y) = \begin{cases} \Theta(A) & \text{si } \langle\langle A, A_i \rangle\rangle \leq 1 - y_i, i = 1 \ldots k \\ +\infty & \text{en caso contrario.} \end{cases}$$

(d) Explicite el dual (D) mostrando que tiene solución. Pruebe asimismo que (P) posee solución única y que si $y^* = (\lambda_1, \ldots, \lambda_k)$ es una solución del dual, entonces la solución de (P) viene dada por
$$A = \left[\sum_{i=1}^k \lambda_i x_i x_i^\tau\right]^{-1}.$$

Problema 18 (Fórmula de Lax-Oleinik)**.** Sean $f, \theta \in \Gamma_0(\mathbf{R}^N)$ con $\theta(\cdot)$ finita, estrictamente convexa y diferenciable en todo $\mathbf{R}^N$, tal que $\lim_{\|y\| \to \infty} \theta(y)/\|y\| = \infty$. Consideremos la ecuación de Hamilton-Jacobi

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) + \theta^*(\nabla_x u(t,x)) &= 0 \qquad && x \in \mathbf{R}^N, t > 0, \\ u(0,x) &= f(x) && x \in \mathbf{R}^N. \end{aligned}$$

Se desea probar que
$$u(t,x) = \min_{y \in \mathbf{R}^n} f(x-y) + t\theta(y/t)$$
es una solución de tal ecuación. Para ello se propone el siguiente esquema (a) _ (f)

(a) Sea $\varphi(t,x,y) = f(x-y) + t\theta(y/t)$. Pruebe que φ es convexa y deduzca que u es convexa y continua.

(b) Pruebe que el óptimo en la definición de $u(t,x)$ es alcanzado en un único punto $\bar{y} = \bar{y}(t,x)$, y que se tiene
$$(t^*, x^*) \in \partial u(t,x) \Longleftrightarrow (t^*, x^*, 0) \in \partial\varphi(t,x,\bar{y}).$$

(c) Calcule φ^* en términos de f^* y θ^*.

(d) Deduzca que $(t^*, x^*) \in \partial u(t,x)$ si, y sólo si, se tienen las condiciones siguientes:

   (i) $t^* + \theta^*(x^*) = 0$;
   (ii) $x^* \in \partial f(x - \bar{y})$; y
   (iii) $x^* = \nabla\theta(\bar{y}/t)$.

(e) Concluya que $u(\cdot,\cdot)$ es diferenciable en todo punto $(t,x) \in (0,\infty) \times \mathbf{R}^N$ y satisface $\frac{\partial u}{\partial t}(t,x) + \theta^*(\nabla_x u(t,x)) = 0$.

(f) Pruebe que para todo $x \in \mathbf{R}^N$ se tiene $\lim_{t\to 0^+} u(t,x) = f(x)$.

Finalmente, considere la ecuación

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) + \frac{1}{2}\|\nabla_x u(t,x)\|^2 &= 0 \qquad && x \in \mathbf{R}^N, t > 0, \\ u(0,x) &= f(x) && x \in \mathbf{R}^N. \end{aligned}$$

(g) Encuentre la fórmula explícita para la solución de la ecuación.

Problema 19 (Subgradiente aproximado)**.** Sea $f \in \Gamma_0(X)$ con $(X, \|\cdot\|)$ espacio de Banach. Para $\epsilon \geq 0$ definimos el ϵ-subdiferencial de f en $u \in X$ como el conjunto $\partial_\epsilon f(u)$ formado por los funcionales $u^* \in X^*$ que satisfacen

$$f(u) - \epsilon + \langle u^*, v - u\rangle \leq f(v) \qquad \text{para todo } v \in X.$$

(a) Pruebe que $\partial_\epsilon f(u)$ es convexo, cerrado y *no-vacío*.

(b) Sea $u^* \in \partial_\epsilon f(u)$. Pruebe que existe $v \in X$ y $v^* \in \partial f(v)$ tal que $\|v-u\| \leq \sqrt{\epsilon}$ y $\|v^*-u^*\| \leq \sqrt{\epsilon}$.

(c) Deduzca que $\mathrm{dom}(\partial f)$ es denso en $\mathrm{dom}(f)$.

Suponga que $m = \inf_{u\in X} f(u) > -\infty$ y sea $\epsilon > 0$.

(d) Pruebe que u es un ϵ-mínimo (i.e. $f(u) \leq m + \epsilon$) si, y sólo si, $0 \in \partial_\epsilon f(u)$.

(e) Deduzca la existencia de una sucesión $v_k \in X$ y $v_k^* \in \partial f(v_k)$ tal que $f(v_k) \to m$ y $\|v_k^*\|_* \to 0$.

(f) Se dice que f satisface la condición de Palais-Smale si toda sucesión v_k que satisface $d(0, \partial f(v_k)) \to 0$ con $f(v_k)$ acotada, es relativamente compacta para la topología débil. Deduzca que si f es acotada inferiormente y satisface la condición de Palais-Smale, entonces el mínimo de f es alcanzado.

Problema 20. Sea K un espacio topológico compacto y $\mathcal{C}(K)$ el conjunto de funciones continuas de K en $\mathbf{R}$ con la norma de la convergencia uniforme $\|f\|_\infty = \max_{t\in K} |f(t)|$. Recordemos que el dual de $\mathcal{C}(K)$ es isométricamente isomorfo al conjunto $\mathcal{M}(K)$ de medidas de Borel regulares $\mu : \mathcal{B}(K) \to \mathbf{R}$ con la norma $\|\mu\| = |\mu|(K) < \infty$ donde

$$|\mu|(A) = \sup\{\textstyle\sum_{i=1}^m |\mu(A_i)| : \{A_i\}_{i=1}^m \text{ partición medible de } A\} \qquad \forall\, A \in \mathcal{B}(K).$$

El producto de dualidad entre $\mathcal{C}(K)$ y $\mathcal{M}(K)$ es $\langle \mu, f\rangle = \int_K f\, d\mu$. Notemos que $|\int_K f\, d\mu| \leq \int_K |f|\, d|\mu|$.

(a) Defina $\psi(\cdot) = \|\cdot\|_\infty$ y pruebe que $\psi^* = \delta_{\bar{B}(0,1)}$. Encuentre $\partial\psi(0)$.

(b) Para $f \neq 0$ denotamos $K_f^+ = \{t \in K : f(t) = \|f\|_\infty\}$ y $K_f^- = \{t \in K : f(t) = -\|f\|_\infty\}$. Pruebe que $\mu \in \partial\|\cdot\|_\infty(f)$ si, y sólo si, $\|\mu\| = 1$, $\mathrm{supp}(\mu) \subset K_f^+ \cup K_f^-$, $\mu \geq 0$ en K_f^+ y $\mu \leq 0$ en K_f^-.

(c) Sean $f, \psi_1, \ldots, \psi_n \in \mathcal{C}(K)$ y consideremos el problema de mejor aproximación en el sentido de Tchebychef

$$(P) \qquad \min_{x \in \mathbf{R}^n} \|f - \textstyle\sum_{i=1}^n x_i \psi_i\|_\infty.$$

Calcule el dual de (P) asociado a las perturbaciones $\varphi(x, y) = \|f + y - \sum_{i=1}^n x_i \psi_i\|_\infty$ para $x \in \mathbf{R}^n$ e $y \in \mathcal{C}(K)$.

(d) Pruebe que $v(P) + v(D) = 0$ y que $S(D) \neq \phi$.

(e) Suponiendo que $\psi_1, \ldots, \psi_n$ son linealmente independientes demuestre que el operador $A : \mathcal{M}(K) \to \mathbf{R}^n$ definido por $A(\mu) = (\int_K \psi_1 \, d\mu, \ldots, \int_K \psi_n \, d\mu)$ es sobreyectivo. Deduzca que en tal caso (P) admite soluciones.

(f) Explicite las relaciones de extremalidad que caracterizan las soluciones óptimas $\bar{x} \in S(P)$ y $\bar{\mu} \in S(D)$.

Problema 21 (Regularizada de Moreau-Yosida)**.** Sea H un espacio de Hilbert y $f : H \to \mathbf{R} \cup \{+\infty\}$ una función convexa, s.c.i. y propia. La *Regularizada Moreau-Yoshida* de parámetro $\lambda > 0$ es la función $f_\lambda : H \to \mathbf{R}$ definida mediante

$$f_\lambda(z) = \inf_{x \in H} \left\{ f(x) + \frac{1}{2\lambda} \|z - x\|^2 \right\}.$$

(a) Pruebe que f_λ es convexa, finita, y que el ínfimo es alcanzado en un único punto.

(b) Deduzca que para todo $z \in H$ existe un único $x \in H$ tal que $z \in x + \lambda \partial f(x)$.

(c) Pruebe que f_λ es Gâteaux-diferenciable con $\nabla f_\lambda = \frac{1}{\lambda}(I - J_\lambda)$.

La propiedad (b) permite definir la aplicación $J_\lambda = (I + \lambda \partial f)^{-1} : H \to H$.

(d) Pruebe que ∂f es *monótono*: $x^* \in \partial f(x), y^* \in \partial f(y) \Longrightarrow \langle y^* - x^*, y - x \rangle \geq 0$.

(e) Deduzca que J_λ es una contracción: $\|J_\lambda(z_1) - J_\lambda(z_2)\| \leq \|z_1 - z_2\|$.

(f) A partir de la desigualdad $f(J_\lambda(z)) + \frac{1}{2\lambda}\|J_\lambda(z) - z\|^2 \leq f(z)$, pruebe que para $\lambda \downarrow 0$ se tiene

- (f.1) $\|J_\lambda(z)\|$ se mantiene acotado,
- (f.2) $\|J_\lambda(z) - z\|$ tiende a cero,
- (f.3) $f_\lambda(z)$ converge monótonamente a $f(z)$.

J_λ se llama *resolvente* y $\frac{1}{\lambda}(I - J_\lambda)$ es la *regularizada Yoshida* de ∂f. Usando (c) y (e) se sigue que f_λ es Fréchet-diferenciable con gradiente Lipschitziano.

Problema 22. (Karush-Kuhn-Tucker en el caso diferenciable no convexo)

(a) Consideremos una función $f : \mathbf{R}^n \to \mathbf{R}$ una función diferenciable y acotada inferiormente. Pruebe que para cada $\varepsilon > 0$ existe $x_\varepsilon \in \mathbf{R}^n$ tal que $\|\nabla f(x_\varepsilon)\| \leq \varepsilon$. Indicación: considere x_ε un mínimo global de la función $f + \varepsilon\| \cdot \|$ y suponga que $\|\nabla f(x_\varepsilon)\| > \varepsilon$, pruebe entonces que en ese caso $f'(x_\varepsilon; d_\varepsilon) < -\varepsilon\|d_\varepsilon\|$ con $d_\varepsilon := -\nabla f(x_\varepsilon)$ y que esto contradice la optimalidad de x_ε.

(b) Demuestre el *Teorema de la Alternativa de Gordan*: dados $a^0, a^1, ..., a^m \in \mathbf{R}^n$, uno y sólo uno de los siguientes sistemas (S1) y (S2) admite una solución.

(S1) $\sum_{i=0}^m \lambda_i a^i = 0$, $\sum_{i=0}^m \lambda_i = 1$, $0 \leq \lambda_0, \lambda_1, ..., \lambda_m$.

(S2) $\langle a^i, d \rangle < 0$ para todo $i = 0, 1, ..., m$, $d \in \mathbf{R}^n$.

Indicación: pruebe que si la función $f(x) = \log\left(\sum_{i=0}^m \exp\langle a^i, x\rangle\right)$ es acotada inferiormente entonces (S1) tiene una solución.

