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

Análisis Convexo y Dualidad

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

Doctorado en Ciencias de la Ingeniería,
mención Modelación Matemática

Felipe Álvarez

con la colaboración de

Juan Escobar, Cristopher Hermosilla y Juan Peypouquet

Marzo 2012

Prefacio

El objetivo de estos apuntes es proporcionar los fundamentos clásicos del Análisis Convexo y de la Teoría de la Dualidad en espacios de dimensión finita e infinita, junto con abordar algunas aplicaciones y desarrollos en Optimización Matemática y el Cálculo de Variaciones. El material aquí presentado se basa 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.

En cada ocasión en que he dictado dicho curso, estas notas se han ampliado y mejorado con el aporte de los alumnos y profesores auxiliares. Mis agradecimientos a Juan Escobar, Cristopher Hermosilla y Juan Peypouquet, quienes mientras eran estudiantes del Departamento de Ingeniería Matemática, aceptaron mi invitación para participar 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. Recientemente, Cristopher Hermosilla incorporó varias figuras para ilustrar los conceptos principales, inspiradas en los bosquejos del manuscrito original, y agregó nuevos ejercicios y soluciones a algunos de los problemas propuestos en los primeros capítulos; además, revisó todo el material completando algunos comentarios, referencias y el índice alfabético.

Cabe señalar que los problemas que tienen un asterisco son aquellos cuya solución es discutida en los anexos.

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 parcial proporcionado por el Departamento de Ingeniería Matemática y el Instituto en Sistemas Complejos de Ingeniería de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile.

Felipe Álvarez
Santiago, 18 de marzo de 2012

Índice general

Capítulo 1

Introducción al Análisis Variacional

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

En el análisis de problemas de minimización y maximización es conveniente considerar funciones que toman valores en la *recta real extendida* $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty]$ en lugar de sólo $\mathbb{R} =]-\infty, +\infty[$. 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 $f_0 \colon X \to \mathbb{R}$ es la *función objetivo* a minimizar y $C \subseteq X$ es un *conjunto de restricciones.* Con la topología apropiada, $\overline{\mathbb{R}}$ es un intervalo compacto. En particular, todo subconjunto $A \subseteq \overline{\mathbb{R}}$ admite un ínfimo $\inf A$ (y un supremo $\sup A$). Luego, tomando $A = \{f_0(x) \mid x \in C\}$ tenemos que $\inf\{f_0(x) \mid x \in C\}$ está bien definido en $\overline{\mathbb{R}}$, 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 (análogo para $\max A$).

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

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

Con la convención

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

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 aún, 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

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

donde $f\colon X \to \mathbb{R}\cup\{+\infty\}$ 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 $\mathbb{R} \cup \{-\infty\}$; los problemas de tipo *minimax*, a funciones a valores en $\overline{\mathbb{R}}$.

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

1. $(+\infty) + \alpha = \alpha + (+\infty) = +\infty$, $\forall \alpha \in \mathbb{R} \cup \{+\infty\}$.
2. $(-\infty) + \alpha = \alpha + (-\infty) = -\infty$, $\forall \alpha \in \mathbb{R} \cup \{-\infty\}$.
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 $\lambda \in \mathbb{R}$ y $f, g : X \to \overline{\mathbb{R}}$.

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 $f\colon X \to \overline{\mathbb{R}}$, su *epígrafo* es el subconjunto de $X \times \mathbb{R}$ dado por

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

mientras que el *conjunto de nivel inferior (o subnivel) de parámetro* $\gamma \in \mathbb{R}$ 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

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

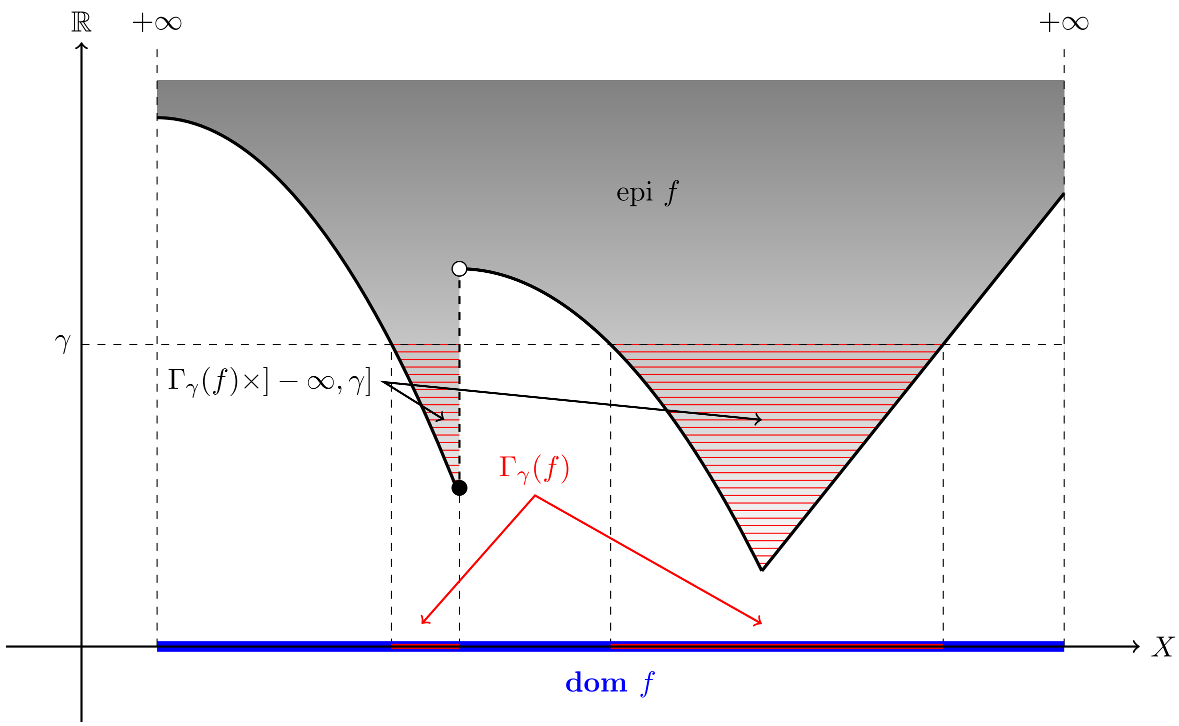


Figura 1.1: Ejemplo de Definición 1.1.1

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 $\overline{\mathbb{R}}$ 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 $f_i\colon X \to \overline{\mathbb{R}}$. Entonces las funciones $\sup_{i\in I} f_i$ e $\inf_{i\in I} f_i$ definidas respectivamente por*

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

e

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

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

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

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:

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

Del mismo modo,

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

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 $i \in I$. Para ver que la igualdad no se tiene en un caso infinito basta considerar $f_i(x) = e^{x-i}$, con $i \in \mathbf{N}$, y ver que $(0,0) \in \operatorname{epi}(\inf f_i) \setminus \operatorname{epi} f_i\ \forall i \in \mathbf{N}$.

□

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 $f\colon X \to \overline{\mathbb{R}}$ 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 $f : X \to \mathbb{R} \cup \{+\infty\}$ se dice *τ-semicontinua inferior (τ-s.c.i.)* en x si

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

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 $f : X \to \mathbb{R} \cup \{+\infty\}$, las siguientes afirmaciones son equivalentes:*

(i) f *es τ-s.c.i. sobre X.*

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

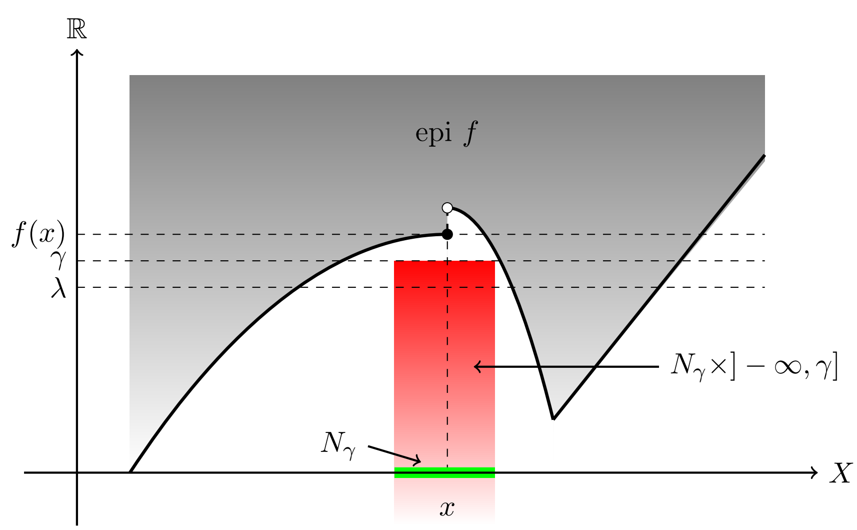
(iii) $\forall \gamma \in \mathbb{R}$, $\Gamma_\gamma(f)$ *es cerrado en (X, τ).*

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

(v) $\forall x \in X,\ f(x) \leq \liminf_{y \to x} f(y) := \sup_{N \in \mathcal{N}_X(\tau)} \inf_{y \in N} f(y)$

Demostración:

- *(i)* $\Rightarrow$ *(ii)* Tomemos $(x, \lambda) \notin \operatorname{epi}(f)$, lo que equivale a $\lambda < f(x)$. Sea $\gamma \in \mathbb{R}$ 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 $(y, \gamma) \notin \operatorname{epi}(f)$. Se sigue que $(N_\gamma \times]-\infty, \gamma[) \cap \operatorname{epi}(f) = \emptyset$. Como $N_\gamma \times]-\infty, \gamma[\in \mathcal{N}_{(x,\lambda)}(\tau \times \tau_{\mathbb{R}})$, concluimos que $\operatorname{epi}(f)^C$ es abierto.



- $(ii) \Rightarrow (iii)$ Como $\Gamma_\gamma(f) \times \{\gamma\} = \text{epi}(f) \cap (X \times \{\gamma\})$, deducimos que $\Gamma_\gamma(f) \times \{\gamma\}$ es cerrado en $X \times \mathbb{R}$, 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 $\gamma \leq \inf_{y \in N} f(y)$, y en particular $\gamma \leq \sup_{N \in \mathcal{N}_x(\tau)} \inf_{y \in N} f(y)$.

  Como lo anterior es válido para todo $\gamma < f(x)$, se sigue (v).
- $(v) \Rightarrow (i)$ Sea $\lambda < f(x)$. Por (v), $\lambda < \sup_{N \in \mathcal{N}_X(\tau)} \inf_{y \in N} f(y)$ y en consecuencia , existe $N \in \mathcal{N}_x(\tau)$: $\lambda < \inf_{y \in N} f(y)$.

□

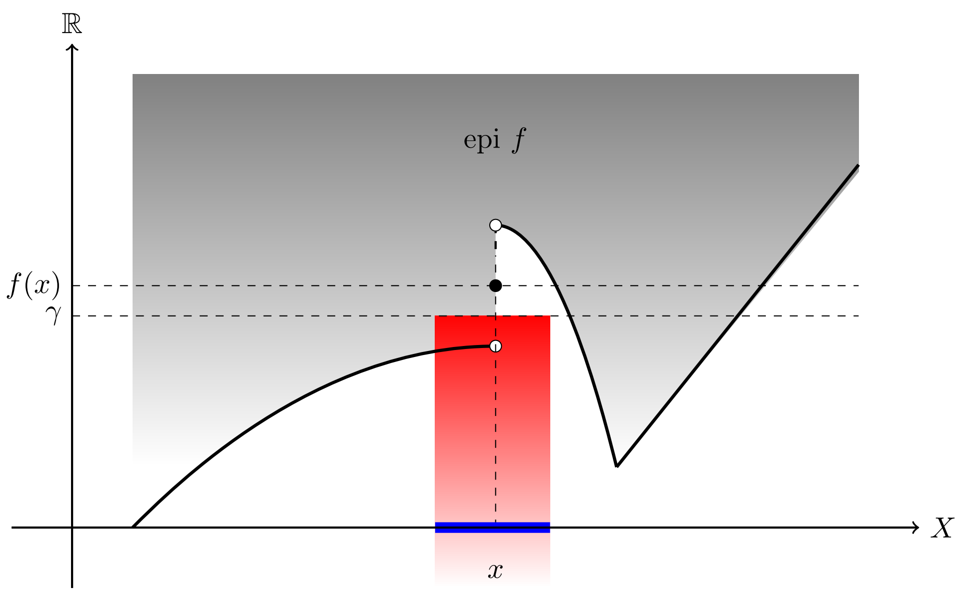


Figura 1.2: Ejemplo de función no s.c.i

Ejercicio 1.2.1. Pruebe que cuando (X, τ) es metrizable, f es τ s.c.i. si y sólo si

$$\forall x \in X \ x_n \to x, \quad f(x) \leq \liminf_{n \to +\infty} f(x_n) := \sup_{n \in \mathbf{N}} \inf_{k \geq n} f(x_k).$$

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., $f_i \colon X \to \mathbb{R} \cup \{+\infty\}$. Entonces $\sup_{i \in I} f_i$ es τ-s.c.i. Si además I es finito, $\min_{i \in I} f_i$ y $\sum_{i \in I} f_i$ 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 $f\colon X \to \mathbb{R} \cup \{+\infty\}$ se dice *τ-inf-compacta* si para todo $\gamma \in \mathbb{R}$, 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 $f\colon V \to \mathbb{R} \cup \{+\infty\}$ se dice *coerciva* si $\lim\limits_{\|x\|\to\infty} f(x) = +\infty$.

Observación 1.2.2. Sea $(V, \|\cdot\|)$ un espacio vectorial normado y $f\colon V \to \mathbb{R} \cup \{+\infty\}$ una función. Entonces son equivalentes (i) f es coerciva y (ii) $\forall \gamma \in \mathbb{R}$, $\Gamma_\gamma(f)$ 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 $f\colon X \to \mathbb{R} \cup \{+\infty\}$ 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
$$f(x_n) \leq \max\{\inf_X f + \frac{1}{n}, -n\} \leq \max\{\inf_X f + 1, -1\} =: \gamma_0 \in \mathbb{R}.$$
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 la función objetivo f,
$$f(x^*) \leq \lim_{k\to\infty} f(x_k) = \inf_X f.$$
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 $\gamma_0 \in \mathbb{R}$ 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 vacíos (¿por qué?) del compacto $\Gamma_{\gamma_0}(f)$. Más aun, esta familia satisface la propiedad 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,

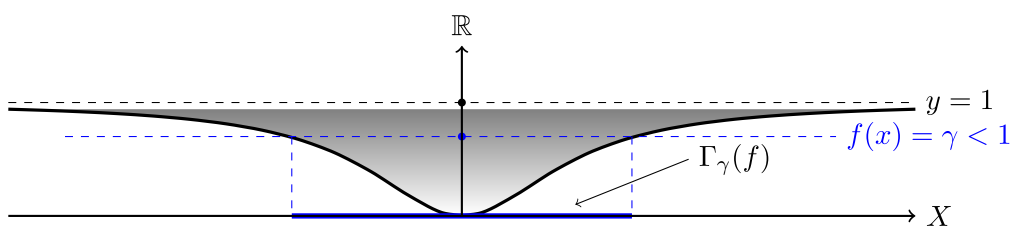
$$\bigcap_{\inf_X f < \gamma < \gamma_0} \Gamma_\gamma(f) \neq \emptyset,$$

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 $f\colon \mathbb{R} \to \mathbb{R}$ por $f(x) = \frac{x^2}{1+x^2}$. Es fácil ver que $\Gamma_\gamma(f)$ es compacto ssi $\gamma < 1$ y $\Gamma_1(f) = \mathbb{R}$ 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. En espacios topológicos más general, la compacidad no implica la existencia de subsucesiones convergentes. Sin embargo, si utilizamos la noción de sucesión generalizada o red, podemos proceder (cuidadosamente) de manera similar.

  Un caso importante de mencionar es cuando la topología utilizada es la topología débil inducida por un espacio de Banach reflexivo. En dimensión infinita, sabemos esta topología no es metrizable, sin embargo, aún podemos extraer una subsucesión convergente. Basta considerar $M = \overline{M_0}$ con M_0 el espacio vectorial generado por la sucesión $(x_n)_n$. Es fácil ver que M también es un espacio de Banach reflexivo y con lo cual B_M, la bola unitaria en M, es compacta para la topología débil. Más aún, como M es separable y reflexivo, M^* es separable y reflexivo, por lo que, B_M es metrizable; ver [Bre83];

Ejercicio 1.2.2. Sea $f\colon X \to \mathbb{R} \cup \{+\infty\}$ 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

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

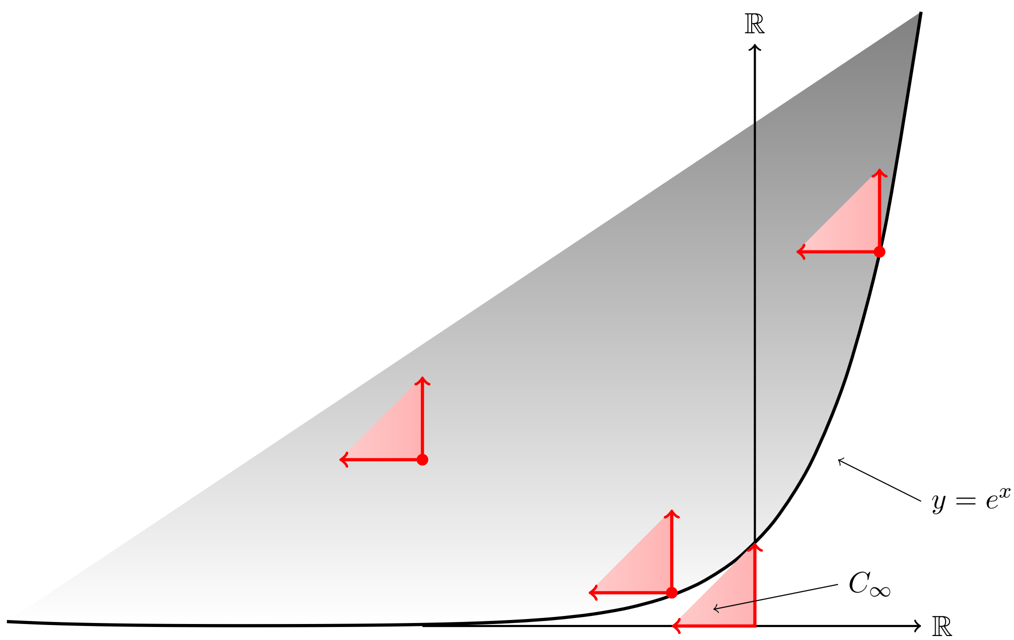


Figura 1.3: Cono de recesión para el conjunto $C = \operatorname{epi} f$ con $f(x) = e^x$.

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, cualquiera que sea el $x \in C$

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

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 un hiperplano cerrado, entonces C_∞ es el subespacio vectorial paralelo a C.

3. Suponga que C es un poliedro convexo definido mediante $C = \{ x \in \mathbb{R}^n \mid Ax \leq b \}$, donde A es una matriz de tamaño $m \times n$ y $b \in \mathbb{R}^m$. Entonces $C_\infty = \{ d \in \mathbb{R}^n \mid Ad \leq 0 \}$.

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 $f : X \to \mathbb{R} \cup \{+\infty\}$ 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 $f, g : X \to \mathbb{R} \cup \{+\infty\}$ 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 $(d, \mu) \in (\operatorname{epi} f)_\infty$. □

Ejercicio 1.2.5. Sea $(f_i)_{i\in I}$ una colección de funciones propias de X en $\mathbb{R}\cup\{+\infty\}$. 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 $f : X \to \mathbb{R}\cup\{+\infty\}$ 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 $\mathbb{R}\cup\{+\infty\}$ y un subconjunto no-vacío S de X. Sea $C = \{\ s \in S \mid f_i(x) \leq 0\ \forall i \in I\ \}$. Entonces*

$$C_\infty \subseteq \{\ d \in S_\infty \mid (f_i)_\infty(d) \leq 0\ \forall i \in I\ \}.$$

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 $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ 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 $\lambda > \inf f$,

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

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 $f_i : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ es s.c.i. y propia. Sea $C = \{\ x \in \mathbb{R}^n \mid f_i(x) \leq 0\ \forall i\ \}$. 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

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

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 2.1 del próximo capítulo presentaremos otros resultados aplicados al caso convexo. En [AuT03] se puede encontrar una exposición más completa y detallada sobre el análisis 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. En [BoZ05] es posible encontrar una amplia gama de aplicaciones de este principio a distintas teorías matemáticas.

Teorema 1.2.3. *Sean (X, d) un espacio métrico completo y $f\colon X \to \mathbb{R}\cup\{+\infty\}$ una función propia, s.c.i. y acotada inferiormente. Consideremos $x_0 \in \operatorname{dom} f$ y $\varepsilon > 0$. Entonces existe $\bar{x} \in X$ 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 $f\colon X \to \mathbb{R}_+ \cup \{+\infty\}$. Consideremos $F : X \to 2^X$ definida por

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

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 $v : \operatorname{dom} f \to \mathbb{R}$ 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 $f\colon X \to \overline{\mathbb{R}}$ y $\varepsilon > 0$, definimos el conjunto de sus *ε-mínimos* por

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

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

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

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

(iii) $f(\bar{x}) < f(x) + \frac{\varepsilon}{\lambda} d(x, \bar{x})$, *para todo* $\forall x \in X \setminus \{\bar{x}\}$

Demostración: Basta aplicar el Teorema 1.2.3, cambiando la métrica d por $\frac{1}{\lambda} d$. □

Observación 1.2.4. La noción de mínimo de una función está fuertemente ligada a la existencia de un hiperplano soportante de pendiente nula. Pero, cuando una función inferiormente acotada no alcanza un mínimo, lo anterior pierde sentido. Sin embargo, el PVE permite reemplazar el concepto de *hiperplano soportante* por el de *cono soportante* asociado a un $\varepsilon-$mínimo. La figura 1.4 evidencia esta afirmación. Para ver otras interpretaciones geométricas del PVE se recomienda ver los primeros capítulos de [BoZ05].

Antes de terminar esta sección recordemos una definición básica del cálculo diferencial.

Definición 1.2.7 (Gâteaux Diferenciabilidad)**.** Sean X, Y espacios de Banach y $f : U \to Y$ una función definida sobre un abierto $U \subseteq X$. Diremos que f es Gâteaux diferenciable en $x \in U$ si

$$\lim_{t \to 0} \frac{f(x + td) - f(x)}{t} = Ad, \quad \forall d \in X,$$

donde $A : X \to Y$ es una función lineal continua. Diremos que f es Gâteaux diferenciable si tal función existe para todo $x \in U$. Esta función A se conoce como la derivada de Gâteaux de f y usualmente es denotada por $Df(x)$.

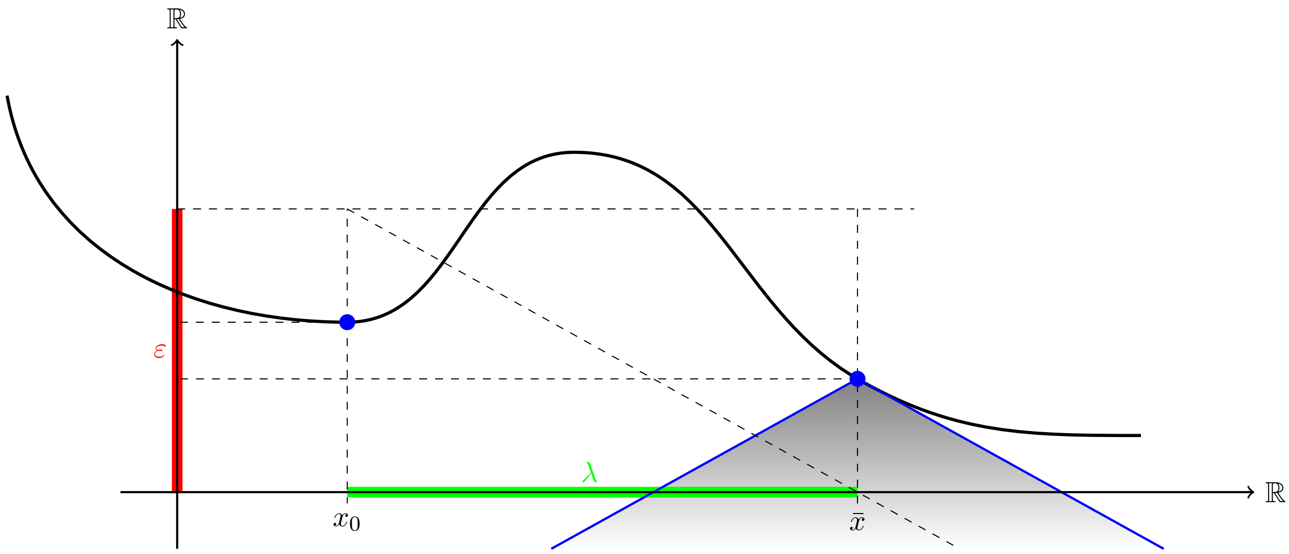


Figura 1.4: Principio Variacional de Ekeland

Corolario 1.2.3. *Supongamos que X es un espacio de Banach. Sea $f : X \to \mathbb{R}$ una función s.c.i, inferiormente acotada y Gâteaux diferenciable en X. Consideremos $\varepsilon > 0$ y sea $x \in \varepsilon - \arg\min f$. Entonces, existe un punto $x_\varepsilon \in X$ tal que*

(i) $f(x_\varepsilon) \leq f(x)$,

(ii) $\|x - x_\varepsilon\| \leq 1$,

(iii) $\|Df(x_\varepsilon)\| \leq \varepsilon$.

Consecuentemente, existe una sucesión minimizante $\{x_n\}$ en X que satisface

$$f(x_n) \to \inf f, \quad y \quad Df(x_n) \to 0.$$

Demostración: En vista del Teorema 1.2.4 con $\lambda = 1$, sólo es necesario probar la afirmación (iii). Sea $d \in X$ con $\|d\| = 1$, como f es Gâteaux diferenciable, tenemos que

$$tDf(x_\varepsilon)d + o(t) = f(x_\varepsilon + td) - f(x_\varepsilon), \quad \text{para todo } t \text{ suficientemente pequeño.}$$

Por la parte (iii) del Teorema 1.2.4, tenemos que

$$f(x_\varepsilon + td) - f(x_\varepsilon) \geq t\varepsilon,$$

por lo tanto

$$Df(x_\varepsilon)d \geq -\varepsilon, \quad \forall d \in X, \|d\| = 1,$$

lo que implica que

$$\|Df(x_\varepsilon)\| = \sup_{\|d\|=1} Df(x_\varepsilon)d \leq \varepsilon.$$

Finalmente, para cada $n \in \mathbf{N}$ tomemos $y_n \in \frac{1}{n} - \arg\min f$. Luego por lo recientemente probado sabemos que existe x_n que satisface

$$f(x_n) \leq f(y_n) \leq \inf f + \frac{1}{n} \quad \text{y} \quad \|Df(x_n)\| \leq \frac{1}{n}.$$

□

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

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

son continuas para las topologías producto $\tau \times \tau$ sobre $V \times V$ y $\tau_{\mathbb{R}} \times V$ sobre $\mathbb{R} \times V$ 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 $\mathbb{R}$) 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 $\langle \cdot, \cdot \rangle : V \times V^* \to \mathbb{R}$ 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* $\langle v, \cdot \rangle : V^* \to \mathbb{R}$ es una forma lineal sobre V^*.

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 $f \colon V \to \mathbb{R}$ es un funcional lineal no nulo y $\alpha \in \mathbb{R}$.

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).

[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.

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

Demostración: Se deja como ejercicio. □

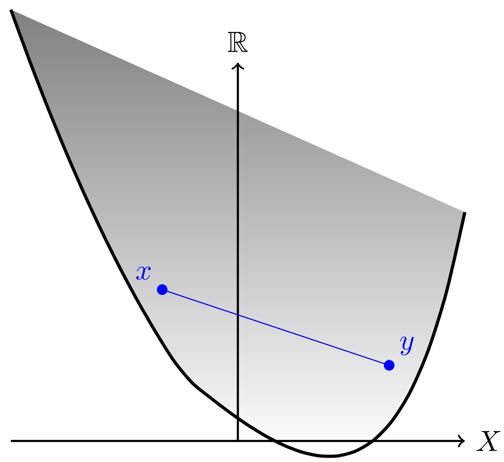


Figura 1.5: Ejemplo de función convexa

Teorema 1.3.1 (Hahn-Banach)**.** *Sean (V,τ) un e.v.t. y $A, B \subseteq V$ dos subconjuntos convexos, 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 $\alpha \in \mathbb{R}$ 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 $\ell \in (V,\tau)^*$, $\alpha \in \mathbb{R}$ 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. separado (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.

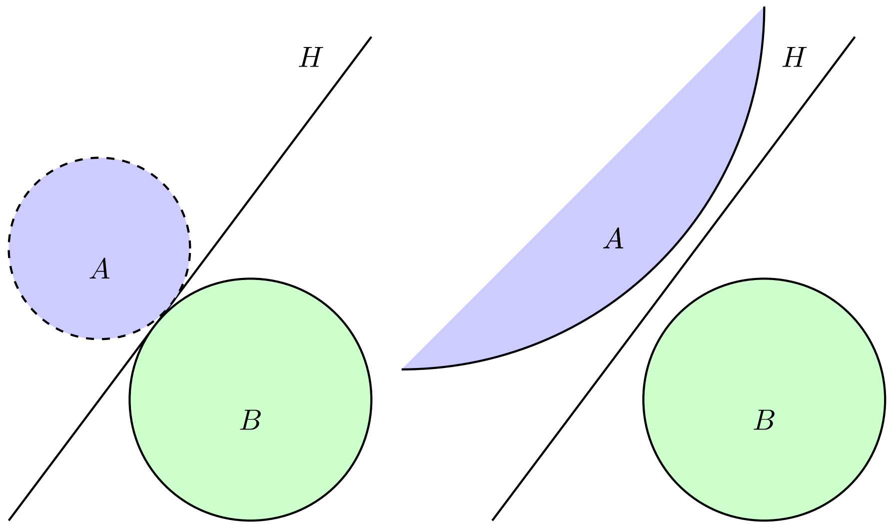


Figura 1.6: Teorema de Hanh-Banach

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 $\ell : V^* \to \mathbb{R}$ 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 $\ell \colon V^* \to \mathbb{R}$ 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

$$\{v^* \in V^* \mid \langle v_i, v^* \rangle \leq \varepsilon,\ \forall i = 1, ..., n\} \subseteq U.$$

En particular,

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

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

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

y sea $L \colon F(V^*) \to \mathbb{R}$ 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 $\mathbb{R}^n$, y en consecuencia se puede extender a $\mathbb{R}^n$ 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 $f \colon V \to \mathbb{R} \cup \{+\infty\}$ 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 $V \times \mathbb{R}$, y que f es τ-s.c.i si, y sólo si, epi(f) es cerrado para $\tau \times \tau_{\mathbb{R}}$ en $V \times \mathbb{R}$.

□

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 $f\colon V \to \mathbb{R} \cup \{+\infty\}$ 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 $f\colon X \to \mathbb{R} \cup \{+\infty\}$ 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 $(V, \|\cdot\|) = (C([0,1]; \mathbb{R}), \|\cdot\|_\infty)$, 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 $\operatorname{epi}(\delta_C) = C \times [0, +\infty[$, 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 $v_n\colon [0,1] \to \mathbb{R}$ 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 $d(0, C) = 1$.