(c) Sean $g_0, g_1, ..., g_m : C \subset \mathbf{R}^n \to \mathbf{R}$ funciones continuas en C y diferenciables en $\bar{x} \in int(C)$. Se definen $g(x) := \max\limits_{0 \leq i \leq m} \{g_i(x)\}$ y $J := \{i \mid g_i(\bar{x}) = g(\bar{x})\}$. Pruebe que $\forall d \in \mathbf{R}^n$:

(c.1) $\liminf\limits_{t \to 0^+} \frac{1}{t}\left[g(\bar{x} + td) - g(\bar{x})\right] \geq \langle \nabla g_i(\bar{x}), d\rangle$, $\forall i \in J$.

(c.2) $\limsup\limits_{t \to 0^+} \frac{1}{t}\left[g(\bar{x} + td) - g(\bar{x})\right] \leq \max\limits_{i \in J}\{\langle \nabla g_i(\bar{x}), d\rangle\}$.
Indicación: razone por contradicción probando que si lo anterior no fuese cierto entonces existirían $\varepsilon > 0$ y $j \in J$ tales que para una sucesión $t_k \to 0^+$ se tendría que $\frac{1}{t_k}(g(\bar{x} + t_k d) - g(\bar{x})) \geq \max\limits_{i \in J}\{\langle \nabla g_i(\bar{x}), d\rangle\} + \varepsilon$.

(c.3) $g'(\bar{x}; d) = \max\limits_{i \in J}\{\langle \nabla g_i(\bar{x}), d\rangle\}$.

(d) Considere el siguiente problema de minimización con restricciones:

$$(P) \qquad \min\{f(x) \mid g_i(x) \leq 0 \text{ para } i = 1, 2, ..., m,\ x \in C\},$$

donde $f, g_1, ..., g_m : C \subset \mathbf{R}^n \to \mathbf{R}$ son funciones continuas. Suponga que (P) tiene un minimizador local $\bar{x} \in int(C)$ y que f, g_i $(i \in I(\bar{x}) := \{i \mid g_i(\bar{x}) = 0\})$ son diferenciables en $\bar{x}$. Pruebe que:

(d.1) (condiciones de Fritz-John) existen $\lambda_0, \lambda_i \in \mathbf{R}_+$ $(i \in I(\bar{x}))$, no todos nulos, que satisfacen

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0. \tag{3.2}$$

Indicación: verifique que $g'(\bar{x}; d) \geq 0$ para una función g escogida apropiadamente y luego aplique el Teorema de la Alternativa de Gordan.

(d.2) (condiciones de Karush-Kuhn-Tucker) si además se satisface la condición de calificación de Fromovitz-Mangasarian :

$$(MF) \qquad \exists d \in \mathbf{R}^n : \forall i \in I(\bar{x}), \langle \nabla g_i(\bar{x}), d\rangle < 0,$$

entonces es posible tomar $\lambda_0 = 1$ en (3.2).

Capítulo 4

Aplicaciones al Cálculo de Variaciones

En esta sección revisaremos algunos ejemplos del Cálculo de Variaciones y del Análisis Variacional. Las técnicas usadas son aplicaciones de la Dualidad Convexa, revisada en la sección anterior, a los espacios de Sobolev. Es recomendable tener frescos algunos resultados de Teoría de las Distribuciones y de espacios de Sobolev. Se recomienda revisar [Bre83].

4.1. Problema de Dirichlet

Sea $\Omega \subset \mathbf{R}^N$ un abierto acotado y consideremos los espacios $X := H_0^1(\Omega)$, $Y = L^2(\Omega)^N$ y sus respectivos duales topológicos $X^* = H^{-1}(\Omega)$, $Y^* = L^2(\Omega)^N$ (identificación). Nos interesa el problema

$$(\mathcal{P}) \inf_{u \in H_0^1(\Omega)} \left\{ \frac{1}{2} \int_\Omega |\nabla u(x)|^2 dx - \int_\Omega f(x)u(x)dx \right\},$$

donde ∇ denota el gradiente débil y $f \in L^2(\Omega)$. Sean $F\colon H_0^1(\Omega) \to \mathbf{R}$ definida por $F(u) = -\int_\Omega f(x)u(x)dx$ y $G\colon L^2(\Omega)^N \to \mathbf{R}$ con $G(p) = \frac{1}{2}\|p\|^2_{L^2(\Omega)^N} = \frac{1}{2}\int_\Omega \sum_{i=1}^N p_i^2(x)dx$, de modo que podemos escribir $(\mathcal{P})$ de manera equivalente por

$$\inf_{u \in X} F(u) + G(Au)$$

donde $A\colon H_0^1(\Omega) \to L^2(\Omega)^N$ está definida por $Au = \nabla u$ que resulta ser lineal y continua. Es fácil ver que

$$F^*(u^*) = \begin{cases} 0 & \text{si } u^* = -f, \\ +\infty & \text{si no.} \end{cases}$$

y que $G^*(p^*) = \frac{1}{2}\|p^*\|^2_{L^2(\Omega)^N}$. Podemos carecterizar $A^*\colon L^2(\Omega)^N \to H^{-1}(\Omega)$ por

$$\forall u \in H_0^1(\Omega), \quad \langle A^*p^*, u\rangle_{H^{-1}(\Omega),H_0^1(\Omega)} = \langle p^*, Au\rangle_{L^2(\Omega)^N,L^2(\Omega)^N} = -\langle \operatorname{div} p^*, u\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}.$$

Así $A^*p^* = -\operatorname{div} p^*$ (en el sentido de las distribuciones). Luego, el problema dual de $(\mathcal{P})$ es

$$(\mathcal{D}) \quad \inf \left\{ \frac{1}{2}\|p^*\|^2_{L^2(\Omega)^N} \;\middle|\; p^* \in L^2(\Omega)^N, \operatorname{div} p^* = -f \right\}.$$

La calificación primal es $0 \in \operatorname{int}(\operatorname{dom}(G) - A\operatorname{dom}(F))$ que es equivalente a $0 \in \operatorname{int}(L^2(\Omega)^N) = L^2(\Omega)^N$. Luego se satisface la calificación primal y como $\inf(\mathcal{P}) \leq 0$ y $F(v) + G(Av)$ es coerciva (el lector debe verificar esto) se tiene que $S(\mathcal{D}) \neq \emptyset$ y, más aun, $(\mathcal{D})$ tiene solución única. Sea $\bar{u}$ solución de $(\mathcal{P})$ y $\bar{p}^*$ la solución de $(\mathcal{D})$. Las ecuaciones de Euler-Lagrange son

$$\operatorname{div}(\bar{p}^*) \in \partial F(\bar{u}) = \{-f\}$$

$$\bar{p}^* \in \partial G(\nabla \bar{u}) = \{\nabla \bar{u}\},$$

y equivalen a

$$\begin{aligned} \operatorname{div}(\bar{p}^*) &= -f \\ \bar{p}^* &= \nabla \bar{u}. \end{aligned}$$

Luego $\bar{u}$ es solución (débil) de

$$\begin{aligned} -\Delta \bar{u} &= f \text{ en } \Omega, \\ u &= 0 \text{ sobre } \partial\Omega. \end{aligned}$$

4.2. Problema de Stokes

Sean $\Omega \subset \mathbf{R}$ abierto acotado y $f \in L^2(\Omega)^3$. El problema de Stokes consiste en encontrar un campo de velocidades $u = (u_1, u_2, u_3) \in H^1(\Omega)^3$ y una función de presión $p \in L^2(\Omega)$ tales que

$$(S) \quad \left\{ \begin{array}{ccll} -\Delta u + \nabla p & = & f & \text{en } \Omega, \\ \operatorname{div}(u) & = & 0 & \text{en } \Omega, \\ u & = & 0 & \text{sobre } \partial\Omega, \end{array} \right.$$

donde $\Delta u = (\Delta u_1, \Delta u_2, \Delta u_3)$. Este problema corresponde a un modelo simplificado de un fluido homogéneo e incompresible con $\Omega \subset \mathbf{R}^3$ la región que contiene al fluido, que supondremos de frontera $\partial\Omega$ regular.

Veamos la formulación variacional del problema. Dada $\varphi = (\varphi_1, \varphi_2, \varphi_3) \in D(\Omega)^3$ (donde $D(\Omega)$ denota el conjunto de funciones de clase C^∞ sobre Ω, a valores reales y con soporte compacto). Notemos que si (u, p) son soluciones clásicas, entonces

$$\sum_{i=1}^{3} \int_\Omega \nabla u_i \nabla \varphi - \int_\Omega p \operatorname{div}(\varphi) = \int_\Omega f v$$

y definimos una solucion débil de (S) como un par $(u, p) \in H^1(\Omega) \times L^2(\Omega)$ tal que

$$(*) \quad \forall v \in H_0^1(\Omega)^3, \quad \sum_{i=1}^{3} \int_\Omega \nabla u_i \nabla v_i - \int_\Omega p \operatorname{div}(v) = \int_\Omega f v.$$

Consideremos el espacio $V = \{v \in H_0^1(\Omega)^3 \mid \operatorname{div}(v) = 0\}$ dotado del producto interno $((u, v)) := \sum_{i=1}^{3} \int_\Omega \nabla u_i \nabla v_i$ y la norma $\|\cdot\|$ inducida por tal producto interno. Es fácil ver que $(V, \|\cdot\|)$ es un espacio de Hilbert. Además, $u \in V$ es solución de

$$(\mathcal{P}) \quad \inf_{v \in V} \frac{1}{2}\|v\|^2 - \int_\Omega f v$$

si, y sólo si,

$$\forall v \in V, \quad ((u, v)) = \int_\Omega fv$$

que corresponde a $(*)$ cuando hacemos $\mathrm{div}(v) = 0$.

Pero, ¿cómo recuperamos la presión? Construiremos un problema dual asociado a $(\mathcal{P})$ que nos permitirá obtener la presión p que perdimos en el camino. Sean $X = H_0^1(\Omega)^3$, $Y = L^2(\Omega) = Y^*$ y $X^* = H^{-1}(\Omega)$. Definamos las funciones

$$\begin{array}{rcl} F : X & \longrightarrow & \mathbf{R} \\ v & \mapsto & \frac{1}{2}\|v\|^2 - \int_\Omega fv, \end{array} \qquad \begin{array}{rcl} G : Y & \longrightarrow & \mathbf{R} \\ y & \mapsto & \delta_{\{0\}}(y), \end{array} \qquad \begin{array}{rcl} A : X & \longrightarrow & \mathbf{R} \\ v & \mapsto & \mathrm{div}\, v. \end{array}$$

Así podemos escribir el problema primal como

$$(\mathcal{P}) \quad \inf_{v \in X} F(v) + G(Av).$$

El dual en el sentido de Fenchel-Rockafellar es

$$(\mathcal{D}) \quad \inf_{p^* \in Y^*} F^*(-A^*p^*) + G^*(p^*)$$

donde $G^*(p^*) = 0$ y

$$\begin{aligned} F^*(-A^*p^*) &= \sup_{v \in X} \left\{ \langle -p^*, \mathrm{div}(v) \rangle_{L^2(\Omega)} - \frac{1}{2}\|v\|^2 + \int_\Omega fv \right\} \\ &= - \inf_{v \in X} \{ \frac{1}{2}\|v\|^2 - \int_\Omega fv + \int_\Omega p^* \,\mathrm{div}(v) \}. \end{aligned}$$

Este último problema es coercivo (verifíquelo) y estrictamente convexo, luego para cada $p^* \in Y^*$ existe una única solución $v_{p^*} \in X$ que está caracterizada por:

$$\forall h = (h_1, h_2, h_3) \in X, \quad ((v_{p^*}, h)) - \langle f, h \rangle_{L^2(\Omega)} + \langle p^*, \mathrm{div}(h) \rangle_{L^2(\Omega)} = 0.$$