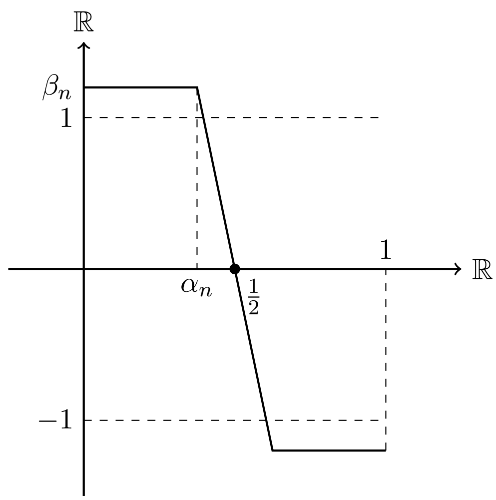


Figura 1.7: Familia de funciones v_n

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 $(C([0,1]; \mathbb{R}), \|\cdot\|_\infty)$ 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,

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

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 $f\colon X \to \mathbb{R} \cup \{+\infty\}$ 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 $f\colon X \to \mathbb{R} \cup \{+\infty\}$ es la función definida por

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

Proposición 1.5.1. *Si $f\colon X \to \mathbb{R} \cup \{+\infty\}$, 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 $f \le \mathrm{cl}_\tau(f)$ ssi $f = \mathrm{cl}_\tau(f)$.*

Demostración: (i) Notemos que $A \subseteq X \times \mathbb{R}$ 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 $\mathbb{R}$.

Así, es fácil verificar que $A = \mathrm{cl}(\mathrm{epi}(f))$ es un epígrafo, es decir, existe $g\colon X \to \mathbb{R} \cup \{+\infty\}$ 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 $g \le f$. Así, g es un minorante τ-s.c.i de f y en consecuencia $g \le \mathrm{cl}_\tau(f)$. Sea ahora $h\colon X \to \mathbb{R} \cup \{+\infty\}$ 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 $g \ge h$. 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* $f\colon X \to \mathbb{R}\cup\{+\infty\}$, *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* $f\colon X \to \mathbb{R}\cup\{+\infty\}$. *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* $f\colon X \to \mathbb{R}\cup\{+\infty\}$ *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* $\mathrm{cl}_\tau(f)$ *en* X.

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

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

□

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

Ejercicio 1.5.1. Sea $F\colon (X,\tau) \to \mathbb{R}\cup\{+\infty\}$ y $G\colon (X,\tau) \to \mathbb{R}$ 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:

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

donde $\Omega \subseteq \mathbb{R}^N$ 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 $F\colon L^p(\Omega) \to \mathbb{R}\cup\{+\infty\}$ 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 $v \in \mathrm{dom}(\overline{F})$ 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, $\mathrm{dom}(\overline{F}) \subseteq W_0^{1,p}(\Omega)$.

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, $\mathrm{dom}(\overline{F}) = W_0^{1,p}(\Omega)$ 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 $\Phi\colon L^p(\Omega)^N \to \mathbb{R}$ 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 $G\colon L^p(\Omega) \to \mathbb{R}$ 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 $\overline{\mathbb{R}}$. 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 $\overline{\mathbb{R}}$. 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 $\overline{\mathbb{R}}$ 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 $\operatorname{ínf} F > -\infty$ y para $\varepsilon > 0$ consideremos u_ε un ε-mínimo de F:

$$F(u_\varepsilon) \leq \operatorname{ínf} F + \varepsilon.$$

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

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

y, de la arbitrariedad de $\varepsilon > 0$, concluimos que $\limsup_{n\to+\infty}(\operatorname{ínf} F_n) \leq \operatorname{ínf} F$. Si $\operatorname{ínf} F = -\infty$ el argumento es similar.
Sea $u_k = u_{n_k}$ la subsucesión convergente a u tal que $u_k \in \arg\operatorname{mín}(F_{n_k})$. De la Γ-convergencia, se tiene que

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

de donde se sigue que

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

lo que implica, junto con la primera parte del teorema, que $u \in \arg\operatorname{mín} F$ y que $\operatorname{lím}_{k\to+\infty} \operatorname{ínf} F_{n_k} = \operatorname{ínf} F$. $\square$

1.6. Problemas

Problema 1.1. 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 $\mathbb{R} \cup \{+\infty\}$ definida sobre τ y $\mathcal{N}_x(\tau)$ denota el conjunto de τ-vecindades de x. Muestre que f es s.c.i.

Problema 1.2. Sean V un espacio de Banach y $U \colon V \to \mathbb{R}$ 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 1.3 (*)**.** Sea X un e.v.t. Dado $A \subseteq X$ definimos la *envoltura convexa* de A como

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

y la *envoltura convexa cerrada* de A como $\overline{\mathrm{co}}A = \mathrm{cl}(\mathrm{co}A)$.

(a) Pruebe que

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

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

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

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

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

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

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

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

Problema 1.4 (∗)**.** (Un teorema de Min-Max). Sean X y Y dos e.v.t., considere $A \subseteq X$ y $B \subseteq Y$ dos convexos, compactos no vacíos. Sea $\phi : A \times B \to \mathbb{R}$ una función que cumple las siguientes propiedades:

(i) Para cada $y \in B$, el mapeo $x \mapsto \phi(x, y)$ es continuo sobre A.

(ii) Si $\lambda_i \geq 0$ para $i = 1, ..., n$ tal que $\sum_{i=1}^n \lambda_i = 1$, entonces

$$(ii.a) \qquad \phi\left(\sum_{i=1}^n \lambda_i x_i, y\right) = \sum_{i=1}^n \lambda_i \phi(x_i, y) \quad \text{para todo } x_i \in A,$$

$$(ii.b) \qquad \phi\left(x, \sum_{i=1}^n \lambda_i y_i\right) = \sum_{i=1}^n \lambda_i \phi(x, y_i) \quad \text{para todo } y_i \in B$$

El objetivo del problema es mostrar que

$$\sup_{x\in A} \inf_{y\in B} \phi(x, y) = \inf_{y\in B} \sup_{x\in A} \phi(x, y).$$

Para ello proceda como sigue:

(a) Muestre que

$$\sup_{x\in A} \inf_{y\in B} \phi(x, y) \leq \inf_{y\in B} \sup_{x\in A} \phi(x, y).$$

(b) Denotemos por $c = \inf_{y\in B} \sup_{x\in A} \phi(x, y)$. Muestre que para probar la otra desigualdad es suficiente con probar que si

$$A_y = \{x \in A : \ \phi(x, y) \geq c\}$$

entonces $\forall I \subseteq B$ finito se tiene que $\displaystyle\bigcap_{y\in I} A_y \neq \emptyset$.

(c) Suponga que existen $y_1, ..., y_n \in B$, tales que $\displaystyle\bigcap_{i=1}^n A_{y_i} = \emptyset$. Considere el mapeo $f : A \to \mathbb{R}^n$ definido por $f(x) = (\phi(x, y_1) - c, ..., \phi(x, y_n) - c)$. Muestre que existe $\ell \in \mathbb{R}^n \setminus \{0\}$ y $\alpha \in \mathbb{R}$ tal que

$$\langle \ell, u\rangle < \alpha \leq \langle \ell, v\rangle \qquad \forall u \in f(A), \forall v \in \mathbb{R}^n_+.$$

Concluya el resultado pedido.

Problema 1.5 (*)**.** Para un conjunto $K \subseteq \mathbb{R}^n$ cualquiera definimos el cono polar de K como el siguiente conjunto,

$$K^\circ = \{y \in \mathbb{R}^n : \langle x, y\rangle \leq 0, \ \forall x \in K\}.$$

y la función soporte de K como,

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

Sea $C \subseteq \mathbb{R}^n$ no vacío y denotemos por C_∞° el cono polar de C_∞. Pruebe las siguientes relaciones:

(a) $\operatorname{dom} \sigma_C \subseteq C_\infty^\circ$.

(b) Si $\operatorname{int}(C_\infty^\circ) \neq \emptyset$, entonces $\operatorname{int}(C_\infty^\circ) \subseteq \operatorname{dom} \sigma_C$.

(c) Si C es convexo, entonces $(\operatorname{dom} \sigma_C)^\circ = C_\infty$.

Problema 1.6 (*)**.** (Algunos teorema de Punto Fijo)**.** Sea (X, d) un espacio métrico.

(a) (Generalización del teorema de punto fijo de Banach) Consideremos una función $\phi : X \to X$. Diremos que ϕ es una contracción direccional si es continua y existe $k \in (0, 1)$ tal que para cualquier $x \neq \phi(x)$ existe $z \in [x, \phi(x)] \setminus \{x\}$ y se tiene

$$d(\phi(x), \phi(z)) \leq kd(x, z).$$

Muestre que si ϕ es una contracción direccional entonces ϕ tiene un punto fijo.

(b) (Caristi) Sea $F : X \to 2^X$ una multiaplicación. Definamos el *grafo* de F como

$$\operatorname{Gr}(F) = \{(x, y) \in X \times X : y \in F(x)\}.$$

Supongamos que $\operatorname{Gr}(F)$ es cerrado y que existe una función $f : X \to \mathbb{R} \cup \{+\infty\}$ acotada inferiormente y s.c.i. tal que

$$f(y) + d(y, x) \leq f(x) \quad \forall (x, y) \in \operatorname{Gr}(F).$$

Muestre que F posee al menos un punto fijo, es decir, existe $\bar{x} \in X$ tal que $\bar{x} \in F(\bar{x})$.

Problema 1.7 (Convergencia Variacional de Puntos Sillas)**.** Diremos que una sucesión $\{F_n : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}, n \in \mathbf{N}\}$ *hipo/epi-converge* a una funcion $F : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ si para cualquier $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ 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 $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ 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 $f\colon V \to \overline{\mathbb{R}}$ se dice convexa si $\operatorname{epi} f$ es un subconjunto convexo de $V \times \mathbb{R}$. De manera equivalente, f es convexa si, y sólo si, $\forall n \geq 2$, $\forall v_1, ..., v_n \in V$, $\forall \lambda_1, ..., \lambda_n \in \mathbb{R}_+ \cup \{0\}$ tales que $\sum\limits_{i=1}^{n} \lambda_i = 1$, se tiene

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

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

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

(ii) *Si $g\colon V \to \overline{\mathbb{R}}$ 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 $\theta\colon \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ es convexa <u>no-decreciente</u> entonces $\theta \circ f$ es convexa.*

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

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

(vii) *Si $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ es convexa y $F\colon V \to \mathbb{R}^m$ es de la forma $F(x) = (f_1(x), ..., f_m(x))$ con $f_i\colon V \to \mathbb{R}$ 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. □

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 $f\colon V \to \overline{\mathbb{R}}$ 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* $f\colon V \to \mathbb{R} \cup \{+\infty\}$ *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 $M \in \mathbb{R}$ 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 $|f(x_1) - f(x_2)| \leq \frac{4\bar{M}}{\varepsilon}\|x_1 - x_2\|$.

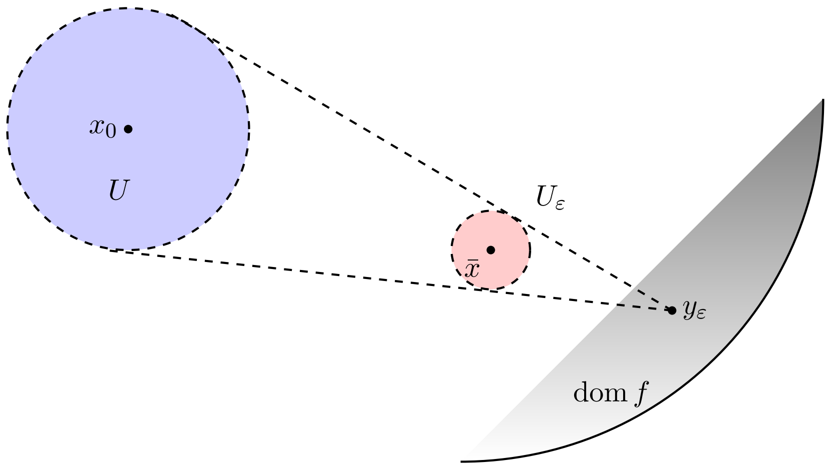
- $(ii) \Rightarrow (iii)$ y $(iii) \Rightarrow (i)$ son inmediatos.

- $(i) \Rightarrow (iv)$. Sea $\bar{x} \in \mathrm{int}(\mathrm{dom}(f))$. 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, $y_\varepsilon \in \mathrm{dom}(f)$. 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 $\gamma \in \mathbb{R}$ 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 $\forall x \in U, \; f(x) \leq \frac{a+b}{2} < +\infty$.

□

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

Corolario 2.1.1. *Si $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es convexa y propia entonces f es localmente Lipschitz en* $\text{int}(\text{dom}(f))$*. En particular, si $f\colon \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbb{R}^n$ y $x_i := x_0 + \varepsilon \hat{e}_i$ para $i = 1, ..., n$. El conjunto

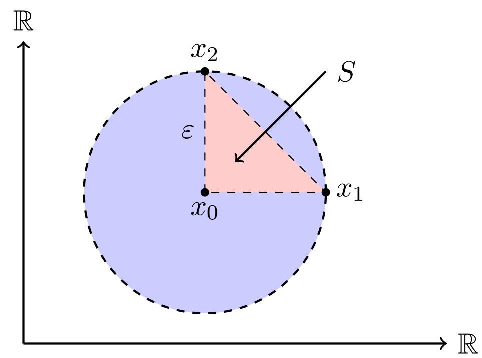
$$S = \left\{ \sum_{i=0}^n \lambda_i x_i \;\middle|\; \sum_{i=0}^n \lambda_i = 1, \; \lambda_i > 0, \; i \in \{0, 1, ..., n\} \right\}$$

es una vecindad abierta de x_0, y por convexidad

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

Por lo tanto f está acotada superiormente en S.

□



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 $\mathbb{R}^n$ pueden obtenerse de un análisis del caso $n = 1$.

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

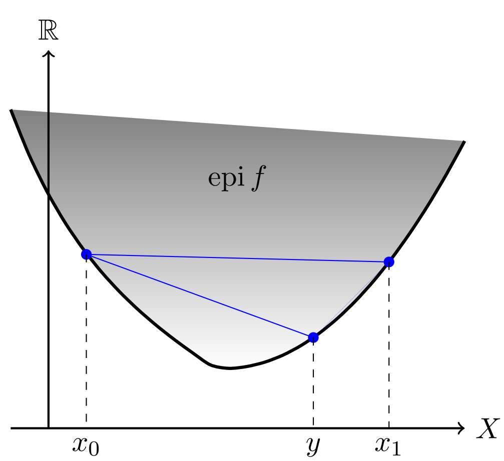
$$\frac{f(y) - f(x_0)}{y - x_0} \leq \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_1) - f(y)}{x_1 - y}.$$

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 $f : I \subseteq \mathbb{R} \to \mathbb{R}$ 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.

$\square$

Ejemplo 2.1.1. Algunas funciones convexas relevantes:

- $f(x) = ax^2 + bx + c$ en $\mathbb{R}$ cuando $a \geq 0$; es estrictamente cuando $a > 0$.
- $f(x) = e^{ax}$ en $\mathbb{R}$; 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 $U \subseteq \mathbb{R}^n$ un conjunto convexo y abierto. Si $f\colon U \to \mathbb{R}$ 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 $z \in \mathbb{R}^n$. □

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 $A \in \mathbb{R}^{n\times n}$ 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 $\langle\cdot,\cdot\rangle\colon X \times Y \to \mathbb{R}$ 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 $\alpha \in \mathbb{R}$ 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., $Y = (X, \tau_{\|\cdot\|})^* = X^*$, $\langle\cdot,\cdot\rangle\colon X \times X^* \to \mathbb{R}$ 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 Teo. 1.3.2.