En particular, para cada $i = 1, 2, 3$ y $w \in H_0^1(\Omega)$ se tiene que

$$\int_\Omega \nabla(v_{p^*})_i \nabla w - \int_\Omega f_i v + \int_\Omega p^* \frac{\partial w}{\partial x_i} = 0,$$

de donde deducimos que

$$-\Delta v_{p^*} - \nabla p^* = f \text{ en } \Omega$$

(en el sentido de las distribuciones). Es directo verificar que

$$F^*(-A^*p^*) = \frac{1}{2} \sum_{i=1}^{3} \|\nabla v_{p^*}\|_{L^2(\Omega)}^2 = \frac{1}{2} \|v_{p^*}\|_{H_0^1(\Omega)^3}^2,$$

de modo que podemos reescribir el dual como

$$(\mathcal{D}) \quad \inf_{p^* \in L^2(\Omega)} \left\{ \frac{1}{2}\|v_{p^*}\|_{H_0^1(\Omega)^3}^2 \;\middle|\; v_{p^*} \in H_0^1(\Omega)^3 \text{ es solución de } -\Delta v_{p^*} - \nabla p^* = f \text{ en } \Omega \right\}.$$

Sea $\bar{u} \in X$ la solución de $(\mathcal{P})$ y supongamos que $(\mathcal{D})$ admite solución $\bar{p}^*$ (la cual no es única, ¿por qué?). Tenemos que la relación de extremalidad es

$$F(\bar{u}) + F^*(-A^*\bar{p}^*) = \langle \bar{u}, -A^*\bar{p}^* \rangle = 0,$$

lo que implica que

$$F^*(-A^*\bar{p}^*) \geq \langle v, -A^*\bar{p}^* \rangle - F(v)$$

y en consecuencia $\bar{u} = v(\bar{p}^*)$. Luego,

$$\begin{aligned} \operatorname{div}(\bar{u}) &= 0 \text{ en } \Omega, \\ -\Delta\bar{u} + \nabla(-\bar{p}^*) &= f \text{ en } \Omega, \\ u &= 0 \in \partial\Omega, \end{aligned}$$

lo que equivale a que $(\bar{u}, -\bar{p}^*)$ sea solución débil del problema de Stokes.

Para finalizar, probemos la existencia de una solución de $(\mathcal{D})$. Veamos la condición de calificación dual:

$$0 \in \operatorname{int}(\operatorname{dom}(F^*) + A^*L^2(\Omega)).$$

Pero $F^*(-A^*p^*) = \frac{1}{2}\|v_{p^*}\|^2$ y se verifica que $-A^*p^* \in \operatorname{int}(\operatorname{dom}(F^*))$ para todo $p^* \in L^2(\Omega)$, de modo que deducimos que $\inf(\mathcal{D}) = -\min(\mathcal{P}) \in \mathbf{R}$. Sea p_n^* una suceción minimizante para $(\mathcal{D})$. En particular tenemos que existe $C \geq 0$ tal que para todo $n \in \mathbf{N}$ se satisface que $\|v_{p_n^*}\|_{H_0^1(\Omega)} \leq C$, lo que implica que $(v_{p_n^*})_n$ está acotada en $H_0^1(\Omega)^3$ y por lo tanto $(\nabla p_n^*)_n$ está acotada en $H^{-1}(\Omega)$ y de aquí concluimos que $(p_n^*)_n$ está acotada en $L^2(\Omega)/\mathbf{R}$. Pasando a una subsucesión si fuese necesario, tenemos que $(p_n^*)_n$ converge débilmente hacia $\bar{p}^* \in L^2(\Omega)/\mathbf{R}$ que resulta ser solución de $(\mathcal{D})$.

4.3. Problema de la torsión elasto-plástica

Dado $f \in L^2(\Omega)$, consideremos el siguiente problema:

$$(\mathcal{P}) \quad \inf_{v \in C} \left\{ \frac{1}{2} \int_\Omega |\nabla v|^2 dx - \int_\Omega f v dx \right\}$$

donde $C := \{v \in H_0^1(\Omega) \mid \ |\nabla v| \leq 1, \text{ en } \Omega - c.t.p\}$, que es un convexo cerrado y no-vacío . Como el problema es coercivo, se deduce que $S(\mathcal{P}) \neq \emptyset$. Es fácil ver que $u \in S(\mathcal{P})$ si, y sólo si, u es solución de la siguiente desigualdad variacional:

$$(VI) \quad \begin{cases} u \in C, \\ \forall v \in C, \quad ((u, v-u)) - (f, v-u) \geq 0, \end{cases}$$

donde $((\cdot,\cdot))$ denota el producto interno en $H_0^1(\Omega)$, es decir, $((u,v)) = \int_\Omega \nabla u \nabla v$. Notemos que de aquí se deduce la unicidad de la solución de $(\mathcal{P})$. Sea $\bar{u}$ tal solución y recordemos el Teorema de Brezis-Stampacchia:

Teorema 4.3.1. *Si $f \in L^p(\Omega)$ con $1 < p < +\infty$ y Ω es regular (de clase C^2), entonces $\bar{u} \in W^{2,p}(\Omega)$. Si $f \in L^{+\infty}$, entonces $\bar{u} \in W^{2,\alpha}(\Omega)$, $\forall \alpha \in [1, +\infty[$. De las inyecciones de Sobolev, $\bar{u} \in C^1(\Omega)$.*

Tomemos $X = H_0^1(\Omega)$, $X^* = H^{-1}(\Omega)$, $Y = L^2(\Omega)^N = Y^*$. Definamos las funciones $A\colon X \to Y$, $Av = \nabla v$ y $F\colon X \to \mathbf{R}$, $G\colon Y \to \mathbf{R}$ dadas por

$$F(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_\Omega fv dx \quad \text{y} \quad G(p) = \delta_S(p),$$

con $S = \{p \in L^2(\Omega)^N \mid \ |p(x)| \leq 1, \text{ en } \Omega - c.t.p.\}$. Luego, $(\mathcal{P})$ equivale a

$$\inf_{v \in X} F(v) + G(Av).$$

Es fácil ver que

$$F^*(v^*) = \frac{1}{2}\|v^* + f\|_{H^{-1}}^2, \quad G^*(p^*) = \int_\Omega |p^*|$$

de donde formulamos el problema dual por

$$(\mathcal{D}) \qquad \inf_{p^* \in L^2(\Omega)^N} \left\{ \frac{1}{2}\|\operatorname{div} p^* + f\|_{H^{-1}(\Omega)} + \int_\Omega |p^*| \right\}.$$

La calificación dual equivale a $0 \in \operatorname{int}(H^{-1}(\Omega) + A^* \operatorname{dom}(G^*))$ por lo que $S(\mathcal{P}) \neq \emptyset$ (ya lo sabíamos) y $\min(\mathcal{P}) = -\inf(\mathcal{D})$.

Si $\bar{p}^* \in S(\mathcal{D})$, las relaciones de extremalidad conducen a

1. $F(\bar{u}) + F^*(-A^*\bar{p}^*) = \langle \operatorname{div} \bar{p}^*, \bar{u} \rangle$ lo que equivale a $-\Delta \bar{u} - \operatorname{div} \bar{p}^* = f$ (en $H^{-1}(\Omega)$).
2. $G(A\bar{u}) + G^*(\bar{p}^*) = \langle \bar{p}^*, A\bar{u} \rangle$ lo que equivale a $|\nabla \bar{u}| \leq 1$ c.t.p. y $\int_\Omega |\bar{p}^*| = \int_\Omega \bar{p}^* \nabla \bar{u}$. Esto implica que $|\bar{p}^*(x)| = \bar{p}^*(x)|\nabla \bar{u}(x)|$ en $\Omega - c.t.p.$.

De esta relación deducimos que en $\Omega - c.t.p.$ se tiene que

$$\bar{p}^*(x) = \begin{cases} 0 & \text{si } |\nabla \bar{u}(x)| < 1, \\ \lambda(x)\nabla \bar{u}(x) & \text{si } |\nabla \bar{u}| = 1, \end{cases}$$

donde $\lambda(x) = |\bar{p}^*(x)|$. En consecuencia,

$$(*) - \Delta \bar{u} - \operatorname{div}(\lambda(x)\nabla \bar{u}) = f \text{ en } H^{-1}(\Omega).$$

Recíprocamente, si existe $\lambda \in L^2(\Omega)$ tal que $\lambda(x)(|\nabla \bar{u}(x)| - 1) = 0$ y que verifique $(*)$, entonces definiendo $\bar{p}^*(x) := \lambda(x)\nabla \bar{u}(x)$ se verifican las condiciones de extremalidad y en consecuencia $\bar{p}^* \in S(\mathcal{D})$. El problema es que $(\mathcal{D})$ no admite necesariamente una solución. (Notar que $(\mathcal{D})$ es coercivo en $L^1(\Omega)^N$ pero este espacio no es reflexivo).

4.4. Problemas

Problema 23 (Problema de visco-plasticidad)**.** Consideremos el *problema de visco-plasticidad* definido por

$$(P) \qquad \min_{u \in H_0^1(\Omega)} \left\{ \frac{\alpha}{2}\int_\Omega |\nabla u(x)|^2 d\lambda(x) + \beta \int_\Omega |\nabla u(x)| d\lambda(x) - \int_\Omega f(x)u(x)d\lambda(x) \right\},$$

donde Ω es un abierto acotado de $\mathbf{R}^N$, $\alpha, \beta > 0$ y $f \in L^2(\Omega)$. Este problema se puede escribir en el formato de la dualidad de Fenchel-Rockafellar $\min_{u \in H_0^1(\Omega)} \Phi(u) + \Psi(Au)$ tomando $\Phi : H_0^1(\Omega) \to \mathbf{R}$, $\Psi : L^2(\Omega)^N \to \mathbf{R}$ y $A : H_0^1(\Omega) \to L^2(\Omega)^N$ definidos mediante $\Phi(u) = \int_\Omega fu\, d\lambda$, $\Psi(p) = \frac{\alpha}{2}\int_\Omega \|p\|^2 d\lambda + \beta \int_\Omega \|p\| d\lambda$, y $Au = \nabla u$.

(a) Pruebe que el dual correspondiente está dado por

$$(D) \qquad \min_{p^* \in L^2(\Omega)^N} \left\{ \frac{1}{2\alpha} \int_\Omega [|p^*(x)| - \beta]_+^2 d\lambda(x) \;\middle|\; \mathrm{div}(p^*) = -f \right\}.$$

(b) Pruebe que $v(P) + v(D) = 0$, que (P) admite una única solución y que (D) admite soluciones.

(c) Pruebe que $\bar{u} \in S(P)$ y $\bar{p}^* \in S(D)$ si, y sólo si, $\mathrm{div}(\bar{p}^*) = -f$ (en el sentido de distribuciones) y además se tiene que $\nabla \bar{u}(x) = \frac{1}{\alpha}[|\bar{p}^*(x)| - \beta]_+ \bar{p}^*(x)/|\bar{p}^*(x)|$ c.t.p. en Ω.

Problema 24. Sea $\Omega \subset \mathbf{R}^N$ un abierto acotado de frontera regular. Consideremos $f \in L^2(\Omega)$ y $\beta : \mathbf{R} \to \mathbf{R}$ una función continua no-decreciente tal que $|\beta(u)| \leq a + b|u|$ para todo $u \in \mathbf{R}$ con a y b constantes positivas. Para cada condición inicial $u_0 \in L^2(\Omega)$ consideremos la ecuación de evolución

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u - \beta(u) + f \\ u(0) = u_0 \end{cases}$$

(a) Pruebe que esta ecuación admite una única solución $u \in L^\infty(0, \infty; L^2(\Omega))$ tal que $u(t) \in H_0^1(\Omega)$ para $t > 0$.

(b) Pruebe que cuando $t \to \infty$, la trayectoria $t \mapsto u(t)$ converge débilmente en $L^2(\Omega)$ hacia la única solución de

$$(P) \qquad \min_{u \in H_0^1(\Omega)} \left\{ \frac{1}{2} \int_\Omega |\nabla u(x)|^2 dx + \int_\Omega g(u(x)) dx - \int_\Omega f(x) u(x) dx \right\},$$

donde $g(u) = \int_0^u \beta(\xi) d\xi$.

Capítulo 5

Penalización en Optimización Convexa

5.1. Preliminares.

Dado un conjunto finito I consideremos un programa convexo de la forma

$$(\mathcal{P}) \quad v = \min\{f_0(x) \mid f_i(x) \leq 0, \quad i \in I\},$$

y un problema sin restricciones

$$(\mathcal{P}_r) \quad v_r = \min\left\{ f_0(x) + r \sum_{i \in I} \theta\left(\frac{f_i(x)}{r}\right) \;\middle|\; x \in \mathbf{R}^n \right\}$$

definido para cada $r > 0$. La idea es que $(\mathcal{P}_r)$ penaliza la violación de las restricciones de $(\mathcal{P})$ mediante la función θ (más adelante mostraremos hipótesis adecuadas que formalizan esta idea). El objetivo de este capítulo es analizar cuán bien $(\mathcal{P}_r)$ aproxima a $(\mathcal{P})$. Más específicamente, estudiamos la convergencia de los valores de los problemas aproximados $(\mathcal{P}_r)$ al valor del problema $(\mathcal{P})$ (esto es, estudiamos la convergencia de v_r a v) y, además, nos preguntamos por la convergencia de las soluciones generadas a partir de $(\mathcal{P}_r)$. Terminamos el capítulo con el estudio de problemas duales generados a partir de este esquema.

Presentaremos ahora una clase de funciones que será especialmente importante en nuestro análisis.

Definición 5.1.1. Diremos que una función $f\colon \mathbf{R}^n \to \mathbf{R}$ pertenece a la clase $\mathcal{Q}$ si es convexa y satisface la siguiente propiedad:

(*) Si f es constante en el segmento de recta $[x, y]$, entonces es constante en toda la recta que pasa por x e y.

Ejercicio 5.1.1. Suponga que $f \in \mathcal{Q}$ y que el convexo $C \subseteq \mathbf{R}^n$ es tal que f es constante en C. Muestre que f es constante en $\operatorname{aff}(C)$.

Ejemplo 5.1.1. Algunas funciones relevantes:

- Las funciones lineales, las cuadráticas, y las reales-analíticas tienen la propiedad (*).
- Si $f \in \mathcal{Q}$, A es lineal, y $\sigma\colon \mathbf{R} \to \mathbf{R}$ es estrictamente convexa y creciente, entonces $\sigma \circ f \circ A \in \mathcal{Q}$.

- Si f es estrictamente convexa, entonces $f \in \mathcal{Q}$.
- Si $f, g \in \mathcal{Q}$, entonces no necesariamente $\max\{f, g\}, f + g \in \mathcal{Q}$.

A lo largo del capítulo supondremos que

(a) $f_0, f_i \colon \mathbf{R}^n \to \mathbf{R}$ son funciones convexas, I es un conjunto finito;

(b) El conjunto de soluciones $S(\mathcal{P})$ es no-vacío y acotado;

(c) $f_0, f_i \in \mathcal{Q}$;

(d) $\theta \colon (-\infty, \kappa) \to \mathbf{R}$, con $\kappa \in [0, +\infty]$, es una función estrictamente convexa diferenciable tal que $\theta'(u) > 0$, con $\theta'(u) \to 0$ cuando $u \to -\infty$, y $\theta'(u) \to +\infty$ cuando $u \to \kappa-$;

(e) Si $\kappa = 0$ supondremos la condición de Slater: $\exists \hat{x}$ tal que $f_i(\hat{x}) < 0$ $\forall i \in I$.

Diremos que θ es una *función de penalización* si satisface (d) y la extendemos a todo $\mathbf{R}$ mediante

$$\theta(u) = \begin{cases} \lim_{u \to \kappa-} \theta(u) & \text{si } u = \kappa; \\ +\infty & \text{si } u > \kappa. \end{cases}$$

En este contexto, el conjunto de hipótesis (a) ... (e) lo denotaremos (H_0).

Ejemplo 5.1.2 (Funciones de Penalización).

- Penalización Logarítmica

$$\theta(u) = \begin{cases} -\ln(-u) & \text{si } u < 0, \\ +\infty & \text{si no.} \end{cases}$$

- Penalización Inversa

$$\theta(u) = \begin{cases} -\frac{1}{u} & \text{si } u < 0, \\ +\infty & \text{si no.} \end{cases}$$

- Penalización Exponencial

$$\theta(u) = e^u.$$

- Penalización Logarítmica Desplazada

$$\theta(u) = \begin{cases} -\ln(1-u) & \text{si } u < 1, \\ +\infty & \text{si no.} \end{cases}$$

- Penalización Hiperbólica

$$\theta(u) = \begin{cases} \frac{1}{1-u} & \text{si } u < 1, \\ +\infty & \text{si no.} \end{cases}$$

- Penalización Raíz

$$\theta(u) = \begin{cases} -(-u)^{1/2} & \text{si } u \leq 0, \\ +\infty & \text{si no.} \end{cases}$$

5.2. Algunos Resultados de Convergencia

Lema 5.2.1. *Sea $C \subset \mathbf{R}^n$ un convexo no-vacío tal que f_0, f_i son constantes en C. Entonces C tiene un solo elemento.*

Demostración: Supongamos que existen $x, y \in C$, con $x \neq y$. Las funciones f_i, $i \in I \cup \{0\}$, son constantes en $[x, y]$ y luego son también constantes en $\mathrm{aff}(\{x, y\})$. Sea $d = y - x$. Para $t \in \mathbf{R}$ se sigue que $f_i(x + td) = f_i(x)$. Luego[1] $f_i^\infty(d) = 0$ y d es dirección de constancia para f_i, $i \in I \cup \{0\}$.

Consideremos ahora $x^* \in S(\mathcal{P})$. Luego $f_0(x^* + td) = f_0(x^*) = v$ y $f_i(x^* + td) = f_i(x^*) \leq 0$ $\forall t > 0$. Se sigue que $S(\mathcal{P})$ es no acotado, lo cual es una contradicción. □

Proposición 5.2.1. *Para cada $r > 0$ existe un único mínimo $x(r) \in S(\mathcal{P}_r)$. Además $v_r \to v$ y $dist(x(r), S(\mathcal{P})) \to 0$ cuando $r \to 0^+$.*

Demostración: Fijemos $r > 0$ y sea $f_r(x) = f_0(x) + r \sum_{i \in I} \theta(\frac{f_i(x)}{r})$. Probaremos que f_r es inf-compacta o, de manera equivalente, que $f_r^\infty(d) > 0$ para todo $d \neq 0$.

No es difícil mostrar que

$$f_r^\infty(d) = \begin{cases} +\infty & \text{si para algún } i, f_i^\infty(d) > 0, \\ f_0^\infty(d) & \text{si no.} \end{cases}$$

Si $d \neq 0$ es tal que $f_r^\infty(d) \leq 0$, entonces d satisface que $f_i^\infty(d) \leq 0$ para todo $i \in I \cup \{0\}$. Tomando $x^* \in S(\mathcal{P})$ se sigue que $x^* + td \in S(\mathcal{P})$ para todo $t > 0$ contradiciendo así el acotamiento de $S(\mathcal{P})$. En consecuencia, $S(\mathcal{P}_r)$ es convexo, no-vacío y acotado.

Veamos ahora que $S(\mathcal{P}_r)$ tiene un solo elemento. Si no, existen $x, y \in S(\mathcal{P}_r)$, con $x \neq y$. Luego $[x, y] \in S(\mathcal{P}_r)$. Se tiene que $f_i(x) = f_i(y)$ $\forall i \in I$. De lo contrario, tomando $z = (x + y)/2$,

$$\begin{aligned} f_r(z) &\leq \frac{1}{2}\big(f_0(x) + f_0(y)\big) + r \sum_{i \in I} \theta\left(\frac{(f_i(x) + f_i(y))/2}{r}\right) \\ &< \frac{1}{2}\big(f_0(x) + f_0(y)\big) + r \sum_{i \in I} \frac{1}{2}\theta\left(\frac{f_i(x)}{r} + \frac{f_i(y)}{r}\right) \\ &= v_r \end{aligned}$$

donde la segunda desigualdad es estricta ya que estamos suponiendo que existe $i \in I$ tal que $f_i(x) \neq f_i(y)$. Pero lo anterior no puede ser ya que v_r es el ínfimo de $(\mathcal{P}_r)$. Se sigue que $f_i(x) = f_i(y)$ $\forall i \in I$ y luego $f_0(x) = f_0(y)$.

De este modo las funciones f_i son constantes en $S(\mathcal{P}_r)$ para $i \in I \cup \{0\}$ y, en virtud del lema anterior, $S(\mathcal{P}_r) = \{x(r)\}$.

Estudiemos la convergencia. Consideremos $x^* \in S(\mathcal{P})$ y $x_\epsilon^* = (1 - \epsilon)x^* + \epsilon\hat{x}$, con $\hat{x} = x^*$ si $\kappa > 0$, y $\hat{x}$ un punto de Slater si $\kappa = 0$. Ya que $v_r \leq f_r(x_\epsilon^*)$, se tiene que

$$\limsup_{r \to 0^+} v_r \leq \limsup_{r \to 0^+} f_r(x_\epsilon^*) \leq f_0(x_\epsilon^*) \leq (1 - \epsilon)f_0(x^*) + \epsilon f_0(\hat{x}).$$

Como $v = f_0(x^*)$, se sigue que $\limsup_{r \to 0^+} v_r \leq v$.

[1] Para simplificar la notació, en este capítulo denotaremos la función de recesión mediante el símbolo ∞ como superíndice, a diferencia de lo que hicimos en la Definición 1.2.5.

Veamos ahora que $x(r)$ es acotada. De lo contrario, existe $r_k \to 0^+$ tal que $\|x(r_k)\| \to +\infty$ y $\frac{x(r_k)}{\|x(r_k)\|} \to d \neq 0$. Se tiene que $f_{r_k}(x(r_k)) \leq f_{r_k}(x^*_\epsilon)$ y luego

$$\frac{f_0(x(r_k)) + r_k \sum_{i\in I} \theta(\frac{f_i(x(r_k))}{r_k})}{\|x(r_k)\|} \leq \frac{f_{r_k}(x^*_\epsilon)}{\|x(r_k)\|}.$$

Esto equivale a

$$\frac{f_0(x(r_k))}{\|x(r_k)\|} + \sum_{i\in I} \frac{\theta(\frac{f_i(x(r_k))}{\|x(r_k)\|}\frac{\|x(r_k)\|}{r_k})}{\frac{\|x(r_k)\|}{r_k}} \leq \frac{f_{r_k}(x^*_\epsilon)}{\|x(r_k)\|},$$

de donde se sigue que $f_0^\infty(d) + \sum_{i\in I} \theta^\infty(f_i^\infty(d)) \leq 0$. En consecuencia, $f_i^\infty(d) \leq 0$ $i \in I \cup \{0\}$ contradiciendo de este modo el acotamiento de $S(\mathcal{P})$.

Consideremos $x_0 = \lim_{k\to\infty} x(r_k)$ un punto de acumulación de $(x(r))_{r>0}$. Se tiene que $f_{r_k}(x(r_k)) = v_{r_k}$ y, tomando límite, deducimos que $f_0(x_0) + \sum_{i\in I} \theta^\infty(f_i(x_0)) \leq v$. Se sigue que $f_i(x_0) \leq 0$ $\forall i \in I$ y $f_0(x_0) \leq v$ lo que equivale a $x_0 \in S(\mathcal{P})$.

Finalmente mostremos que $v_r \to v$. Tomemos r_k tal que $v_{r_k} \to \liminf_{r\to 0^+} v_r$. Sin pérdida de generalidad, podemos suponer que $x(r_k) \to x_0$. Del argumento anterior

$$v_{r_k} \to f_0(x_0) + \sum_{i\in I} \theta^\infty(f_i(x_0)) = v,$$

y luego $\liminf_{r\to 0^+} v_r = v$. □

Hemos probado que $v_r \to v$ y que cualquier punto de acumulación de la sucesión $x(r)$ es solución del problema $(\mathcal{P})$. Sin embargo, es de especial interés analizar la convergencia de $x(r)$ (obviamente, esto no es trivial sólo si $S(\mathcal{P})$ tiene más de un elemento).