- Más generalmente, si X e Y son e.v. y $\langle \cdot, \cdot \rangle \colon X \times Y \to \mathbb{R}$ 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 $f \colon X \to \overline{\mathbb{R}}$ 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 $\alpha \in \mathbb{R}$ 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 $f \colon X \to \overline{\mathbb{R}}$ es la funcion $f^* \colon Y \to \overline{\mathbb{R}}$ 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 $\mathrm{dom}(f) = \emptyset$.

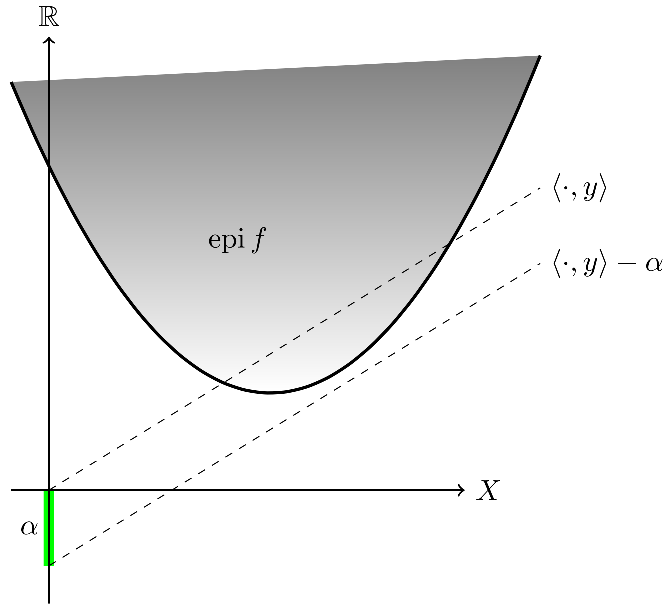


Figura 2.1: Ejemplo de minorante afín

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 $f^*(0) = -\operatorname{ínf}_X f$ y que, más generalmente, $-f^*(y) = \operatorname{ínf}_{x\in X}\{f(x) - \langle x, y\rangle\}$ 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 $\overline{\mathbb{R}}$ entonces

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

3. Dado $\lambda > 0$, $(\lambda f)^*(y) = \lambda f^*(\frac{y}{\lambda})$.

4. *Dado* $\alpha \in \mathbb{R}$, $(f+\alpha)^* = f^* - \alpha$.
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 $\mathbb{R}$. 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 $\varphi\colon \mathbb{R} \to \overline{\mathbb{R}}$ es una función convexa y $(V, \|\cdot\|)$ es una e.v.n. en dualidad con V^* entonces la conjugada de $f\colon V \to \overline{\mathbb{R}}$ definida por $f(v) = \varphi(\|v\|)$ está dada por

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

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

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

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

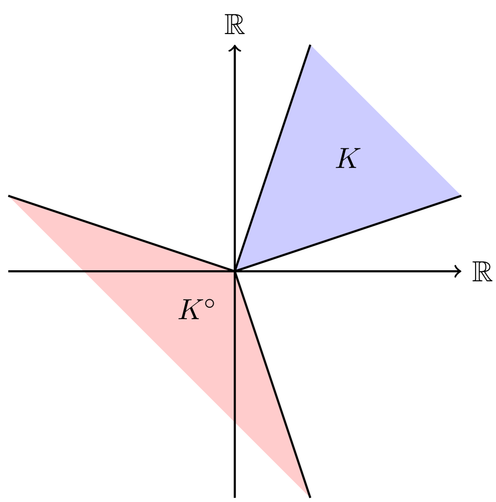


Figura 2.2: Cono polar en $\mathbb{R}^2$

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 $(X, Y, \langle\cdot,\cdot\rangle)$.

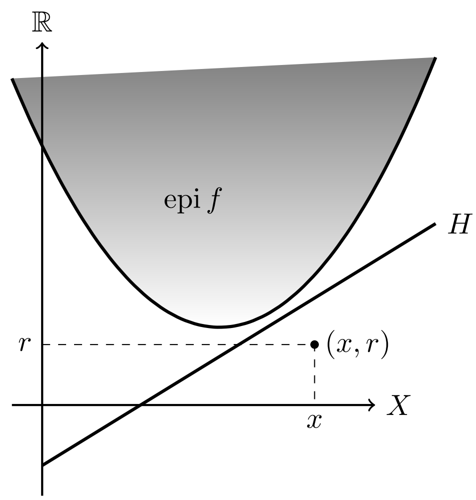
Definición 2.3.2. A continuación definiremos dos espacios muy importantes en análisis convexo.

- $\Gamma(X) = \{f : X \to \overline{\mathbb{R}} \mid f$ 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 $(y, s) \in Y \times \mathbb{R} \setminus \{(0, 0)\}$ y $\alpha \in \mathbb{R}$ 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 consecuencia, $s(f(x) - r) > 0$ 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. $x \notin \operatorname{dom}(f)$. 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

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

Razonando como antes (pues $\operatorname{dom}(f) \neq \phi$) 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 $f \colon X \to \overline{\mathbb{R}}$ se tiene que $f^* \in \Gamma(X)$.

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

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

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 $(y_i, r_i) \in Y \times \mathbb{R}$ 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 $g^{**} = g$. $\square$

Observación 2.3.2. 1. Si $\bar{f}$ denota la regularizada s.c.i. de $f\colon X \to \overline{\mathbb{R}}$, entonces

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

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

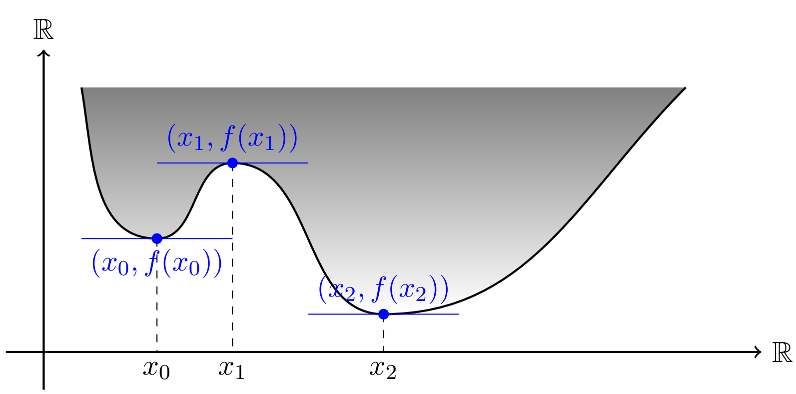
$$f(x) = \begin{cases} -\infty & x \in C, \\ +\infty & x \notin C, \end{cases}$$

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 $f\colon \mathbb{R} \to \mathbb{R}$ 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 (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

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

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 $f(x_0) \in \mathbb{R}$ para que esto ocurra. En efecto, basta considerar

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

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

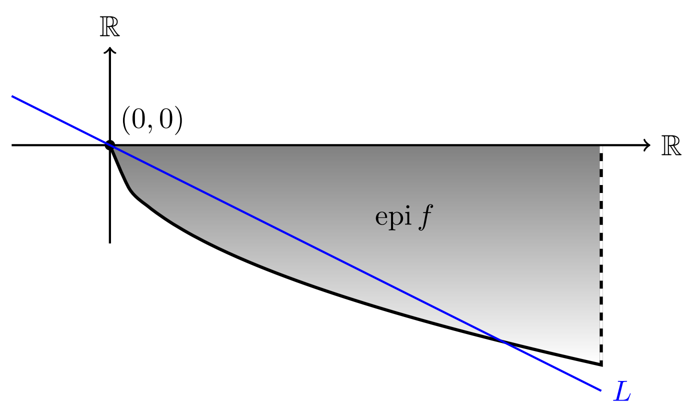


Figura 2.3: Epígrafo de la función $f(x) = -\sqrt{x}$.

Definición 2.4.1. Sean $f \colon X \to \mathbb{R}$ 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.

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

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

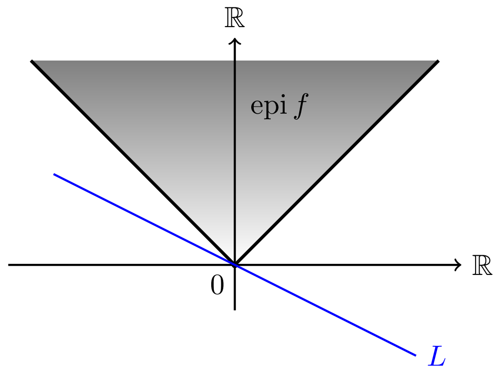


Figura 2.4: Epígrafo de la función $f(x) = |x|$.

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 $f \colon X \to \mathbb{R} \cup \{+\infty\}$ 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) $f(x) + f^*(y) = \langle x, y\rangle$.

M'as 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 $\Omega \subseteq \mathbb{R}^n$ un abierto convexo no vacío y $f\colon \Omega \to \mathbb{R}$ una función convexa tal que $\nabla f\colon \Omega \to U \subseteq \mathbb{R}^n$ 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 $f\colon X \to \mathbb{R} \cup \{+\infty\}$ 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. $\square$

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 $f\colon X \to \overline{\mathbb{R}}$ convexa y $x_0 \in X$ tales que $f(x_0) \in \mathbb{R}$ 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, $\operatorname{epi}(f)$ es un convexo de interior no vacío. En particular, $(x_0, f(x_0))$ no pertenece al convexo $\operatorname{int}(\operatorname{epi}(f))$ 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 $\operatorname{dom}(f) \neq \emptyset$ 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, $\|x^*\|_* \leq L$. $\square$

Teorema 2.4.2 (Regla de Fermat I). *Sea $f\colon X \to \mathbb{R}\cup\{+\infty\}$ 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 $f\colon X \to \mathbb{R}\cup\{+\infty\}$ convexa y propia. Sean $x_0 \in \operatorname{dom}(f)$ y $d \in X$. Entonces la derivada direccional en x_0 según d está definida en $\overline{\mathbb{R}}$ 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 $f'(x_0; d) \in \mathbb{R}, \ \forall d \in X$. 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 $x_0 \in \operatorname{dom}(f)$ 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 $f \in \Gamma_0(\mathbb{R}^n)$ y $x_0 \in \operatorname{dom}(f)$ 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 $L \in \overline{\mathbb{R}}$ 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 $x_0 \in \operatorname{dom}(f)$, 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 $f\colon U \to \mathbb{R}$ 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 $f\colon X \to \mathbb{R}\cup\{+\infty\}$ 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 $A\colon X \to 2^Y$ suelen llamarse *multiaplicaciones* y muchas veces son denotadas por $A\colon X \rightrightarrows Y$. Cuando una multiaplicación tiene la propiedad anterior, decimos que es *monótona*.

Más aún, si definimos el grafo de una multiaplicación como el siguiente conjunto

$$\mathrm{Gr}(A) := \{(x, x^*) \in X \times Y : x^* \in A(x)\},$$

entonces diremos que la multiaplicación monótona es *maximal* si su grafo no está estrictamente contenido en el grafo de otra multiaplicación monótona.

Luego, es natural preguntarse si el subdiferencial de una función en la clase $\Gamma_0(X)$ es un operador maximal monótono. La respuesta es afirmativa al menos en el caso que X es un espacio de Banach y su demostración está estrechamente relacionada con el Principio Variacional de Ekeland. Existe una demostración más sencilla en el caso que X es un espacio de Hilbert que basa en el Teorema de Minty y en la regularizada de Moreau-Yosida; ver [BaCo11, Capítulo 21] y ejercicio (2.7), respectivamente.

Otra pregunta interesante que surge de lo anterior es saber si todo operador maximal monótono corresponde al subdiferencial de una función en la clase $\Gamma_0(X)$. La respuesta en este caso es negativa, sin embargo existe una subclase de operadores maximales monótonos que responden afirmativamente a esta pregunta; ver ejercicio (2.12).

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

Ejercicio 2.4.1.

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

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

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, $\mathrm{int}(C_1) = \{(y, \lambda) \in X \times \mathbb{R} \mid g(y) < \lambda\} \neq \emptyset$ 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 $\alpha \in \mathbb{R}$. 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 $x_0 \in \operatorname{int}(C)$ 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 $x_0 \in \operatorname{int}(C)$ 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 $y_0 \in \mathrm{dom}(f)$ 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

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

que resulta ser un hiperplano cerrado de $Y \times \mathbb{R}$. Notemos que $(y, \lambda) \in S \cap \mathrm{epi}(f)$ 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 $\mathrm{int}(\mathrm{epi}(f)) \neq \emptyset$ y obviamente $S \cap \mathrm{int}(\mathrm{epi}(f)) = \emptyset$. En consecuencia, podemos separar S de $\mathrm{int}(\mathrm{epi}(f))$ 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 $\alpha \in \mathbb{R}$. 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 2.1. En lo que sigue, X es un e.v.n. y todas las funciones están en $\Gamma_0(X)$. Siguiendo la Definición 1.2.5, denotamos por f_∞ la función de recesión de f.

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

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

(b) Pruebe que $\left(\sup_{i\in I} f_i\right)_\infty = \sup_{i\in I} \{(f_i)_\infty\}$.

(c) 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 $C_\infty = \{\ d \in S_\infty \mid (f_i)_\infty(d) \leq 0\ \forall i \in I\ \}$.

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

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

Problema 2.2 (∗). Sea X un e.v.n. en dualidad con X^*, y $f \in \Gamma_0(X)$. Pruebe que las siguientes afirmaciones son equivalentes.

(a) f es coerciva;

(b) $\liminf_{\|x\|\to+\infty} \frac{f(x)}{\|x\|} > 0$.

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

La equivalencia entre (a) y (c) se conoce como *Teorema de Moreau*. Muestre además que si f es coerciva, entonces $\bigcup_{x\in X} \partial f(x)$ es una vecindad del origen.

Problema 2.3. 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 2.4 ($*$)**.** Sea $\mathcal{S}^n$ el conjunto de matrices simétricas de orden n. Dotamos a $\mathcal{S}^n$ de el producto escalar $\langle A, B\rangle = \mathrm{tr}(AB)$. Considere la función $f : \mathcal{S}^n \to \mathbb{R}^n$ dada por

$$f(A) := \frac{1}{n}\|A\|^2 - \frac{1}{n^2}\,\mathrm{tr}(A)^2.$$

(a) Pruebe que f es convexa sobre $\mathcal{S}^n$.

(b) Muestre que $\mathrm{dom}\, f^* = \{S \in \mathcal{S}^n : \ \mathrm{tr}(S) = 0\}$ y calcule f^*.

Problema 2.5. Sea $(X, Y, \langle\cdot,\cdot\rangle)$ una dualidad entre e.v.t.l.c. Dadas dos funciones convexas $f, g\colon X \to \overline{\mathbb{R}}$, se define la *inf-convolución* de f y g mediante

$$(f\Box g)(x) := \mathrm{ínf}\{f(x_1) + g(x_2) \mid x_1 + x_2 = x\}.$$

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

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

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

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

Problema 2.6 ($*$)**.** Sea $f : \mathbb{R}^n \to \mathbb{R}$ convexa. Diremos que f es $\beta-$fuertemente convexa si y sólo si para todo $x, y \in \mathbb{R}^n$, $\lambda \in (0{,}1)$

$$\frac{1}{2}\beta\lambda(1-\lambda)\|x-y\|^2 + f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$$

(a) Pruebe que las siguientes son equivalentes:

(a.1) f es $\beta-$fuertemente convexa

(a.2) $f(x) + \langle x^*, y - x\rangle + \dfrac{\beta}{2}\|y - x\|^2 \le f(y)$, para todo $x^* \in \partial f(x)$, $y \in \mathbb{R}^n$

(a.3) $\beta\|x - y\|^2 \le \langle x^* - y^*, x - y\rangle$, para todo $x^* \in \partial f(x)$, $y^* \in \partial f(y)$.