Ejemplo 5.2.1. *Esquema de Penalización de Tikhonov.* Sea el programa convexo

$$(\mathcal{P}) \quad v = \min\{\varphi(x) \mid x \in \mathbf{R}^n\},$$

donde $\varphi \in \Gamma_0(\mathbf{R}^n)$, y consideremos $x(\epsilon)$ la única solución del problema

$$v_\epsilon = \min\left\{\varphi(x) + \frac{\epsilon}{2}\|x\|^2 \mid x \in \mathbf{R}^n\right\},$$

con $\epsilon > 0$. Supongamos que $S(\mathcal{P}) \neq \emptyset$ y consideremos $\bar{x} \in S(\mathcal{P})$. Es fácil ver que

$$\varphi(x(\epsilon)) + \frac{\epsilon}{2}\|x(\epsilon)\|^2 \leq \varphi(\bar{x}) + \frac{\epsilon}{2}\|\bar{x}\|^2 \leq \varphi(x(\epsilon)) + \frac{\epsilon}{2}\|\bar{x}\|^2.$$

Luego $\|x(\epsilon)\| \leq \|\bar{x}\|$ y $x(\epsilon)$ es una sucesión acotada y converge (vía subsucesión) a x^*. Del mismo modo $\varphi(x(\epsilon)) \leq \varphi(\bar{x}) + \frac{\epsilon}{2}\|\bar{x}\|^2$. Tomando límite cuando $\epsilon \to 0^+$, deducimos que $\varphi(x^*) \leq \varphi(\bar{x}) = v$. Así $\operatorname{dist}(x(\epsilon), S(\mathcal{P})) \to 0$. Más aun, ya que $\|x^*\| \leq \|\bar{x}\|$ $\forall \bar{x} \in S(\mathcal{P})$, se sigue que $x(\epsilon)$ *converge* al elemento de norma mínima de $S(\mathcal{P})$:

$$x(\epsilon) \to \mathrm{Proy}_{S(\mathcal{P})}(0), \quad \epsilon \to 0^+.$$

Volvamos al análisis de la convergencia de las soluciones primales en nuestro esquema de penalización. Sea $x^* = \lim_{k\to+\infty} x(r_k)$ un punto de acumulación de $x(r)$. Luego $x^* \in S(\mathcal{P})$. Sea $\bar{x} \in S(\mathcal{P})$ y $\widetilde{x}_k = x(r_k) - x^* + \bar{x} \to \bar{x}$ cuando $k \to +\infty$. Se tiene que f_0 es constante en $[x^*, \bar{x}]$ y luego es también constante en $\operatorname{aff}([x^*, \bar{x}])$. Se sigue que $d = \bar{x} - x^*$ es dirección de constancia para f_0 y en consecuencia $f_0(\widetilde{x}_k) = f_0(x(r_k) + d) = f_0(x(r_k))$. De la desigualdad $f_{r_k}(x(r_k)) \le f_{r_k}(\widetilde{x}_k)$ se sigue que

$$\sum_{i\in I} \theta\left(\frac{f_i(x(r_k))}{r_k}\right) \le \sum_{i\in I} \theta\left(\frac{f_i(\widetilde{x}_k)}{r_k}\right).$$

Del mismo modo podemos definir $I_0 = \{i \in I \mid f_i \text{ es constante en } S(\mathcal{P})\}$ y deducir que para $i \in I_0$ la dirección d es de constancia para f_i. De este modo se tiene que

$$\sum_{i\notin I_0} \theta\left(\frac{f_i(x(r_k))}{r_k}\right) \le \sum_{i\notin I_0} \theta\left(\frac{f_i(\widetilde{x}_k)}{r_k}\right).$$

Motivados por el Ejemplo 5.2.1, quisiéramos ahora, mediante un proceso límite, obtener información del tipo $\Gamma(x_0) \le \Gamma(x^*)$, donde Γ es algún criterio que tendremos que definir. En la próxima sección presentaremos las herramientas que nos permitirán abordar esta pregunta más adelante.

5.3. Medias Asintóticas

En lo que resta del capítulo, suponemos que θ es tal que para $y \in]-\infty, 0[^k$ la función

$$A_\theta(y) = \lim_{r\to 0^+} r\theta^{-1}\left(\frac{1}{k}\sum_{i=1}^{k} \theta\left(\frac{y_i}{r}\right)\right)$$

está bien definida.

Ejemplo 5.3.1.

- Para $\theta(y) = -\ln(-y)$, $A_\theta(y) = -\left[\Pi_{i=1}^k(-y_i)\right]^{1/k}$ es la media geométrica.
- Para $\theta(y) = -\frac{1}{y}$, $A_\theta(y) = \frac{1}{\frac{1}{k}\sum_{i=1}^k \frac{1}{y_i}}$ es la media armónica.
- Para $\theta(y) = e^y$, $A_\theta(y) = \max_{j=1,\dots,k} y_j$.

Ejercicio 5.3.1. Pruebe las afirmaciones del ejemplo anterior.

Observación 5.3.1. Si $z \mapsto M(z) = \theta^{-1}\left(\frac{1}{k}\sum_{i=1}^k \theta(z_i)\right)$ es convexa, entonces $A_\theta \equiv M^\infty$ y en particular, A_θ está bien definida.

Proposición 5.3.1. *La función $A_\theta\colon]-\infty, 0[^k \to \mathbf{R}$ es simétrica, positivamente homogénea, no-decreciente por componentes, convexa y satisface*

$$\frac{1}{k}\sum_{i=1}^{k} y_i \le A_\theta(y) \le \max_{i=1,\dots,k} y_i.$$

Demostración: Todas las propiedades son fáciles de demostrar salvo la convexidad. Veamos primero que A_θ es cuasi-cónvexa (tiene conjuntos de nivel convexos). Para esto basta probar que la función

$$A^r(y) = r\theta^{-1}\left(\frac{1}{k}\sum_{i=1}^{k}\theta\left(\frac{y_i}{r}\right)\right)$$

es cuasi-convexa (pues la cuasi-convexidad se preserva bajo límite). Es directo ver que dados $y \in]-\infty, 0[^k$ y $\lambda \in \mathbf{R}$ se tiene que

$$\{y \mid A^r(y) \leq \lambda\} = \left\{y \;\middle|\; \frac{1}{k}\sum_{i=1}^{k}\theta\left(\frac{y_i}{r}\right)\theta\left(\frac{\lambda}{r}\right)\right\}$$

de donde la propiedad se sigue en virtud de la convexidad de la función de penalización θ. Finalmente, la convexidad de A_θ se sigue del lema de más abajo. □

Lema 5.3.1 (Crouzeix)**.** *Sea $\delta\colon P \to \mathbf{R}$ una función positivamente homogénea y cuasi-convexa, con P un cono convexo. Si $\delta(y) < 0$ $\forall y \in P$ (o bien $\delta(y) > 0$ $\forall y \in P$), entonces δ es convexa.*

Extendamos continuamente ahora la función A_θ a todo $]-\infty, 0]^k$ y notemos que tal extensión es única y la denotamos también A_θ. En efecto, tal extensión está dada por la fórmula

$$A_\theta(y) = \lim_{\epsilon \to 0^+} A_\theta(y - \epsilon\mathbf{1}),$$

donde $\mathbf{1}$ es un vector que contiene 1 en todas sus componentes. Es directo ver que A_θ es simétrica, positivamente homogénea, convexa, continua, no-decreciente por componentes, y satisface

$$\frac{1}{k}\sum_{i=1}^{k} y_i \leq A_\theta(y) \leq \max_{j=1,\ldots,k} y_j.$$

Ejercicio 5.3.2. Muestre que efectivamente A_θ es continua en $]-\infty, 0]^k$.

La siguiente propiedad muestra que es posible definir clases de equivalencias para funciones de penalización que comparten función de media asintótica.

Proposición 5.3.2. *Sean $\theta_1, \theta_2\colon \mathbf{R} \to \mathbf{R} \cup \{+\infty\}$ convexas, estrictamente crecientes y finitas en $]-\infty, 0]$. Si se tienen las siguientes condiciones*

(a) $\inf\theta_1 = \inf\theta_2$*;*

(b) $\lim_{u\to-\infty}\frac{\theta_2^{-1}(\theta_1(u))}{u} = \beta > 0$*;*

entonces A_{θ_1} y A_{θ_2} coinciden (o ninguna de las dos existe).

Demostración: Para cada $\rho > 1$, existe $r(\rho) > 0$ tal que

$$\frac{\beta}{\rho} < \frac{\theta_2^{-1}(\theta_1(y_i/r))}{y_i/r} < \beta\rho \quad \forall r \in]0, r(\rho)[.$$

Suponiendo que $A_{\theta_2}(y)$ existe, de lo anterior se sigue que

$$\begin{aligned} A_{\theta_2}(\beta\rho y) &= \beta\rho A_{\theta_2}(y) \\ &\leq \liminf_{r\to 0^+} \Psi_r \\ &\leq \limsup_{r\to 0^+} \Psi_r \\ &\leq A_{\theta_2}(\frac{\beta}{\rho}y) \\ &= \frac{\beta}{\rho}A_{\theta_2}(y), \end{aligned}$$

con $\Psi_r = r\theta_2^{-1}(\frac{1}{k}\sum_{i=1}^k \theta_1(\frac{y_i}{r}))$.

Por otro lado, se tiene que $\frac{1}{k}\sum_{i=1}^k \theta_1(\frac{y_i}{r}) \to \inf \theta_1 = \gamma$ cuando $r \to 0^+$. Además, de (b) se sigue que

$$\frac{\theta_2^{-1}(x)}{\theta_1^{-1}(x)} \to \beta$$

cuando $x \to \gamma$. De estas dos observaciones deducimos que

$$\frac{r\theta_2^{-1}(\frac{1}{k}\sum_{i=1}^k \theta_1(\frac{y_i}{r}))}{r\theta_1^{-1}(\frac{1}{k}\sum_{i=1}^k \theta_1(\frac{y_i}{r}))} \to \beta$$

cuando $r \to 0^+$. En consecuencia,

$$\liminf_{r\to 0^+} \Psi_r = \beta \liminf_{r\to 0^+} r\theta_1^{-1}\left(\frac{1}{k}\sum_{i=1}^k \theta_1\left(\frac{y_i}{r}\right)\right)$$

y

$$\limsup_{r\to 0^+} \Psi_r = \beta \limsup_{r\to 0^+} r\theta_1^{-1}\left(\frac{1}{k}\sum_{i=1}^k \theta_1\left(\frac{y_i}{r}\right)\right).$$

De esto se deduce que

$$\begin{aligned} \beta\rho A_{\theta_2}(y) &\leq \beta \liminf_{r\to 0^+} r\theta_1^{-1}\left(\frac{1}{k}\sum_{i=1}^k \theta_1\left(\frac{y_i}{r}\right)\right) \\ &\leq \beta \limsup_{r\to 0^+} r\theta_1^{-1}\left(\frac{1}{k}\sum_{i=1}^k \theta_1\left(\frac{y_i}{r}\right)\right) \\ &\leq \frac{\beta}{\rho}A_{\theta_2}(y). \end{aligned}$$

Tomando $\rho \to 1$ se concluye que $A_{\theta_1}(y)$ existe y coincide con $A_{\theta_2}(y)$.