(b) Pruebe que f es $\beta-$fuertemente convexa si y sólo si $\mathrm{dom}\, f^* = \mathbb{R}^n$, que f^* es diferenciable y que ∇f^* es Lipschitz de constante $\frac{1}{\beta}$.

Indicación: para la demostrar $(\Leftarrow)$ muestre que

$$f(y) \ge f(x) + \sup_{y^* \in \mathbb{R}^n}\left\{\langle y - x, y^*\rangle - \frac{1}{2\beta}\|x^* - y^*\|^2\right\}$$

Problema 2.7 (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 2.8 (*)**.** Sea X un e.v.n. y $f \in \Gamma_0(X)$. Decimos que f es

1. *Esencialmente suave* si $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$ y $\partial f(x)$ es un síngleton para todo $x \in \operatorname{dom} \partial f$.

2. *Esencialmente convexa estricta* si f es estric. convexa en cada subconjunto convexo de $\operatorname{dom} \partial f$.

(a) Pruebe que $\partial f(x)$ es un síngleton en su dominio ssi f^* es estrictamente convexa en segmentos de $\operatorname{ran} \partial f$. Concluya que f es esencialmente suave ssi f^* es esencialmente convexa estricta.

(b) Sean $f, g \in \Gamma_0(X)$ con $\operatorname{int}(\operatorname{dom} f^*) \cap \operatorname{int}(\operatorname{dom} g^*) \neq \emptyset$. Muestre que si f es esencialmente suave entonces lo mismo ocurre para la inf-convolución $f \Box g$.

(c) Pruebe que si f es esencialmente suave, entonces $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$ y $\|\nabla f(x_n)\| \to +\infty$ para toda sucesión $\{x_n\}_n \subseteq \operatorname{int}(\operatorname{dom} f)$ que converge a un punto x en la frontera de $\operatorname{dom} f$.

Problema 2.9 (*)**.** (Función distancia)**.**

Considere que X es un espacio de Hilbert y considere $C \subseteq X$ un convexo, cerrado y no vacío.

(a) Muestre que $\operatorname{dist}_C(x) = \max_{\|x^*\|_* \leq 1} \{\langle x^*, x \rangle - \sup_{y \in C} \langle x^*, y \rangle\},\ \forall x \in X$.

(b) Caracterize los elementos de $\partial \operatorname{dist}_C(x)$, $\forall x \in X$.

Problema 2.10. (Valor propio maximal)**.** Considere $(\mathcal{S}^n, \langle \cdot, \cdot \rangle)$ como en el problema 2.4. Considere $\lambda_{\text{máx}} \colon \mathcal{S}^n \to \mathbb{R}$ la aplicación valor propio máximo.

(a) Demuestre que la función $\lambda_{\text{máx}}$ es una función convexa.
Indicación: Pruebe que $\lambda_{\text{máx}}(A) = \sup\limits_{\|v\|=1} v^\top A v$.

(b) Demuestre que la conjugada de $\lambda_{\text{máx}}$ está dada por

$$\lambda_{\text{máx}}^*(A^*) = \begin{cases} 0 & \text{si } A^* \in \mathcal{S}^n_+ \text{ y } \operatorname{tr}(A^*) = 1 \\ +\infty & \text{en caso contrario.} \end{cases}$$

(c) Demuestre que el subdiferencial de $\lambda_{\text{máx}}$ está dado por

$$\partial \lambda_{\text{máx}}(A) = \operatorname{co}\{vv^\top \colon \|v\|_2 = 1,\, Av = \lambda_{\text{máx}}(A)v\}$$

(d) Demuestre que

(d.1) $\partial \lambda_{\text{máx}}(0) = \{S \in \mathcal{S}^n_+ \colon \operatorname{tr}(S) = 1\}$.

(d.2) La derivada direccional de $\lambda_{\text{máx}}$ en $A \in \mathcal{S}^n$ en la dirección $D \in \mathcal{S}^n$ viene dada por

$$\lambda'_{\text{máx}}(A; D) = \lambda_{\text{máx}}(V^\top D V),$$

con V tal que sus columnas forman una base ortonormal del espacio propio asociado a $\lambda_{\text{máx}}(A)$.

Problema 2.11 (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 $\operatorname{dom}(\partial f)$ es denso en $\operatorname{dom}(f)$. Este resultado es conocido como el Teorema de Brondsted-Rockafellar

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 2.12 (∗). (Operadores maximales monótonos y funciones convexas).
Sea X un espacio de Banach en dualidad con X^*

(a) (Teorema de integración de Rockafellar). Diremos que un operador $T : X \rightrightarrows X$ es cíclicamente monótono si para todo $n \in \mathcal{N}$, $\{(x_i, x_i^*)\}_{i=0}^n \subseteq \mathrm{Gr}(T)$, se tiene:

$$\sum_{i=0}^{n} \langle x_{i+1} - x_i, x_i^* \rangle \leq 0 \qquad \text{con } x_{n+1} = x_0.$$

En particular, tomando $n = 1$ se tiene que T es monótono.

Sea X un espacio de Banach, si $T : X \rightrightarrows X^*$ es maximal y cíclicamente monótono, con $\operatorname{dom} T \neq \emptyset$, sea $x_0^* \in T(x_0)$, pruebe que $\exists f \in \Gamma_0(X)$ tal que $f(x_0) = 0$ y $T = \partial f$.

Indicación: considere

$$f(x) := \sup \left\{ \sum_{i=0}^{n} \langle x_i^*, x_{i+1} - x_i \rangle : \ x_{n+1} = x, \ x_i^* \in T(x_i) \ \forall i = 0, ..., n \right\}.$$

(b) Sea $T : X \rightrightarrows X$ un operador maximal monótono. Muestre que $N_{\operatorname{dom} T}(x) = (T(x))_\infty$, para todo $x \in \operatorname{dom} T$. En particular, muestre que si $f \in \Gamma_0(X)$ entonces:

(b.1) $N_{\operatorname{dom} f}(x) = (\partial f(x))_\infty$, para todo $x \in \operatorname{dom} \partial f$.

(b.2) $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$ y $x \in \operatorname{dom} \partial f \setminus \operatorname{int}(\operatorname{dom} f)$ entonces $\partial f(x)$ es no acotado.

Problema 2.13. 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 $\|y\|_* \leq K$, $\forall y \in \hat{\partial} f(x)$.

(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.)

Problema 2.14 (∗)**.** (Fórmula de Lax-Oleinik)**.** Sean $f, \theta \in \Gamma_0(\mathbf{R}^n)$ con $\theta(\cdot)$ finita y diferenciable en todo $\mathbf{R}^n$, tal que

$$\lim_{\|y\|\to\infty} \frac{\theta(y)}{\|y\|} = \infty.$$

Consideremos la ecuación de Hamilton-Jacobi

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) + \theta^*(\nabla_x u(t,x)) &= 0 \qquad && x \in \mathbf{R}^n, t > 0, \\ u(0,x) &= f(x) && x \in \mathbf{R}^n. \end{aligned}$$

Consideremos $\varphi(t,x,y) = f(x-y) + t\theta\left(\frac{y}{t}\right) + \delta_{(-\infty,0]}(t)$, se desea probar que

$$u(t,x) = \begin{cases} \inf\limits_{y\in\mathbf{R}^n} \varphi(t,x,y) & \text{si } t > 0 \\ f(x) & \text{si } t = 0 \\ +\infty & \text{si } t < 0 \end{cases}$$

es una solución de la ecuación anterior. Para ello se propone el siguiente esquema:

(a) Pruebe que φ es convexa. Deduzca que u es convexa y continua en $[0,\infty] \times \mathbb{R}^n$.

(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

$$(t^*, x^*) \in \partial u(t,x) \iff (t^*, x^*, 0) \in \partial\varphi(t,x,\bar{y}).$$

(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

$$\frac{\partial u}{\partial t}(t,x) + \theta^*(\nabla_x u(t,x)) = 0.$$

Finalmente, considere la ecuación

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) + \frac{1}{2}\|\nabla_x u(t,x)\|^2 &= 0 \qquad && x \in \mathbf{R}^n, t > 0, \\ u(0,x) &= f(x) && x \in \mathbf{R}^n. \end{aligned}$$

(f) Encuentre la fórmula explícita para la solución entregada por la fórmula de Lax-Oleinik de la ecuación.

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

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

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

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

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

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

En particular, el problema de optimización

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

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 $\varphi\colon X \times Y \to \mathbb{R} \cup \{+\infty\}$ 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 $v\colon Y \to \overline{\mathbb{R}}$ y $w\colon X^* \to \overline{\mathbb{R}}$ 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*

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

donde f y $(g_i)_i$ son funciones convexas de $\mathbb{R}^n$ en $\mathbb{R}$. 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, $\overline{v}(0) = v(0) \in \mathbb{R}$ de modo que $\overline{v}$ es propia, es decir $\overline{v} \in \Gamma_0(Y)$ y en consecuencia $\overline{v} = v^{**}$ y en particular $-\beta = v^{**}(0) = v(0) = \alpha \in \mathbb{R}$.
Mostremos ahora la necesidad. Como $\alpha + \beta = 0$ necesariamente α y β son finitos y en consecuencia $v(0) \in \mathbb{R}$. 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 $f\colon Y \to \overline{\mathbb{R}}$ es convexa, entonces

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

   En particular, si $f\colon Y \to \mathbb{R} \cup \{+\infty\}$ 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*

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

es

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

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 $v(0) \in \mathbb{R}$ 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 $v(0) \in \mathbb{R}$, 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 $\text{ínf}(\mathcal{P}) + \text{ínf}(\mathcal{D}) = 0$. □

Corolario 3.1.2. *Si $v(0) \in \mathbb{R}$ y $\partial v(0) \neq \emptyset$ entonces $S(\mathcal{D}) = \partial v(0) \neq \emptyset$ y $\text{ínf}(\mathcal{P}) + \text{ínf}(\mathcal{D}) = 0$. Del mismo modo, si $w(0) \in \mathbb{R}$ 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 $v(0) \in \mathbb{R}$, $\alpha + \beta = 0$ 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 $v(0) \in \mathbb{R}$, 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 "$f \colon X \to \overline{\mathbb{R}}$ 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 $\text{ínf}_{Y^*} v^*$, 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

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

y tomando

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

obtuvimos que

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

Evidentemente $\inf(\mathcal{D}) \geq \inf(\mathcal{P})$. Supongamos que existe $x_0 \in \mathbb{R}^n$ 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 $\alpha = \inf(\mathcal{P}) = -\inf(\mathcal{D}) = \min(\mathcal{D}) \in \mathbb{R}$ y el mínimo se alcanza; en efecto, se tiene que $\{x \in \mathbb{R}^n \mid Ax \leq b\} \neq \emptyset$ y $S(\mathcal{P}) = \{x \in \mathbb{R}^n \mid Ax \leq b\} \cap \{x \in \mathbb{R}^n \mid c^T x = \alpha\}$ 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,

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

3.2. Dualidad Lagrangeana

Consideremos el siguiente problema de programación matemática

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

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

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

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

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

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

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

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

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

□

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

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

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

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

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

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

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

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

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

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

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 $(\hat{\lambda}, \hat{\mu}) \in \mathbb{R}^p_+ \times \mathbb{R}^q$ 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,*

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

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 $\lambda_i g_i(\hat{x}) = 0,\ i = 1, \ldots, p$. $\square$

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 $Y^* := \mathbb{R}^p \times \mathbb{R}^q$ 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 $\varphi \colon X \times Y \to \mathbb{R} \cup \{+\infty\}$ 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 $\varphi \colon X \times Y \to \mathbb{R} \cup \{+\infty\}$ se define la función Lagrangeana $L \colon X \times Y^* \to \overline{\mathbb{R}}$ 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 $\overline{\mathbb{R}}$.*
- *Si φ es convexa, entonces $\forall y^* \in Y^*,\ x \mapsto L(x, y^*)$ es convexa de X en $\overline{\mathbb{R}}$, 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 $L(\hat{x},\hat{y}^*)\in\mathbb{R}$), 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. $\square$

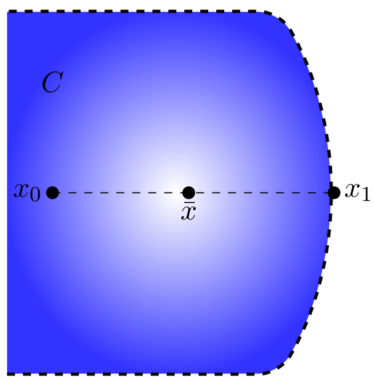
3.3. Teoremas de Dualidad de Fenchel-Rockafellar y de Attouch-Brezis

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

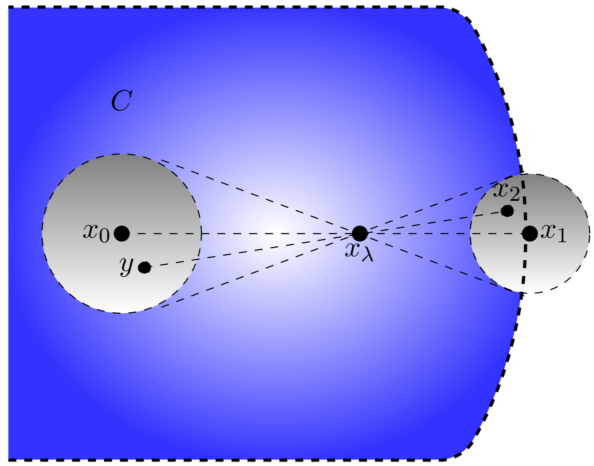
$$\exists x_2 \in C \cap \left[\frac{1}{1-\lambda}x_\lambda - \frac{\lambda}{1-\lambda}B(x_0, r)\right].$$

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 $\mathrm{int}(C) \neq \emptyset$ es esencial. Por ejemplo, si C es un hiperplano denso, entonces $\mathrm{int}(C) = \emptyset$ pero $\overline{C} = X$ de modo que $\mathrm{int}(\overline{C}) = X$.

Ejercicio 3.3.1. Si X es un e.v.n. y $C \subseteq X$ es un convexo con $\mathrm{int}(C) \neq \emptyset$, entonces

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

El interés del siguiente resultado es que no requiere saber a priori que $\mathrm{int}(C) \neq \emptyset$.

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* $\mathrm{int}(C_Y) = \mathrm{int}(\overline{C_Y})$.