Finalmente, para probar la recíproca basta notar que

$$\lim_{u\to -\infty} \frac{\theta_1^{-1}(\theta_2(u))}{u} = \lim_{v\to -\infty} \frac{v}{\theta_2^{-1}(\theta_1(v))} = \frac{1}{\beta}$$

e intercambiar los roles de $A_{\theta_2}(y)$ y $A_{\theta_1}(y)$ $\square$

Corolario 5.3.1. *Sea θ como en la proposición anterior. Entonces*

(i) Si $\lim_{u\to-\infty} u\theta(u) < 0$, *entonces*

$$A_\theta(y) = \left[\frac{1}{k}\sum_{i=1}^{k}\frac{1}{y_i}\right]^{-1}.$$

(ii) Si $\lim_{u\to-\infty} ue^{\theta(u)} < 0$, *entonces*

$$A_\theta(y) = -\left[\Pi_{i=1}^{k}(-y_i)\right]^{1/k}.$$

(iii) Si $\lim_{u\to-\infty} \frac{\ln(\theta(u))}{u} < 0$, *entonces*

$$A_\theta(y) = \max_{i=1,\dots,k} y_i.$$

Ejercicio 5.3.3. Pruebe que si θ es además de clase C^2 en $\mathbf{R}_-$, con $\lim_{u\to-\infty}\frac{\theta''(u)}{\theta'(u)} > 0$, entonces $A_\theta(y) = \max_{i=1,\dots,k} y_i$.

5.4. Convergencia Primal del Método de Penalización

Adicionalmente en el resto del análisis consideramos la hipótesis

$$(H_1) \quad \forall u, v \in [-\infty, 0]^k,\ \max_{i=1,\dots,k} u_i \neq \max_{i=1,\dots,k} v_j \Rightarrow \inf_{w\in[u,v]} A_\theta(w) < \max\{A_\theta(u), A_\theta(v)\}.$$

Definición 5.4.1. Para cada convexo cerrado no-vacío $C \subseteq S(\mathcal{P})$ definimos

$$I_C = \{i \in I \mid f_i \text{ no es constante en } C\}.$$

Cuando $I_C \neq \emptyset$ definimos $\varphi_C\colon \text{aff}(C) \to \mathbf{R}\cup\{+\infty\}$ mediante

$$\varphi_C(x) = \begin{cases} A_\theta((f_i(x) \mid i \in I_C)) & \text{si } x \in C, \\ +\infty & \text{si no.} \end{cases}$$

Del mismo modo definimos $v_C = \inf \varphi_C$ y $S_C = \arg\min \varphi_C$.

Lema 5.4.1. *Consideremos $C \subseteq S(\mathcal{P})$ convexo cerrado no-vacío tal que $S_C \neq \emptyset$. Entonces*

(i) $\exists \hat{x} \in C$ *tal que* $f_i(\hat{x}) < 0\ \forall i \in I_C$.

(ii) $\forall x, y \in C, \forall i \notin I_C$ *se tiene que* $f_i^\infty(x-y) = 0$.

(iii) $I_C = \emptyset$ *si, y sólo si, C tiene un solo elemento.*

(iv) Si $v^j \to v, r_j \to 0^+, v \in]-\infty, 0]^k$, entonces

$$A_\theta(v) \leq \liminf_{k\to+\infty} r_k\theta^{-1}\big(\frac{1}{k}\sum_{i=1}^{k}\theta(\frac{v_i^j}{r_k})\big).$$

Si $v \in]-\infty, 0[^k$, entonces el límite existe y

$$A_\theta(v) = \lim_{k\to+\infty} r_k\theta^{-1}\big(\frac{1}{k}\sum_{i=1}^{k}\theta(\frac{v_i^j}{r_k})\big).$$

(v) Si $I_C \neq \emptyset$, entonces $S_C \subset C$ y $\exists i \in I_C$ $\exists \beta < 0$ tales que $f_i(x) = \beta$ $\forall x \in S_C$.

Demostración: (i) Dado $x \in C$, $f_i(x) \leq 0$ $\forall i \in I_C$. Como para cada $i \in I_C$, f_i no es constante en C, existe $x_i \in C$ tal que $f_i(x) < 0$ y $f_j(x_i) \leq 0 \quad \forall j \in I_C \setminus \{i\}$. Tomando $\hat{x} = \frac{1}{|I_C|}\sum_{i\in I_C} x_i \in C$ se tiene que $f_i(\hat{x}) < 0$ $\forall i \in I_C$.

(ii) Para $i \notin I_C$ y $x, y \in C$ se tiene que f_i es constante en $[x, y]$. Luego f_i es también constante en aff$(\{x, y\})$ de donde se sigue que $f_i^\infty(x - y) = 0$.

(iii) Supongamos que $I_C = \emptyset$ y que C no es un síngleton. Luego existen $x, y \in C, x \neq y$. De (ii) se sigue que $x - y$ es dirección de constancia para f_i, $i \in I \cup \{0\}$, y luego $S(\mathcal{P})$ es no acotado, lo cual es una contradicción. La otra implicancia es evidente.

(iv) Sean $\epsilon > 0$ y j suficientemente grande de modo tal que $v - \epsilon\mathbf{1} \leq v^j$. Se tiene que

$$r_j\theta^{-1}\left(\frac{1}{k}\sum_{i=1}^{k}\theta\left(\frac{v_i - \epsilon}{r_j}\right)\right) \leq r_j\theta^{-1}\left(\frac{1}{k}\sum_{i=1}^{k}\theta\left(\frac{v_i^j}{r_j}\right)\right)$$

de donde deducimos que

$$A_\theta(v - \epsilon\mathbf{1}) \leq \liminf_{j\to+\infty} r_j\theta^{-1}\left(\frac{1}{k}\sum_{i=1}^{k}\theta\left(\frac{v_i^j}{r_j}\right)\right).$$

La primera parte de la propiedad se sigue entonces de la arbitariedad de $\epsilon > 0$.

De la misma forma se tiene que $v + \epsilon\mathbf{1} < 0$ y $v^j < v + \epsilon\mathbf{1}$, para ϵ y j adecuadamente escogidos. Luego, si $v < 0$,

$$\limsup_{j\to+\infty} r_j\theta^{-1}\left(\frac{1}{k}\sum_{i=1}^{k}\theta\left(\frac{v_i^j}{r_j}\right)\right) \leq A_\theta(v + \epsilon\mathbf{1})$$

de donde se sigue la igualdad deseada.

(v) Para probar que $S_C \neq C$ basta mostrar la existencia de $i \in I_C \setminus I_{S_C}$, es decir, basta ver que para algún $i \in I_C$, f_i es constante sobre S_C.

Se tiene que $\max_{i\in I_C} f_i$ es constante en S_C. De lo contrario existirían $x, y \in S_C$ tales que $\max_{i\in I_C} f_i(x) \neq \max_{i\in I_C} f_i(y)$. Tomando $z = \alpha y + (1-\alpha)x$ se tiene que $f(z) \leq \alpha f(y) + (1-\alpha)f(x)$ de modo que

$$A_\theta(\varphi_C(z)) = A_\theta((f_i(z) \mid i \in I_C)) \leq A_\theta(\alpha(f_i(y) \mid i \in I_C) + (1-\alpha)(f_i(x) \mid i \in I_C)).$$

Notemos que el argumento de A_θ en el término de la izquierda pertenece a $[f(y), f(x)]$. Por otro lado, en virtud de (H_1), se tiene que

$$\inf_{w\in[u,v]} A_\theta(u) < \max\{\varphi_C(x), \varphi_C(y)\} = \inf \varphi_C,$$

lo que entra en contradicción con lo anterior.

Definamos β como el valor de $\max_{i\in I_C} f_i$ sobre S_C. Probemos que existe $i \in I_C$ tal que $f_i(x) = \beta\ \forall x \in S_C$. Supongamos que no. Luego para cada $i \in I_C$ existe $x_i \in S_C$ tal que $f_i(x) < \beta$ y $f_j(x_i) \leq \beta\ \forall j \in I_C \setminus \{i\}$. Tomando $\hat{x} = \frac{1}{|I_C|}\sum_{i\in I_C} x_i$ se tiene que $f_i(\hat{x}) < \beta\ \forall i \in I_C$ y en consecuencia $\max_{i\in I_C} f_i(\hat{x}) < \beta$. □

Corolario 5.4.1. *Sea la sucesión $S^0 = S(\mathcal{P})$ y $S^{k+1} = \arg\min \varphi_{S^k}$. Entonces $S^0 \supset S^1 \supset S^2 \supset \ldots$. Además, definiendo $I^k = I_{S^k}$, se tiene que $I^0 \supset I^1 \supset I^2 \supset \ldots$. En particular, existe $\bar{k}$ tal que $I_{S^{\bar{k}}} = \emptyset$ y por lo tanto $S^{\bar{k}} = \{x^\theta\}$.*

A continuación mostramos el principal resultado concerniente a la convergencia primal del método de penalización.

Teorema 5.4.1. *Supongamos $(H_0), (H_1)$. Entonces $(\mathcal{P}_r)$ admite una única solución $x(r)$ y se tiene que $v_r \to v$ y $x(r) \to x^\theta$ cuando $r \to 0^+$.*

Demostración: Sea $\bar{x} = \lim_{j\to+\infty} x(r_j)$ un punto de acumulación de $x(r)$. Luego $\bar{x} \in S^0$. Probemos inductivamente que $\bar{x} \in S^{k+1}$ con lo cual obtendremos que $\bar{x} = x^\theta$.

Supongamos que $\bar{x} \in S^k$. Luego $x_j^\theta = x_j - \bar{x} + x^\theta \to x^\theta$. Notemos que $\bar{x}, x^\theta \in S^k$ de modo que $[\bar{x}, x^\theta] \in S^k$. Además, para $i \notin I^k$, f_i es constante en S^k y, en consecuencia, es también constante en $\mathrm{aff}(\{\bar{x}, x^\theta\})$. De esto se sigue que $\bar{x} - x^\theta$ es dirección de constancia para f_i con $i \notin I^k$. En consecuencia, $f_i(x(r_j)) = f_i(x_j^\theta)\ \forall i \notin I^k$. Por otro lado, de la optimalidad de $x(r_j)$, $f_{r_j}(x(r_j)) \leq f_{r_j}(x_j^\theta)$. Así

$$\frac{1}{r_j}\theta^{-1}\left(\frac{1}{|I^k|}\sum_{i\in I^k}\theta\left(\frac{f_i(x(r_j))}{r_j}\right)\right) \leq \frac{1}{r_j}\theta^{-1}\left(\frac{1}{|I^k|}\sum_{i\in I^k}\theta\left(\frac{f_i(x_j^\theta)}{r_j}\right)\right).$$

Consideramos separadamente dos casos. Primero, si $f_i(x^\theta) < 0\ \forall i \in I^k$. De la desigualdad anterior y del Lema 5.4.1 deducimos que

$$\begin{aligned}
\varphi_{S^k}(\bar{x}) &\leq \liminf_{j\to+\infty} \frac{1}{r_j}\theta^{-1}\left(\frac{1}{|I^k|}\sum_{i\in I^k}\theta\left(\frac{f_i(x(r_j))}{r_j}\right)\right)\\
&\leq \lim_{j\to\infty} \frac{1}{r_j}\theta^{-1}\left(\frac{1}{|I^k|}\sum_{i\in I^k}\theta\left(\frac{f_i(x_j^\theta)}{r_j}\right)\right)\\
&\leq \varphi_{S^k}(x^\theta)\\
&= \inf \varphi_{S^k}.
\end{aligned}$$

De este modo $\bar{x} \in S^{k+1}$.

Consideremos ahora el caso en que $f_i(x^\theta) \geq 0$ para algún $i \in I^k$. Podemos reemplazar x^θ por $(1-\alpha)x^\theta + \alpha\hat{x}$, donde $\hat{x} \in S^k$ es tal que $f_i(\hat{x}) < 0$, $\forall i \in S^k$ (véase el lema anterior parte (i)). Reproduciendo el análisis previo,

$$\varphi_{S^k}(\hat{x}) \leq (1-\alpha)\varphi_{S^k}(x^\theta) + \alpha\varphi_{S^k}(\hat{x}).$$

Tomando $\alpha \to 0^+$, se concluye que $\bar{x} \in S^{k+1}$. □

Ejercicio 5.4.1. Pruebe que si A_θ es estrictamente convexa, entonces $S^1 = \{x^\theta\}$.

5.5. Convergencia Dual del Método de Penalización

Consideremos la función $\varphi\colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R} \cup \{+\infty\}$ definida mediante

$$\varphi(x,y) = \begin{cases} f_0(x) & \text{si } f_i(x) + y_i \leq 0 \ \forall i \in I, \\ +\infty & \text{si no,} \end{cases}$$

la cual es una función de perturbación para $(\mathcal{P})$. Es fácil probar (véase Ejemplo 3.1.1 y Ejercicio 3.1.2) que en este caso el dual esta dado por

$$(\mathcal{D}) \quad \min_{\lambda \geq 0} p(\lambda),$$

con $p(\lambda) = \inf_{x \in \mathbf{R}^n} f_0(x) + \sum_{i \in I} \lambda_i f_i(x)$. Definimos además

$$\varphi_r(x,y) = f_0(x) + r \sum_{i \in I} \theta\left(\frac{f_i(x) + y_i}{r}\right)$$

que resulta ser una función de perturbación para $(\mathcal{P}_r)$. No es difícil mostrar que

$$\varphi_r^*(0,\lambda) = p(\lambda) + r \sum_{i \in I} \theta^*(\lambda_i)$$

y, en consecuencia, el dual generado a partir de esta perturbación es

$$(\mathcal{D}_r) \quad \min_{\lambda \in \mathbf{R}^m} p(\lambda) + r \sum_{i \in I} \theta^*(\lambda_i).$$

El problema $(\mathcal{D}_r)$ puede también entenderse como un método de barrera para el problema $(\mathcal{D})$ ya que $\theta^*(\lambda) = \infty$ para $\lambda < 0$. Notemos además que $\theta^*(0)$ puede ser finito o infinito. Más específicamente, $\theta^*(0)$ es finito si, y sólo si, θ es acotada inferiormente. Aun en este caso, θ^* actua como una barrera pues $\theta^{*\prime}(\lambda) = (\theta')^{-1}(\lambda) \to -\infty$ cuando $\lambda \to 0^+$.

Ejercicio 5.5.1. Pruebe que de la convexidad estricta de θ se sigue la diferenciabilidad de θ^* en $]0,+\infty[$. Muestre además que de la diferenciabilidad de θ en $]-\infty,\kappa[$ se sigue la convexidad estricta de θ^* en $]0,+\infty[$.

De todo lo anterior se deduce el siguiente resultado.

Proposición 5.5.1. *$(\mathcal{D}_r)$ admite a lo más una solución. Si $\lambda(r)$ es tal solución, entonces $\lambda_i(r) > 0$ $\forall i \in I$.*

Ejercicio 5.5.2. Pruebe que $\varphi_r(x, \cdot)$ es finita y continua en 0 para algún $x \in \mathbf{R}^n$. Deduzca la existencia de solución para $(\mathcal{D}_r)$

Proposición 5.5.2. *$(\mathcal{D}_r)$ admite una única solución $\lambda(r)$. Además, $\lambda(r) > 0$ y*

$$\lambda_i(r) = \theta'\left(\frac{f_i(x(r))}{r}\right),$$

donde $x(r)$ es la única solución de $(\mathcal{P}_r)$.

Demostración: Basta probar la fórmula, que se obtiene de la relación $S(\mathcal{D}_r) = \{\lambda \mid \lambda_i = \frac{\partial \varphi_r}{\partial y_i}(x(r), 0)\}$. □

Estudiemos ahora la convergencia de los problemas duales. Para ello haremos uso del siguiente resultado.

Teorema 5.5.1. *Sean $f, g \in \Gamma_0(\mathbf{R}^n)$ tales que*

1. $\arg\min f \cap \mathrm{dom}(g) \neq \emptyset$;
2. *al menos una de ellas tiene conjuntos de nivel acotados;*
3. *g es estrictamente convexa en su dominio;*
4. $g^\infty(d) \geq 0$ $\forall d \in \mathbf{R}^n$.

Supongamos que $z(r)$ es solución de

$$\min_{z \in \mathbf{R}^n} f(z) + rg(z).$$

Entonces $z(r)$ permanece acotada y converge a $z^* = \arg\min\{g(z) \mid z \in \arg\min f\}$.

Demostración: Es directo verificar que $(f + rg)^\infty(d) > 0$ y luego $f + rg$ es una función inf-compacta. En consecuencia, $z(r)$ existe.

Además, se tiene que el problema

$$\min\{g(z) \mid z \in \arg\min f\}$$

admite una única solución z^*.

Probemos que $z(r)$ es acotada. De lo contrario, existiría $r_k \to 0^+$ tal que $\|z(r_k)\| \to +\infty$. Sin pérdida de generalidad supongamos que $d_k := \frac{z(r_k)}{\|z(r_k)\|} \to d \neq 0$. Notemos que $f(z(r_k)) + r_k g(z(r_k)) \leq f(z^*) + r_k g(z^*) \leq f(z(r_k)) + r_k g(z^*)$. Luego $g(z(r_k)) \leq g(z^*)$ y en consecuencia $g^\infty(d) \geq 0$. Además

$$\frac{f(z(r_k))}{z(r_k)} + r_k \frac{g(z(r_k))}{z(r_k)} \leq \frac{f(z^*) + r_k g(z^*)}{\|z(r_k)\|},$$

lo cual, pasando al límite, entra en contradicción con 2. De este modo $z(r)$ es acotada.

Veamos que todos los puntos de acumulación de $z(r)$ coinciden con z^* (que es la única solución del problema de más arriba). Sea $\bar{z} = \lim_{k\to+\infty} z(r_k)$ para algún $r_k \to 0^+$. De la desigualdad $f(z(r_k)) + r_k g(z(r_k)) \le f(z^*) + r_k g(z^*)$ deducimos que

$$\liminf_{k\to+\infty} f(z(r_k)) + r_k g(z(r_k)) \le f(z^*)$$

Además, es fácil ver que $g(z(r_k))$ está acotada y luego $r_k g(z(r_k)) \to 0$. En consecuencia, $f(\bar{z}) \le f(z^*)$ y $\bar{z} \in \arg\min f$. Como además se tiene que

$$g(\bar{z}) \le \liminf_{k\to+\infty} g(z(r_k)) \le g(z^*),$$

deducimos finalmente que $\bar{z} \in \arg\min\{g(z) \mid z \in \arg\min f\}$. $\square$

A continuación presentamos el principal resultado de esta sección.

Teorema 5.5.2. *Bajo* $(H_0), (H_1)$, *supongamos que* $\theta^*(0)$ *es finito y* $S(\mathcal{D}) \neq \emptyset$. *Entonves* $(\mathcal{D}_r)$ *admite una única solución* $\lambda(r)$ *que converge a* $\lambda^\theta \in S(\mathcal{D})$ *cuando* $r \to 0^+$, *siendo* λ^θ *la única solución de*

$$\min_{\lambda\in S(\mathcal{D})} \sum_{i\in I} \theta^*(\lambda_i).$$

Demostración: Verifiquemos las hipótesis del teorema anterior con $f(\lambda) = p(\lambda) + \delta_{\mathbf{R}^m_+}(\lambda)$ y $g(\lambda) = \sum_{i\in I} \theta^*(\lambda_i)$.

La primera condición es satisfecha pues $\arg\min f \cap \mathrm{dom}(g) = S(\mathcal{D}) \cap \{\lambda \mid \theta^*(\lambda_i) < \infty\} = S(\mathcal{D}) \neq \emptyset$.

Para ver la segunda hipótesis distiguimos dos casos. Primero, si $\kappa = 0$, entonces existe un punto de Slater para $(\mathcal{P})$ y luego $S(\mathcal{D}) \neq \emptyset$ es acotado. En consecuencia, $p(\lambda) + \delta_{\mathbf{R}^m_+}(\lambda)$ tiene conjuntos de nivel acotados. Por otro lado, si $\kappa > 0$, entonces

$$\theta^*(\lambda) \ge \lambda y - \theta(y),$$

con $y > 0, \theta(y) < +\infty$. Luego $\theta(\lambda) \to \infty$ si $\lambda \to \infty$. Además, $\theta^*(\lambda) = \infty$ si $\lambda < 0$. En consecuencia, $g(\lambda)$ tiene conjuntos de nivel acotados.

Por otro lado, como θ^* es estrictamente convexa, $g(\cdot)$ es estrictamente convexa en $]0, \infty[^m$.

Finalmente verifiquemos la cuarta condición del teorema anterior. Ya que $g^\infty(d) = \sum_{i\in I} (\theta^*)^\infty(d_i)$, basta probar que $(\theta^*)^\infty(\pm 1) \ge 0$. Pero esto último es evidente pues

$$(\theta^*)^\infty(1) = \lim_{t\to+\infty} \frac{\theta^*(t)}{t} \ge \kappa,$$

y

$$(\theta^*)^\infty(-1) = \lim_{t\to+\infty} \frac{\theta^*(0-t) - \theta^*(0)}{t} = +\infty.$$

$\square$

Cuando $\theta^*(0) = +\infty$ falla la primera condición del teorema salvo que $(\mathcal{D})$ admita solución $\lambda^* > 0$. Pero esto raramente ocurre. De hecho es usual que exista $i \in I$ tal que $\lambda_i = 0$ para todo $\lambda \in S(\mathcal{D})$. En el caso de la programación lineal

$$(\mathcal{P}) \quad \min_{a_i x \le b_i, i\in I} c^T x$$

con dual

$$(\mathcal{D}) \qquad \min_{c+A^T\lambda=0,\lambda\geq 0} b^T\lambda$$

y esquema de penalización

$$(\mathcal{P}_r) \qquad \min c^T x + r\sum_{i\in I}\theta\Big(\frac{a_i x - b_i}{r}\Big),$$

$$(\mathcal{D}_r) \qquad \min_{c+A^T\lambda=0} b^T\lambda + r\sum_{i\in I}\theta^*(\lambda_i)$$

se tiene el siguiente resultado que presentamos sin demostración (véase la sección de problemas).

Teorema 5.5.3. *Bajo $(H_0),(H_1)$, supongamos que $\theta^*(0)=\infty$ y $S(\mathcal{D})\neq\emptyset$. Entonces $(\mathcal{D}_r)$ admite una única solución $\lambda(r)$ que converge a $\lambda^\theta\in S(\mathcal{D})$ cuando $r\to 0^+$, siendo λ^θ la única solución de*

$$\min_{\lambda\in S(\mathcal{D})}\sum_{i\in I_0}\theta^*(\lambda_i),$$

donde $I_0=\{i\in I\mid \exists\lambda\in S(D),\lambda_i>0\}$.

5.6. Problemas

Problema 25. Sea I finito y para $i\in I$ sean $f_i\colon \mathbf{R}^n\to\mathbf{R}$ funciones convexas. Considere el siguiente problema de minimización

$$(\mathcal{P}) \qquad \inf_{x\in\mathbf{R}^n}\max_{i\in I}\{f_i(x)\}.$$

(a) Pruebe que $x^*\in S(\mathcal{P})$ si, y sólo si, $(\mu^*,x^*)\in\mathbf{R}\times\mathbf{R}^n$, con $\mu^*=\max_{i\in I}\{f_i(x^*)\}$, es solución de

$$\inf_{(\mu,x)\in\mathbf{R}\times\mathbf{R}^n}\{\mu\mid \forall i\in I,\ f_i(x)\leq\mu\}.$$

(b) Definamos $\varphi\colon\mathbf{R}^{n+1}\times\mathbf{R}^m\to\mathbf{R}\cup\{+\infty\}$ mediante

$$\varphi(\mu,x,y)=\mu+\sum_{i=1}^m\delta_{]-\infty,0]}(f_i(x)+y_i-\mu).$$

Verifique que φ es una función de perturbación para $(\mathcal{P})$ y que el problema dual correspondiente está dado por

$$(\mathcal{D}) \qquad \inf_{\lambda\in\mathbf{R}^n} p(\lambda),$$

donde

$$p(\lambda)=\begin{cases}-\inf_{x\in\mathbf{R}^n}\{\sum_{i\in I}\lambda_i f_i(x)\} & \text{si } \lambda\in\Delta_m,\\ +\infty & \text{si no,}\end{cases}$$

donde, recordamos, $\Delta_m=\{\lambda\in\mathbf{R}^m_+\mid\sum_{i\in I}\lambda_i=1\}$.