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

$$\begin{aligned}
\|x_{n+1} - x_n\| &= \left\| \frac{\alpha_n}{1+\alpha_n} x_n + \frac{1}{1+\alpha_n} u - x_n \right\| \\
&= \frac{1}{1+\alpha_n} \|u - x_n\| \\
&\leq \frac{\operatorname{diam}(C_X)}{\alpha_n} = \frac{2\operatorname{diam}(C_X)}{\epsilon} \|y_n - y\|
\end{aligned}$$

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 aorbente, entonces $0 \in \operatorname{int}(C)$.*

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 $\operatorname{int}(k_0 C) \neq \emptyset$ lo que implica que $\operatorname{int}(C) \neq \emptyset$. Más aun, podemos suponer que $C = -C$ (si no, podemos tomar el convexo absorbente $C \cap (-C)$), en cuyo caso debe tenerse que $0 \in \operatorname{int}(C)$. $\square$

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 $v(0) \in \mathbb{R}$ 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 $0 \in \operatorname{int}(\overline{C_Y})$. □

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 $\inf(\mathcal{P}) \in \mathbb{R}$ *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 $\inf(\mathcal{D}) \in \mathbb{R}$ *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 $y\in\operatorname{dom}(\varphi(x,\cdot))$ si, y sólo si, $f(x)<+\infty$ y $Ax+y\in\operatorname{dom}(g)$, de modo que

$$\bigcup_{x\in X}\operatorname{dom}(\varphi(x,\cdot)) = \operatorname{dom}(g) - A\operatorname{dom}(f)$$

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. □

3.4. Teoremas de Fritz John y Kuhn-Tucker

Sea X un espacio de Banach y consideremos $f, g_i\colon X\to\overline{\mathbb{R}}$, $i=1,\dots,p$ convexas y propias, h_j, $j=1,\dots,q$ lineales afines y continuas. El programa convexo asociado a $f,(g_i)_i,(h_j)_j$ es el problema de optimización

$$(\mathcal{P})\quad \min_{x\in X}\{f(x)\mid g_i(x)\le 0, i=1,\dots,p,\quad h_j(x)=0, j=1,\dots,q\},$$

que tiene asociada la función lagrangeana $L\colon X\times\mathbb{R}^p\times\mathbb{R}^q\to\overline{\mathbb{R}}$ dada por

$$L(x,\lambda,\mu) = \begin{cases} f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \mu_j h_j(x) & \text{si } \lambda\ge 0, \\ -\infty & \text{si no,}\end{cases}$$

de modo que

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

Suponemos que $v(\mathcal{P})$ es finito. El dual asociado a L es

$$(\mathcal{D})\qquad \sup_{(\lambda,\mu)\in\mathbb{R}^p\times\mathbb{R}^q}\inf_{x\in X} L(x,\lambda,\mu),$$

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* $v(\mathcal{P}) \in \mathbb{R}$, 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* $\hat{\mu}_1, \ldots, \hat{\mu}_q \in \mathbb{R}$ *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* $x^* \in X^* \backslash \{0\}$ *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

$$\overline{C} = \overline{\mathrm{int}(C)} \subseteq [\langle x^*, \cdot \rangle \geq 0].$$

□

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. □

Teorema 3.4.4. *Sea* $F \colon X \to \overline{\mathbb{R}}$ *convexa y propia con* $\mathrm{int}(\mathrm{epi}(F)) \neq \emptyset$. *Sea* $H \subseteq X$ *un subespacio vetorial de* X *y* $L \colon H \to \mathbb{R}$ *una función lineal tal que*

$$\forall x \in X,\ F(x) \geq L(x).$$

Si $0 \in \mathrm{int}(\mathrm{dom}(F) - H)$ *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 $A := \{(x, \lambda) \in X \times \mathbb{R} \mid F(x) < \lambda\}$ y $B := \{(x, \lambda) \in X \times \mathbb{R} \mid x \in H, \lambda = L(x)\}$. 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 $(x^*, s) \in X^* \times \mathbb{R}$ 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. □

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 $f_i\colon X \to \overline{\mathbb{R}},\ i = 1, \ldots, m$ 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

$$A := \{y = (y_1, \ldots, y_m) \in \mathbb{R}^m \mid \exists x \in X, f_1(x) < y_1, \ldots, f_m(x) < y_m\}.$$

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 $\nu = (\nu_1, \ldots, \nu_m) \in \mathbb{R}^m \setminus \{0\}$ 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 $x \in \bigcap_{i=1}^m \operatorname{dom}(f_i)$ 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 $x_0 \in \bigcap_{i=1}^m \operatorname{dom}(f_i)$ 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 $\mu_1, \ldots, \mu_q \in \mathbb{R}$ 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. $\square$

Demostración: [del Teorema de Fritz-John.] Tenemos que $h_j(x) = \langle x_j^*, x \rangle - \alpha_j$ con $x_j^* \in X^*$ y $\alpha_j \in \mathbb{R}$. Como $v(\mathcal{P}) \in \mathbb{R}$, 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

$$F(x) := \max_{1 \le i \le p} \{f(x) - v(\mathcal{P}), g_i(x)\} \ge 0.$$

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 $\mathrm{epi}(F)$ es no-vacío. Como

$$\mathrm{dom}(F) = D = \mathrm{dom}(f) \cap \left(\bigcap_{i=1}^{p} \mathrm{dom}(g_i) \right)$$

y hemos supuesto que $0 \in \mathrm{int}(D - H)$, deducimos que existe $\bar{x}^* \in X^*$ tal que

$$\forall x \in H, \langle \bar{x}^*, x \rangle = 0 \Rightarrow \bar{x}^* = \sum_{j=1}^{q} \bar{\mu}_j x_j^*, \text{ para algunos } \bar{\mu}_j \in \mathbb{R}$$

y

$$\forall x \in X, F(x) - \langle \bar{x}^*, x \rangle \ge 0 \Leftrightarrow \forall x \in X,\ \max_{1 \le i \le p} \{f(x) - v(\mathcal{P}) - \langle \bar{x}^*, x \rangle, g_i(x) - \langle \bar{x}^*, x \rangle\} \ge 0.$$

Del Teorema 3.4.5, deducimos que existen $\hat{\lambda}_0, \ldots, \hat{\lambda}_p \ge 0$ no todos nulos tales que

$$\forall x \in D,\ \hat{\lambda}_0(f(x) - v(\mathcal{P})) + \hat{\lambda}_1 g_1(x) + \cdots + \hat{\lambda}_p g_p(x) - (\hat{\lambda}_0 + \ldots \hat{\lambda}_p)\langle \bar{x}^*, x \rangle \ge 0.$$

Definiendo $\hat{\mu}_j := -(\hat{\lambda}_0 + \ldots \hat{\lambda}_p)\bar{\mu}_j$, se deduce el resultado. $\square$

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

$$(\mathcal{S}) \quad \exists x_0 \in \mathrm{dom}(f)\colon g_i(x_0) < 0, i = 1, \ldots, p,\ \ h_j(x_0) = 0, j = 1, \ldots, q.$$

Entonces existe un par $(\hat{\lambda}, \hat{\mu}) \in \mathbb{R}^p_+ \times \mathbb{R}^q \setminus \{(0,0)\}$ *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* $f_1, \ldots f_m\colon X \to \overline{\mathbb{R}}$ *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 $D = \mathrm{dom}(F) = \bigcap_{i=1}^m \mathrm{dom}(f_i)$ *y*

$$\Lambda_F(x) = \{\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m_+ \mid \forall i = 1, \ldots, m\ \lambda_i(F(x) - f_i(x)) = 0\}.$$

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, $v(\mathcal{P}) \in \mathbb{R}$.

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

$$\partial \delta_H(\hat{x}) = \bigcup_{\mu \in \mathbb{R}^q} \left(\sum_{j=1}^{q} \mu_j \partial h_j(\hat{x}) \right)$$

de donde se obtiene el resultado $\square$

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 $\hat{\mu}_1, \ldots, \hat{\mu}_q \in \mathbb{R}$ 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. $\square$

3.5. Problemas

Problema 3.1 (*)**.** (Regularización de problemas min-max)**.** Considere el problema min-max

$$(P) \qquad \min_{x\in\mathbb{R}^n} \ \max\{f_1(x),\ldots,f_m(x)\}$$

con $f_i:\mathbb{R}^n\to\mathbb{R}$ convexa y $S(P)$ no vacío y acotado. Para $r>0$ considere el problema regularizado

$$(P_r) \qquad \min_{x\in\mathbb{R}^n} \ M_r(f_1(x),\ldots,f_m(x)).$$

donde $M_r(y)=rM(y/r)$ con

$$M(y_1,\ldots,y_m)=\inf_{v\in\mathbb{R}}\left\{v+\sum_{i=1}^m\theta(y_i-v)\right\}$$

y $\theta:\mathbb{R}\to[0,\infty)$ es una función estrict. convexa, creciente, diferenciable y tal que $\theta_\infty(1)=+\infty$.

(a) Muestre que M es convexa y diferenciable. Además, muestre que el ínfimo en la definición de M es alcanzado en un único punto.

(b) Pruebe que $-r\,\theta^*(1)\le M_r(y)-\max y_i\le -r\,m\,\theta^*(1/m)$.

(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 $r\to 0$. Concluir que $v(P_r)\to v(P)$ 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 3.2 (Aproximación de máxima entropía)**.** Sea $\bar u\in L^1\equiv L^1([0,1],\mathbf{R})$ una función desconocida tal que $\bar u(x)\ge 0$ 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 $u\ge 0$ 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

$$\begin{aligned}\Phi(u) &= \begin{cases}\displaystyle\int_0^1 E(u(x))dx & \text{si } E\circ u\in L^1\\ +\infty & \text{si no,}\end{cases}\\ (Au)_i &= \textstyle\int_0^1 x^i u(x)dx, \quad i=0,1,\ldots,n.\end{aligned}$$

(a) Pruebe que $\Phi \in \Gamma_0(L^1)$. 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 3.3 (*). Sea X un espacio de Banach en dualidad con X^*, su dual topológico.

(a) Calcule $\|\cdot\|^*$ y pruebe que para $x \in X \setminus \{0\}$ se tiene que

$$\partial\|\cdot\|(x) = \{x^* \in X^* : \ \|x^*\| = 1, \langle x^*, x\rangle = \|x\|\}.$$

En lo que sigue, sea K un espacio topológico Hausdorff y compacto. Denotemos por $\mathcal{C}(K)$ el conjunto de funciones continuas de K en $\mathbb{R}$ 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 $\mu : \mathcal{B}(K) \to \mathbb{R}$ con la norma $\|\mu\| = |\mu|(K) < \infty$ donde

$$|\mu|(A) = \mu^+(A) + \mu^-(A) \qquad \forall\, A \in \mathcal{B}(K).$$

El producto de dualidad entre $\mathcal{C}(K)$ y $\mathcal{M}(K)$ es $\langle \mu, f\rangle = \int_K f\, d\mu$.

(b) Una medida μ se dice concentrada en $A \in \mathcal{B}(K)$ si $|\mu|(K \setminus A) = 0$. Para $f \in C(K) \setminus \{0\}$ denotamos $K_f^+ = \{t \in K : \ f(t) = \|f\|_\infty\}$ y $K_f^- = \{t \in K : \ f(t) = -\|f\|_\infty\}$. Pruebe que

$$\partial\|\cdot\|_\infty(f) = \{\mu \in \mathcal{M} : \ \|\mu\| = 1,\ \mu^+ \text{ y } \mu^- \text{ están concentradas en } K_f^+ \text{ y } K_f^-, \text{ respectivamente}\}.$$

(c) Sean $f, \psi_1, \ldots, \psi_n \in \mathcal{C}(K)$, supongamos que f no es generado por las funciones ψ_i y consideremos el problema de mejor aproximación en el sentido de Tchebychef

$$(P) \qquad \min_{x\in\mathbb{R}^n} \|f - \textstyle\sum_{i=1}^n x_i\psi_i\|_\infty.$$

Calcule el dual de (P) asociado a la perturbación $\varphi(x, y) = \|f + y - \sum_{i=1}^n x_i\psi_i\|_\infty$.

(d) Pruebe que $v(P) + v(D) = 0$ y que $S(D) \neq \phi$.

(e) Suponiendo que $\psi_1, \ldots, \psi_n$ son l.i. demuestre que el operador $A : \mathcal{M}(K) \to \mathbb{R}^n$ definido por

$$A(\mu) = \left(\int_K \psi_1\, d\mu, \ldots, \int_K \psi_n\, d\mu\right)$$

es sobreyectivo. Deduzca que en tal caso (P) admite soluciones.

(f) Explicite las relaciones de extremalidad que caracterizan las soluciones óptimas primal y dual.

Problema 3.4 (Algunos resultados en dualidad Lagrangeana.)**.** A lo largo de esta pregunta, X e Y serán dos espacios vectoriales normados en dualidad con X^* e Y^*, respectivamente. Pruebe que:

(a) Para un esquema de dualidad de perturbaciones con $\varphi \in \Gamma_0(X \times Y)$, considere el Lagrangeano

$$L(x, y^*) := -[\varphi(x, \cdot)]^*(y^*).$$

Entonces las siguientes aseveraciones son equivalentes:

(i) $\hat{x}$ es solución del primal e $\hat{y}^*$ solución del dual, y además $\alpha + \beta = 0$.

(ii) $\hat{x}$ es solución del primal e $\hat{y}^*$ multiplicador de Lagrange asociado a $\hat{x}$, y además $\alpha + \beta = 0$.

(iii) $(\hat{x}, \hat{y}^*)$ es un punto silla del Lagrangeano L, con $L(\hat{x}, \hat{y}^*) \in \mathbb{R}$.

(b) *Teorema de Min-Max de Von Neuman.* Considere ahora dos compactos no vacíos $K \subset X$ e $M \subset Y^*$ y una función Lagrangeana dada por $L : X \times Y^* \to \mathbb{R}$ que satisface lo siguiente:

(i) Para todo y, la función $L(\cdot, y)$ es convexa y semi-continua inferior.

(ii) Para todo x, la función $L(x, \cdot)$ es concava y semi-continua superior.

Entonces

$$\alpha_L := \min_{x \in K} \max_{y \in M} L(x, y) = \max_{y \in M} \min_{x \in K} L(x, y) =: \beta_L.$$

Problema 3.5 (∗)**.** (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 $S(P)$ necesariamente es no vacío.

(b) Pruebe que $\bar{x} \in C$ es solución óptima de (P) si, y sólo si, $\bar{z} := A^* A \bar{x}$ satisface la inecuación variacional:

$$\langle \bar{x} - x, \bar{z} \rangle \leq 0, \ \forall x \in C.$$

(c) Suponga que el conjunto C tiene la siguiente forma,

$$C = \{x \in X \mid \langle a^i, x \rangle \leq b_i, \ i = 1, \ldots, m\}.$$

Verifique que el conjunto de soluciones de (P) es no vacío.

(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,

$$\exists \lambda \in \mathbb{R}^m_+ \text{ tal que } A^* A \bar{x} = -\sum_{i=1}^{m} \lambda_i a^i \text{ y además } \lambda_i(\langle a^i, x \rangle - b_i) = 0, \ i = 1, \ldots m.$$

Problema 3.6 (Dualidad de Toland-Singer)**.** En lo que sigue, denotamos por $\dot{+}$ la inf-adición en $\overline{\mathbb{R}}$ (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 $f, g : X \to \mathbb{R} \cup \{+\infty\}$, 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 definen el *exceso de Hausdorff* de C sobre D y la *distancia de Hausdorff* entre C y D, respectivamente por,

$$e(C,D) := \sup_{x\in C} d(x,D) \in [0,+\infty]$$
$$h(C,D) := \max\{e(C,D), e(D,C)\} \in [0,+\infty]$$

con $d(x,D) = \inf_{y\in D} \|y-x\|$. Usando (b), pruebe que

$$e(C,D) = \sup_{\|x^*\|\leq 1} \{\sigma_C(x^*)\underset{\cdot}{-}\sigma_D(x^*)\}.$$

Indicación: 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 Hörmander*:

$$h(C,D) = \sup_{x^*\in B^*} |\sigma_C(x^*) - \sigma_D(x^*)|.$$

Problema 3.7 (∗)**.** (Un caso particular de la dualidad de Ekeland-Lasry)**.** 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 $\Phi, \Psi : H \to \mathbb{R} \cup \{+\infty\}$ mediante

$$\Phi(x) = \frac{1}{2}\langle x, Ax\rangle + f(x) \quad \text{y} \quad \Psi(x) = \frac{1}{2}\langle x, Ax\rangle + f^*(-Ax).$$

Note que Φ y Ψ no son necesariamente convexas.

(a) Pruebe que si $\bar{x} \in H$ es un *punto crítico* para Φ, es decir, $0 \in \partial\Phi(\bar{x})$, entonces todo $\bar{y} \in \ker A + \bar{x}$ es un punto crítico para Ψ, es decir, $0 \in \partial\Psi(\bar{y})$.

(b) Demuestre que si 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 aún $\Phi(\bar{x}) + \Psi(\bar{y}) = 0$.