(c) Suponga que $\inf(\mathcal{P})\in\mathbf{R}$. Pruebe que $S(\mathcal{D})\neq\emptyset$ y que no hay salto de dualidad. Pruebe que $x^*\in S(\mathcal{P})$ y $\lambda^*\in S(\mathcal{D})$ si, y sólo si, $\lambda^*\in\Delta_m$ y $\lambda_i^*(f_i(x^*)-\max_{i\in I}\{f_i(x^*)\})=0$.

(d) Dado $r > 0$, considere la función de penalización exponencial $\bar{\varphi}_r \colon \mathbf{R}^{n+1} \times \mathbf{R}^m \to \mathbf{R}$ definida por
$$\bar{\varphi}_r(\mu, x, y) = \mu + r \sum_{i \in I} \exp\left(\frac{f_i(x) + y_i - \mu}{r}\right)$$
y definamos $\varphi_r(x, y) = \min_{\mu \in \mathbf{R}} \bar{\varphi}_r(\mu, x, y) - r$. Verifique que $\varphi_r \in \Gamma_0(\mathbf{R}^n \times \mathbf{R}^m)$ y que más aun
$$\varphi_r(x, y) = r \ln\left(\sum_{i \in I} \exp\left(\frac{f_i(x) + y_i}{r}\right)\right).$$
Sea $F_r(x) := \varphi_r(x, 0)$ para $r > 0$ y $F_0(x) := F(x) := \max_{i \in I}\{f_i(x)\}$. Pruebe que para $r > 0$ se tiene que
$$F(x) < F_r(x) \leq F(x) + r \ln(m), \quad \forall x \in \mathbf{R}^n.$$

Considere la familia parametrizada de problemas
$$(\mathcal{P}_r) \qquad \inf_{x \in \mathbf{R}^n} F_r(x)$$
y en lo que sigue suponga que $S(\mathcal{P}) = \arg\min F$ es no-vacío y compacto.

(e) Pruebe que $\forall r > 0$, $\min(\mathcal{P}) \leq \inf(\mathcal{P}_r) \leq \min(\mathcal{P}) + r \ln(m)$ y que, más aún, el conjunto $S(\mathcal{P}_r) = \arg\min F_r$ es no-vacío y compacto.

(f) Sea $\gamma > \min(\mathcal{P})$. Pruebe que si $0 < r < \frac{1}{\ln(m)}(\gamma - \min(\mathcal{P}))$, entonces
$$\Gamma_{\min(\mathcal{P}) + r \ln(m)}(F_r) \subseteq \Gamma_\gamma(F).$$
Deduzca que toda selección $r \mapsto x(r) \in S(\mathcal{P}_r)$ permanece acotada cuando $r \to 0^+$ y demuestre que todo punto de acumulación de $x(r)$ pertenece a $S(\mathcal{P})$.

En lo que resta suponga que $f_i \in \mathcal{Q}$, para todo $i \in I$.

(g) Pruebe que el problema $(\mathcal{P}_r)$ admite solución única. Para ello demuestre que para todo $i \in I$, f_i es constante sobre $S(\mathcal{P}_r)$.

(h) Verifique que el problema dual asociado a la perturbación φ_r de $(\mathcal{P}_r)$ es
$$(\mathcal{D}_r) \qquad \inf_{\lambda \in \mathbf{R}^m} \left\{p(\lambda) + r \sum_{i \in I} \lambda_i \ln(\lambda_i)\right\}$$
con la convención $0 \ln(0) = 0$ y $\lambda \ln(\lambda) = +\infty$ para $\lambda < 0$. Pruebe que $\inf(\mathcal{P}_r) + \inf(\mathcal{D}_r) = 0$ y que (D_r) admite solución única $\lambda(r)$.

Para simplificar, supongamos además que $f_i \in C^1(\mathbf{R})$.

(i) Pruebe que
$$\lambda_i(r) = \frac{\exp(f_i(x(r))/r)}{\sum_{i \in I} \exp(f_i(x(r))/r)}.$$
Muestre que, cuando $r \to 0^+$, $\min(\mathcal{D}_r) \to \min(\mathcal{D})$ y todo punto de acumulación de $\lambda(r)$ pertenece a $S(\mathcal{D})$. (Es posible demostrar que $(x(r), \lambda(r))$ en realidad convergen a $(x^*, \lambda^*) \in S(\mathcal{P}) \times S(\mathcal{D})$.)

Problema 26. Sea $f \in \mathcal{Q} \cap \Gamma_0(\mathbf{R}^n)$ (ver Definición 5.1.1). Suponemos que existe $x_0 \in \mathrm{dom}(f)$ con $\|x_0\|_\infty \leq 1$, y consideramos el problema de optimización

$$(P_\infty) \qquad \min_{x \in \mathbf{R}^n} \{f(x) : \|x\|_\infty \leq 1\}$$

junto con el esquema de aproximación (sin restricciones)

$$(P_p) \qquad \min_{x \in \mathbf{R}^n} \left\{ f(x) + \frac{1}{p}\|x\|_p^p \right\}$$

donde $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$.

(a) Pruebe que para cada $p > 1$ el problema (P_p) admite una solución única $x(p)$, y que la trayectoria $p \to x(p)$ permanece acotada cuando $p \to \infty$.

(b) Sea u_∞ un punto de acumulación de $x(p)$. Pruebe que u_∞ es solución de (P_∞) y además es solución de

$$(P^1_\infty) \qquad \min_{x \in S(P_\infty)} \|x\|_\infty.$$

(c) Sea α_1 el valor óptimo de (P^1_∞). Pruebe que existe un índice i_1 tal que $|x_{i_1}| = \alpha_1$ para todo $x \in S(P^1_\infty)$. Deduzca que, o bien $x_{i_0} = \alpha_1$ para todo $x \in S(P^1_\infty)$, o $x_{i_0} = -\alpha_1$ para todo $x \in S(P^1_\infty)$. Denotemos por I_1 el conjunto de tales índices i_1 y $\Pi_1(x) = (x_i : i \notin I_1)$. Pruebe que todo punto de acumulación u_∞ es solución de

$$(P^2_\infty) \qquad \min_{x \in S(P^1_\infty)} \|\Pi_1(x)\|_\infty.$$

(d) Inspirándose en la parte (c) probar que $x(p)$ converge cuando $p \to \infty$ hacia un punto $x^* \in S(P_\infty)$ caracterizado como la solución de una jerarquía de problemas de optimización encajonados.

Consideremos ahora un sistema lineal sobre-determinado $Ay = b$ donde $b \in \mathbf{R}^m$ y $A \in M^{m \times n}$ con $\mathrm{rg}(A) = n \leq m$. Sin pérdida de generalidad suponemos $\|b\|_\infty \leq 1$. Este sistema no admite necesariamente una solución por lo cual consideramos el problema

$$(Q_\infty) \quad \min_{y \in \mathbf{R}^n} \|Ay - b\|_\infty$$

el cual aproximamos mediante el esquema

$$(Q_p) \quad \min_{y \in \mathbf{R}^n} \|Ay - b\|_p.$$

Mostrar que este último admite una única solución $y(p)$ la cual converge hacia una solución $y^* \in S(Q_\infty)$. Indicación: Considere el conjunto $L = \{Ay - b : y \in \mathbf{R}^n\}$ y la función

$$f(x) := \begin{cases} 0 & \text{si } x \in L \\ +\infty & \text{si no.} \end{cases}$$

Problema 27. Sean $f_0, \ldots, f_m : \mathbf{R}^n \to \mathbf{R}$ convexas y consideremos el esquema de aproximación

$$(P_\epsilon) \qquad \min_{x \in \mathbf{R}^n} f_\epsilon(x) := f_0(x) + \epsilon f_1(x) + \cdots + \epsilon^m f_m(x)$$

definido para $\epsilon > 0$, junto con la jerarquía de problemas de optimización

$$\begin{array}{ll} (P_0) & \min\{f_0(x) : x \in \mathbf{R}^n\} \\ (P_1) & \min\{f_1(x) : x \in S_0\} \\ \vdots & \\ (P_m) & \min\{f_m(x) : x \in S_{m-1}\} \end{array}$$

donde S_i denota el conjunto de soluciones óptimas del problema (P_i). Suponemos $S_m \neq \emptyset$ y f_i convexas de clase C^2 tales que $\max_i d^T \nabla^2 f_i(x) d \geq \alpha \|d\|^2$ para una cierta constante $\alpha > 0$ y todos $x, d \in \mathbf{R}^n$.

(a) Muestre que $f_\epsilon(\cdot)$ es fuertemente convexa y (P_ϵ) admite una solución única $x(\epsilon)$.

(b) Pruebe que $x(\epsilon)$ permanece acotada cuando $\epsilon \to 0$. Indicación: razone por contradicción suponiendo que existe $\epsilon_k \to 0$ con $\|x(\epsilon_k)\| \to \infty$ y $d_k := x(\epsilon_k)/\|x(\epsilon_k)\| \to d$. Muestre que $f_0^\infty(d) \leq 0$ y $f_1^\infty(d) \leq 0$. Luego, utilice la desigualdad $f_\epsilon(x(\epsilon)) \leq f_\epsilon(x(\epsilon) - \|x(\epsilon)\| d)$ para probar recursivamente que $f_i^\infty(d) \leq 0$ para $i = 2, \ldots, m$. Concluya.

(c) Muestre que todo punto de acumulación de $x(\epsilon)$ pertenece a S_0 y a S_1.

(d) Muestre que S_m contiene un único elemento, que denotamos x^*.

(e) Suponiendo que $f_0, \ldots, f_m \in \mathcal{Q}$ pruebe que $\lim_{\epsilon \to 0} x(\epsilon) = x^*$.

Problema 28. El objetivo de este problema es demostrar el Teorema 5.5.3 en el caso $\kappa = 0$.

(a) Muestre que $\lambda(r)$ está acotada. Para ello argumente por contradicción deduciendo la existencia de $\mu \in \mathbf{R}^m$ tal que
$$\mu \geq 0, \; A^T \mu = 0, \; b^T \mu \geq 0.$$

(b) Pruebe que todo punto de acumulación de $\lambda(r)$ pertenece a $S(\mathcal{D})$.

(c) Concluya.

Índice alfabético

Bibliografía

[Aub98] J.P. Aubin, *Optima and Equilibria*, Springer, 1998.

[AuT03] A. Auslander, M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, 2003.

[Att84] H. Attouch, *Variational Convergence for Functions and Operators*, Applicable Mathematics Series, Pitman, London, 1984.

[Bre83] H. Brezis, *Analyse fonctionnelle*, Masson Editeur, París, 1983.

[BoL00] J. M. Borwein, A. S. Lewis, *Convex Analysis and Nonlinear Optimization. Theory and Examples*, CMS Books in Mathematics, Springer-Verlag, New York, 2000.

[Cal83] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[EkT74] I. Ekeland, R. Temam, *Analyse Convexe et Problèmes Variationnels*, Dunod, Paris, 1974.

[HiL93] J.-B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin, 1993.

[Roc70] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[Roc74] R.T. Rockafellar, *Conjugate Duality and Optimization*, Conference Board of Mathematical Sciences Series 16, SIAM Publications, Philadelphia, 1974.

[RoW98] R.T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Grundlehren der matematischen Wissenschaften 317, Springer-Verlag, Berlin, 1998.