Problema 3.8 (∗)**.** Sea X un e.v.n. y X^* su dual. Dado un s.e.v. $M \subseteq X$ denotamos $M^\perp$ el conjunto de los $x^* \in X^*$ tales que $\langle x^*, x\rangle = 0$ para todo $x \in M$. Un funcional $x^* \in X^*$ se dice *alineado* con un vector $x \in X$ si $\langle x^*, x\rangle = \|x\|\|x^*\|_*$.

(a) Dado $z^* \in X^*$ muestre que

$$\inf_{x^*\in M^\perp} \|x^* - z^*\|_* = \sup_{x\in M,\|x\|\le 1} \langle z^*, x\rangle \tag{3.2}$$

y que el mínimo alcanzado en algún $x_0^* \in M^\perp$. Pruebe además que $x_0 \in M$ alcanza el supremo ssi $z^* - x_0^*$ está alineado con x_0.

(b) Sea $D = \{x^* \in X^* : \langle x^*, x_i\rangle = a_i,\ i = 1\ldots, n\}$ con $x_1, \ldots, x_n \in X$ y $a_1, \ldots, a_n \in \mathbb{R}$ dados. Suponiendo que D es no vacío, pruebe que

$$\min_{x^*\in D} \|x^*\|_* = \max_{y\in\mathbb{R}^n} \{\textstyle\sum_{i=1}^n a_i y_i : \|\sum_{i=1}^n y_i x_i\| \le 1\} \tag{3.3}$$

y que el x^* óptimo está alineado con $\sum_{i=1}^n y_i x_i$ donde y alcanza el máximo.

(c) Se desea calcular el flujo de combustible $u(t)$ para un cohete en ascenso vertical hasta alcanzar una altura h en tiempo T minimizando el consumo $\int_0^T |u(t)|dt$. La altura $x(t)$ satisface $\ddot{x}(t) = u(t) - g$ con $x(0) = \dot{x}(0) = 0$, de modo que integrando por partes el problema se plantea como

$$\min_u \{\textstyle\int_0^T |u(t)|dt : \int_0^T (T - t)u(t)dt = h + \frac{1}{2}gT^2\}$$

Si bien el problema se formula naturalmente en $L^1([0, T])$, este espacio no es reflexivo ni tampoco un dual de modo que no podemos garantizar la existencia de mínimos. Relajamos el problema identificando los controles $u(\cdot)$ con derivadas de funciones de variación acotada $v : [0, T] \to \mathbb{R}$ continuas por la derecha con $v(a) = 0$. Este espacio $NBV([0, T])$ se identifica con el dual de $C([0, T])$ mediante el pareo de dualidad $\langle x, v\rangle = \int_a^b x(t)\, dv(t)$, e incluye a L^1 a través de la identificación $dv(t) = u(t)dt$. La norma dual viene dada por la variacion total $\|v\|_* = \int_0^T |dv(t)|$, de modo que el problema relajado resulta ser

$$\alpha = \min_{v\in NBV} \{\textstyle\int_0^T |dv(t)| : \int_0^T (T - t)dv(t) = h + \frac{1}{2}gT^2\}.$$

Usando la parte (b) pruebe que este último mínimo es alcanzado y que

$$\alpha = \max_{y\in\mathbb{R}} \{y(h + \tfrac{1}{2}gT^2) : \|(T - t)y\|_\infty \le 1\|\} = \tfrac{1}{T}(h + \tfrac{1}{2}gT^2)$$

(d) Probar que $v \in NBV([0, T])$ está alineada con $x \in C([0, T])$ ssi v solo varía en el conjunto Γ de puntos $t \in [0, T]$ tal que $|x(t)| = \|x\|_\infty$, con $v(t)$ no-decreciente si $x(t) > 0$ y no-creciente si $x(t) < 0$. Deducir que si Γ es finito, entonces un funcional alineado v es constante por pedazos con un número finito de discontinuidades de salto.

(e) Deduzca que el control óptimo v de la parte (c) solo puede variar en $t = 0$ de modo que v es una función escalón y $u = dv$ es un control impulsivo en $t = 0$. Determine el horizonte T que minimiza α y encuentre el control óptimo v correspondiente.

Problema 3.9. Consideremos el problema de optimización

$$\text{(P)} \qquad \inf f(x) \ ; \ x \in C := \{x \in X \ : \ f_i(x) \leq 0 \quad i = 1, ..., m\}$$

donde $f, f_i : X \to \mathbb{R} \cup \{+\infty\} \cup \{\pm\infty\}$ son convexas y propias, que satisfacen $\text{int}(C) \cap \text{dom} f \neq \emptyset$. Sea $L : X \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\} \cup \{+\infty\}$ el lagrangeano usual asociado a (P). Consideremos el siguiente conjunto

$$\Lambda(P) = \left\{ \lambda \in \mathbb{R}^m_+ \ : \ \inf_{x \in X} L(x, \lambda) = \inf_{x \in X} (f + \Psi_C)(x) \in \mathbb{R} \right\}.$$

(a) Demuestre que si $\bar{x}$ es mínimo de f en C, entonces $\exists \bar{\lambda} \in \Lambda(P)$ tal que $\bar{x}$ es mínimo de $L(\cdot, \bar{\lambda})$.

(b) Definamos el conjunto

$$B = \{(x, y) \in X \times \mathbb{R}^m \ : \ f_i(x) \leq y_i, \forall i = 1, ..., m\},$$

y las funciones

$$F(x, y) = f(x) + \Psi_B(x, y) \quad \text{y} \quad g(y) = \inf_{x \in X} F(x, y).$$

Demuestre que la función g es convexa y que se tiene que $-\partial g(0) = \Lambda(P)$.

Indicación: Para la inclusión $\subseteq$, pruebe que $g'(0; d) \leq 0$, $\forall d \in \mathbb{R}^m$, de donde se concluye la positividad de las componentes de un elemento en $-\partial g(0)$.

Problema 3.10. (Karush-Kuhn-Tucker en el caso diferenciable no convexo)

(a) Consideremos una función $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable y acotada inferiormente. Pruebe que para cada $\varepsilon > 0$ existe $x_\varepsilon \in \mathbb{R}^n$ 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 $a^0, a^1, ..., a^m \in \mathbb{R}^n$, 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 $i = 0, 1, ..., m$, $d \in \mathbb{R}^n$.

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 $g_0, g_1, ..., g_m : C \subset \mathbb{R}^n \to \mathbb{R}$ funciones continuas en C y diferenciables en $\bar{x} \in int(C)$. Se definen

$$g(x) := \max_{0 \leq i \leq m} \{g_i(x)\} \quad \text{y} \quad J := \{i \mid g_i(\bar{x}) = g(\bar{x})\}.$$

Pruebe que $\forall d \in \mathbb{R}^n$:

(c.1) $\liminf\limits_{t \to 0^+} \frac{1}{t} [g(\bar{x} + td) - g(\bar{x})] \geq \langle \nabla g_i(\bar{x}), d \rangle$, $\forall i \in J$.

(c.2) $\limsup_{t\to 0^+} \frac{1}{t}\,[g(\bar{x}+td)-g(\bar{x})] \leq \max_{i\in J}\{\langle \nabla g_i(\bar{x}), d\rangle\}$.

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

$$\frac{1}{t_k}(g(\bar{x}+t_k d)-g(\bar{x})) \geq \max_{i\in J}\{\langle \nabla g_i(\bar{x}), d\rangle\} + \varepsilon.$$

(c.3) $g'(\bar{x}; d) = \max_{i\in J}\{\langle \nabla g_i(\bar{x}), d\rangle\}$.

(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 $f, g_1, ..., g_m : C \subset \mathbb{R}^n \to \mathbb{R}$ 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 $\lambda_0, \lambda_i \in \mathbb{R}_+$ $(i \in I(\bar{x}))$, no todos nulos, que satisfacen

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i\in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0. \tag{3.4}$$

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 Mangasarian-Fromovitz :

$$(MF) \qquad \exists d \in \mathbb{R}^n : \forall i \in I(\bar{x}), \langle \nabla g_i(\bar{x}), d\rangle < 0,$$

entonces es posible tomar $\lambda_0 = 1$ en (3.4).

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 $\Omega \subset \mathbb{R}^N$ 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 $F\colon H_0^1(\Omega) \to \mathbb{R}$ definida por $F(u) = -\int_\Omega f(x)u(x)dx$ y $G\colon L^2(\Omega)^N \to \mathbb{R}$ 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 $\Omega \subset \mathbb{R}$ 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 $\Omega \subset \mathbb{R}^3$ 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 $\operatorname{div}(v) = 0$.

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

$$\begin{array}{rcccl} F & : & X & \longrightarrow & \mathbb{R} \\ & & v & \mapsto & \frac{1}{2}\|v\|^2 - \int_\Omega fv, \end{array} \qquad \begin{array}{rcccl} G & : & Y & \longrightarrow & \mathbb{R} \\ & & y & \mapsto & \delta_{\{0\}}(y), \end{array} \qquad \begin{array}{rcccl} A & : & X & \longrightarrow & \mathbb{R} \\ & & v & \mapsto & \operatorname{div} v. \end{array}$$

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

$$\begin{aligned} F^*(-A^*p^*) &= \sup_{v\in X}\left\{\langle -p^*, \operatorname{div}(v)\rangle_{L^2(\Omega)} - \frac{1}{2}\|v\|^2 + \int_\Omega fv\right\} \\ &= -\inf_{v\in X}\{\frac{1}{2}\|v\|^2 - \int_\Omega fv + \int_\Omega p^* \operatorname{div}(v)\}. \end{aligned}$$

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:

$$\forall h = (h_1,h_2,h_3) \in X, \quad ((v_{p^*},h)) - \langle f,h\rangle_{L^2(\Omega)} + \langle p^*, \operatorname{div}(h)\rangle_{L^2(\Omega)} = 0.$$

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

$$F^*(-A^*p^*) = \frac{1}{2}\sum_{i=1}^3 \|\nabla v_{p^*}\|^2_{L^2(\Omega)} = \frac{1}{2}\|v_{p^*}\|^2_{H_0^1(\Omega)^3},$$

de modo que podemos reescribir el dual como

$$(\mathcal{D}) \quad \inf_{p^*\in L^2(\Omega)}\left\{\frac{1}{2}\|v_{p^*}\|^2_{H_0^1(\Omega)^3} \;\middle|\; v_{p^*} \in H_0^1(\Omega)^3 \text{ es solución de } -\Delta v_{p^*} - \nabla p^* = f \text{ en } \Omega\right\}.$$

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 $\operatorname{ínf}(\mathcal{D}) = -\operatorname{mín}(\mathcal{P}) \in \mathbb{R}$. 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 $L^2(\Omega)/\mathbb{R}$. Pasando a una subsucesión si fuese necesario, tenemos que $(p_n^*)_n$ converge débilmente hacia $\bar{p}^* \in L^2(\Omega)/\mathbb{R}$ 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:

$$(\mathcal{P}) \quad \operatorname*{ínf}_{v \in C} \left\{ \frac{1}{2} \int_\Omega |\nabla v|^2 dx - \int_\Omega f v dx \right\}$$

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 $F\colon X \to \mathbb{R}$, $G\colon Y \to \mathbb{R}$ dadas por

$$F(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_\Omega f v dx \quad \text{y} \quad G(p) = \delta_S(p),$$

con $S = \{p \in L^2(\Omega)^N \mid \ |p(x)| \leq 1, \text{ en } \Omega - c.t.p.\}$. Luego, $(\mathcal{P})$ equivale a

$$\operatorname*{ínf}_{v \in X} F(v) + G(Av).$$

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

$$(\mathcal{D}) \qquad \operatorname*{ínf}_{p^* \in L^2(\Omega)^N} \left\{ \frac{1}{2}\|\operatorname{div} p^* + f\|_{H^{-1}(\Omega)} + \int_\Omega |p^*| \right\}.$$

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 $\operatorname{mín}(\mathcal{P}) = -\operatorname{ínf}(\mathcal{D})$.

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 4.1 (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 $\mathbb{R}^N$, $\alpha, \beta > 0$ 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 $\Phi : H_0^1(\Omega) \to \mathbb{R}$, $\Psi : L^2(\Omega)^N \to \mathbb{R}$ 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

$$(D) \qquad \min_{p^*\in L^2(\Omega)^N} \left\{ \frac{1}{2\alpha}\int_\Omega [|p^*(x)| - \beta]_+^2 d\lambda(x) \;\middle|\; \operatorname{div}(p^*) = -f \right\}.$$

(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, $\operatorname{div}(\bar{p}^*) = -f$ (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 4.2. 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

$$(\mathcal{P}) \quad v = \text{mín}\{f_0(x) \mid f_i(x) \leq 0, \quad i \in I\},$$

y un problema sin restricciones

$$(\mathcal{P}_r) \quad v_r = \text{mín}\left\{ f_0(x) + r\sum_{i\in I} \theta\left(\frac{f_i(x)}{r}\right) \;\middle|\; x \in \mathbb{R}^n \right\}$$

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 $f\colon \mathbb{R}^n \to \mathbb{R}$ pertenece a la clase $\mathcal{Q}$ si es convexa y satisface la siguiente propiedad:

$$(*) \qquad \begin{array}{c}\text{Si } f \text{ es constante en el segmento de recta } [x,y],\\ \text{entonces es constante en toda la recta que pasa por } x \text{ e } y.\end{array}$$

Ejercicio 5.1.1. Suponga que $f \in \mathcal{Q}$ y que el convexo $C \subseteq \mathbb{R}^n$ es tal que f es constante en C. Muestre que f es constante en $\text{aff}(C)$.

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 $\sigma\colon \mathbb{R} \to \mathbb{R}$ 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) $f_0, f_i \colon \mathbb{R}^n \to \mathbb{R}$ 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) $\theta \colon (-\infty, \kappa) \to \mathbb{R}$, 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 $\mathbb{R}$ 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 $C \subset \mathbb{R}^n$ 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 $t \in \mathbb{R}$ 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 desigualad 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 notación, 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

$$(\mathcal{P}) \quad v = \min\{\varphi(x) \mid x \in \mathbb{R}^n\},$$

donde $\varphi \in \Gamma_0(\mathbb{R}^n)$, y consideremos $x(\epsilon)$ la única solución del problema

$$v_\epsilon = \min\left\{\varphi(x) + \frac{\epsilon}{2}\|x\|^2 \mid x \in \mathbb{R}^n\right\},$$

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 $S(\mathcal{P})$:

$$x(\epsilon) \to \operatorname{Proy}_{S(\mathcal{P})}(0), \quad \epsilon \to 0^+.$$

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 $f_{r_k}(x(r_k)) \leq f_{r_k}(\widetilde{x}_k)$ se sigue que

$$\sum_{i\in I} \theta\left(\frac{f_i(x(r_k))}{r_k}\right) \leq \sum_{i\in I} \theta\left(\frac{f_i(\widetilde{x}_k)}{r_k}\right).$$

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

$$\sum_{i\notin I_0} \theta\left(\frac{f_i(x(r_k))}{r_k}\right) \leq \sum_{i\notin I_0} \theta\left(\frac{f_i(\widetilde{x}_k)}{r_k}\right).$$

Motivados por el Ejemplo 5.2.1, quisiéramos ahora, mediante un proceso límite, obtener información del tipo $\Gamma(x_0) \leq \Gamma(x^*)$, 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 $\theta(y) = e^y$, $A_\theta(y) = \max_{j=1,\ldots,k} y_j$.

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 $A_\theta \colon]-\infty, 0[^k \to \mathbb{R}$ es simétrica, positivamente homogénea, no-decreciente por componentes, convexa y satisface*

$$\frac{1}{k}\sum_{i=1}^{k} y_i \leq A_\theta(y) \leq \max_{i=1,\ldots,k} y_i.$$

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 $\lambda \in \mathbb{R}$ 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 $\delta\colon P \to \mathbb{R}$ 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 $\theta_1, \theta_2\colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ 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 $A_{\theta_1}(y)$ □

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 $\mathbb{R}_-$, 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 $\varphi_C\colon \operatorname{aff}(C)\to\mathbb{R}\cup\{+\infty\}$ 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 $\operatorname{aff}(\{x, y\})$ 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 $S^0 \supset S^1 \supset S^2 \supset \dots$. Además, definiendo $I^k = I_{S^k}$, se tiene que $I^0 \supset I^1 \supset I^2 \supset \dots$. 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 $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ 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

$$(\mathcal{D}) \qquad \min_{\lambda \geq 0} p(\lambda),$$

con $p(\lambda) = \inf_{x\in\mathbb{R}^n} f_0(x) + \sum_{i\in I} \lambda_i f_i(x)$. 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

$$(\mathcal{D}_r) \qquad \min_{\lambda\in\mathbb{R}^m} p(\lambda) + r\sum_{i\in I}\theta^*(\lambda_i).$$

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 $x \in \mathbb{R}^n$. 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 $f, g \in \Gamma_0(\mathbb{R}^n)$ tales que*

1. $\arg\min f \cap \operatorname{dom}(g) \neq \emptyset$;
2. *al menos una de ellas tiene conjuntos de nivel acotados;*
3. *g es estrictamente convexa en su dominio;*
4. $g^\infty(d) \geq 0$ $\forall d \in \mathbb{R}^n$.

Supongamos que $z(r)$ es solución de

$$\min_{z\in\mathbb{R}^n} f(z) + rg(z).$$

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 $f(z(r_k)) + r_k g(z(r_k)) \leq f(z^*) + r_k g(z^*)$ deducimos que

$$\liminf_{k\to+\infty} f(z(r_k)) + r_k g(z(r_k)) \leq f(z^*)$$

Además, es fácil ver que $g(z(r_k))$ está acotada y luego $r_k g(z(r_k)) \to 0$. En consecuencia, $f(\bar{z}) \leq f(z^*)$ y $\bar{z} \in \arg\min f$. Como además se tiene que

$$g(\bar{z}) \leq \liminf_{k\to+\infty} g(z(r_k)) \leq g(z^*),$$

deducimos finalmente que $\bar{z} \in \arg\min\{g(z) \mid z \in \arg\min f\}$. □

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 $f(\lambda) = p(\lambda) + \delta_{\mathbb{R}^m_+}(\lambda)$ 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, $p(\lambda) + \delta_{\mathbb{R}^m_+}(\lambda)$ tiene conjuntos de nivel acotados. Por otro lado, si $\kappa > 0$, entonces

$$\theta^*(\lambda) \geq \lambda y - \theta(y),$$

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 $(\theta^*)^\infty(\pm 1) \geq 0$. Pero esto último es evidente pues

$$(\theta^*)^\infty(1) = \lim_{t\to+\infty} \frac{\theta^*(t)}{t} \geq \kappa,$$

y

$$(\theta^*)^\infty(-1) = \lim_{t\to+\infty} \frac{\theta^*(0-t) - \theta^*(0)}{t} = +\infty.$$

□

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

$$(\mathcal{P}) \quad \min_{a_i x \leq b_i, i\in I} c^T x$$

con dual

$$(\mathcal{D}) \qquad \min_{c+A^T\lambda=0,\lambda\geq 0} b^T\lambda$$

y esquema de penalización

$$(\mathcal{P}_r) \quad \min c^T x + r\sum_{i\in I}\theta\Big(\frac{a_i x - b_i}{r}\Big),$$

$$(\mathcal{D}_r) \quad \min_{c+A^T\lambda=0} b^T\lambda + r\sum_{i\in I}\theta^*(\lambda_i)$$

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 5.1. Sea I finito y para $i \in I$ sean $f_i\colon \mathbb{R}^n \to \mathbb{R}$ funciones convexas. Considere el siguiente problema de minimización

$$(\mathcal{P}) \qquad \inf_{x\in\mathbb{R}^n} \max_{i\in I}\{f_i(x)\}.$$

(a) Pruebe que $x^* \in S(\mathcal{P})$ si, y sólo si, $(\mu^*, x^*) \in \mathbb{R}\times\mathbb{R}^n$, con $\mu^* = \max_{i\in I}\{f_i(x^*)\}$, es solución de

$$\inf_{(\mu,x)\in\mathbb{R}\times\mathbb{R}^n}\{\mu \mid \forall i \in I,\ f_i(x) \le \mu\}.$$

(b) Definamos $\varphi\colon \mathbb{R}^{n+1}\times\mathbb{R}^m \to \mathbb{R}\cup\{+\infty\}$ 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

$$(\mathcal{D}) \qquad \inf_{\lambda\in\mathbb{R}^n} p(\lambda),$$

donde

$$p(\lambda) = \begin{cases} -\inf_{x\in\mathbb{R}^n}\{\sum_{i\in I}\lambda_i f_i(x)\} & \text{si } \lambda \in \Delta_m, \\ +\infty & \text{si no,} \end{cases}$$

donde, recordamos, $\Delta_m = \{\lambda \in \mathbb{R}^m_+ \mid \sum_{i\in I}\lambda_i = 1\}$.

(c) Suponga que $\inf(\mathcal{P}) \in \mathbb{R}$. 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 $\bar{\varphi}_r\colon \mathbb{R}^{n+1}\times\mathbb{R}^m \to \mathbb{R}$ 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 $\varphi_r(x, y) = \min_{\mu\in\mathbb{R}} \bar{\varphi}_r(\mu, x, y) - r$. Verifique que $\varphi_r \in \Gamma_0(\mathbb{R}^n\times\mathbb{R}^m)$ 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

$$F(x) < F_r(x) \le F(x) + r\ln(m), \quad \forall x \in \mathbb{R}^n.$$

Considere la familia parametrizada de problemas

$$(\mathcal{P}_r) \qquad \inf_{x\in\mathbb{R}^n} F_r(x)$$

y en lo que sigue suponga que $S(\mathcal{P}) = \arg\min F$ es no-vacío y compacto.

(e) Pruebe que $\forall r > 0$, $\text{mín}(\mathcal{P}) \leq \text{ínf}(\mathcal{P}_r) \leq \text{mín}(\mathcal{P}) + r\ln(m)$ y que, más aún, el conjunto $S(\mathcal{P}_r) = \arg\text{mín}\, F_r$ es no-vacío y compacto.

(f) Sea $\gamma > \text{mín}(\mathcal{P})$. Pruebe que si $0 < r < \frac{1}{\ln(m)}(\gamma - \text{mín}(\mathcal{P}))$, entonces

$$\Gamma_{\text{mín}(\mathcal{P})+r\ln(m)}(F_r) \subseteq \Gamma_\gamma(F).$$

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

$$(\mathcal{D}_r) \qquad \inf_{\lambda \in \mathbb{R}^m} \left\{ p(\lambda) + r\sum_{i\in I} \lambda_i \ln(\lambda_i) \right\}$$

con la convención $0\ln(0) = 0$ y $\lambda\ln(\lambda) = +\infty$ para $\lambda < 0$. Pruebe que $\text{ínf}(\mathcal{P}_r) + \text{ínf}(\mathcal{D}_r) = 0$ y que (D_r) admite solución única $\lambda(r)$.

Para simplificar, supongamos además que $f_i \in C^1(\mathbb{R})$.

(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 $r \to 0^+$, $\text{mín}(\mathcal{D}_r) \to \text{mín}(\mathcal{D})$ 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 5.2. Sea $f \in \mathcal{Q} \cap \Gamma_0(\mathbb{R}^n)$ (ver Definición 5.1.1). Suponemos que existe $x_0 \in \text{dom}(f)$ con $\|x_0\|_\infty \leq 1$, y consideramos el problema de optimización

$$(P_\infty) \qquad \min_{x\in\mathbb{R}^n} \{f(x) : \|x\|_\infty \leq 1\}$$

junto con el esquema de aproximación (sin restricciones)

$$(P_p) \qquad \min_{x\in\mathbb{R}^n} \left\{ f(x) + \frac{1}{p}\|x\|_p^p \right\}$$

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

$$(P_\infty^1) \qquad \min_{x \in S(P_\infty)} \|x\|_\infty.$$

(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 $x \in S(P_\infty^1)$. Deduzca que, o bien $x_{i_0} = \alpha_1$ para todo $x \in S(P_\infty^1)$, o $x_{i_0} = -\alpha_1$ para todo $x \in S(P_\infty^1)$. 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

$$(P_\infty^2) \qquad \min_{x \in S(P_\infty^1)} \|\Pi_1(x)\|_\infty.$$

(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 $b \in \mathbb{R}^m$ 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

$$(Q_\infty) \quad \min_{y \in \mathbb{R}^n} \|Ay - b\|_\infty$$

el cual aproximamos mediante el esquema

$$(Q_p) \quad \min_{y \in \mathbb{R}^n} \|Ay - b\|_p.$$

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 $L = \{Ay - b : y \in \mathbb{R}^n\}$ y la función

$$f(x) := \begin{cases} 0 & \text{si } x \in L \\ +\infty & \text{si no.} \end{cases}$$

Problema 5.3. Sean $f_0, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ convexas y consideremos el esquema de aproximación

$$(P_\epsilon) \qquad \min_{x \in \mathbb{R}^n} f_\epsilon(x) := f_0(x) + \epsilon f_1(x) + \cdots + \epsilon^m f_m(x)$$

definido para $\epsilon > 0$, junto con la jerarquía de problemas de optimización

$$\begin{array}{ll} (P_0) & \min\{f_0(x) : x \in \mathbb{R}^n\} \\ (P_1) & \min\{f_1(x) : x \in S_0\} \\ \vdots & \\ (P_m) & \min\{f_m(x) : x \in S_{m-1}\} \end{array}$$

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 $x, d \in \mathbb{R}^n$.

(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 5.4. 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 $\mu \in \mathbb{R}^m$ 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.

Apéndice A

Resolución de algunos problemas

A.1. Introducción al Análisis Variacional

Problema 1.3

(a) Consideremos el conjunto

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

Es fácil ver que B es un conjunto convexo y que $A \subseteq B$, luego $\mathrm{co}(A) \subseteq B$. Sea ahora C un convexo cualquiera que contiene a A y sea $x \in B$. Luego, como $A \subseteq C$ y C es convexo entonces $x \in C$, esto es directo pues sabemos que $x = \sum \lambda_i x_i$ con $x_i \in A$, $\lambda_i \geq 0$ y $\sum \lambda_i = 1$. Luego $B \subseteq C$ para cualquier convexo que contiene a A y por lo tanto $B \subseteq \mathrm{co}(A)$.

(b) Sea $x \in \mathrm{co}(A)$, luego, sin pérdida de generalidad $x = \sum_{i=1}^{k} \lambda_i x_i$ con $x_i \in A$, $\lambda_i > 0$ y $\sum_{i=1}^{k} \lambda_i = 1$. Supongamos que $k > n+1$, dado que la dimensión del espacio es n, cualquier familiar de $l > n$ vectores será linealmente dependiente, en particular podemos considerar $l = k-1$ y los siguiente $k-1$ vectores:

$$x_2 - x_1, ..., x_k - x_1.$$

Por lo tanto, existen $\mu_2, ..., \mu_k \in \mathbb{R}$ no todos nulos tales que $\sum_{i=2}^{k} \mu_i (x_i - x_1) = 0$. Definamos $\mu_1 = -\sum_{i=2}^{k} \mu_i$, con esto tenemos que $\sum_{i=1}^{k} \mu_i x_i = 0$ y $\sum_{i=1}^{k} \mu_i = 0$. En particular, para todo $\alpha \in \mathbb{R}$ se tiene que $x = \sum_{i=i}^{k} (\lambda_i - \alpha \mu_i) x_i$. Ahora, dado que los μ_i suman cero pero no son todo nulos, al menos uno de ellos debe ser positivo. Escojamos α de la siguiente manera:

$$\alpha = \min \left\{ \frac{\lambda_i}{\mu_i} : \mu_i > 0 \right\}.$$

Denotaremos por I el índice donde se alcanza el mínimo. Luego es directo que $\lambda_i - \alpha \mu_i \geq 0$ para todo i y en particular, $\lambda_I - \alpha \mu_I = 0$. Finalmente, $x = \sum_{i=i}^{k-1} (\lambda_i - \alpha \mu_i) x_i$ con $\lambda_i - \alpha \mu_i \geq 0$, es decir, x se puede escribir como combinación convexa de $k-1$ términos. Luego, repitiendo el mismo argumento $k-(n+1)$ veces obtenemos el resultado pedido.

(c) Sean $x, y \in \overline{C}$ y $\lambda \in [0,1]$, por lo tanto existen $\{x_n\}, \{y_n\} \subseteq C$ tales que $x_n \to x$ e $y_n \to y$. Más aún, como C es convexo $\lambda x_n + (1-\lambda)y_n \in C$ y $\lambda x_n + (1-\lambda)y_n \to \lambda x + (1-\lambda)y$. Sigue que $\lambda x + (1-\lambda)y \in \overline{C}$, de donde se deduce que $\overline{C}$ es convexo.

Consideremos el siguiente conjunto,

$$B = \bigcap \{\ C \mid C \text{ es un convexo cerrado, } A \subseteq C\}.$$

Como $\overline{\text{co}}(A)$ es convexo, cerrado y contiene a A, sigue que $B \subseteq \overline{\text{co}}(A)$. Por otra parte

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

Por lo tanto $\overline{\text{co}}(A) \subseteq B$, de donde se concluye lo pedido.

(d) Sea $x \in \text{co}(A)$, luego existen $x_i \in A$, $\lambda_i > 0$ y $\sum_{i=1}^k \lambda_i = 1$, tales que $x = \sum_{i=1}^k \lambda_i x_i$. Como A es abierto, entonces existen vecindades $U_i \in \mathcal{N}_{x_i}$ tales que $U_i \subseteq A$. Sea $U = \lambda_1 U_1 + ... + \lambda_k U_k$. Finalmente, como $\text{co}(A)$ es convexo y $U_i \subseteq A$ entonces $U \subseteq \text{co}(A)$, además es claro que U es una vecindad de x, por lo tanto $\text{co}(A)$ es abierto.

Por otra parte, supongamos que C es convexo. Tomemos $x, y \in \text{int}(C)$ y $\lambda \in [0,1]$, luego existen vecindades $U \in \mathcal{N}_x$ y $V \in \mathcal{N}_y$ tales que $U, V \subseteq C$. Luego sigue que $W = \lambda U + (1-\lambda)V \subseteq C$, por convexidad. Como W es una vecindad de $\lambda x + (1-\lambda)y$ se concluye el resultado.

(e) Consideremos $\Delta = \{s \in \mathbb{R}^n : \ \sum s_i = 1,\ s_i \geq 0\}$ y $A = A_1 \times ... \times A_2$. Luego tenemos que $\Delta \times A$ es compacto. Consideremos la función $f : \Delta \times A \to X$ definida por

$$f(s,a) = \sum_{i=1}^n s_i a_i.$$

Como X es un e.v.t. tenemos que f es continua y por lo tanto $f(\Delta \times A)$ es compacto. Por otra parte, es fácil ver que $A_i \subseteq f(\Delta \times A)$ para todo $i = 1...n$, luego $A_1 \cup \cdots \cup A_n \subseteq f(\Delta \times A)$. Además, no es difícil ver que $f(\Delta \times A)$ es convexo, por lo tanto, tomando envoltura convexa en la última inclusión obtenemos que

$$\text{co}(A_1 \cup \cdots \cup A_n) \subseteq f(\Delta \times A).$$

Finalmente, por definición se tiene que $f(\Delta \times A) \subseteq \text{co}(A_1 \cup \cdots \cup A_n)$. Luego $\text{co}(A_1 \cup \cdots \cup A_n)$ es compacto.

(f) Supongamos que A es totalmente acotado, luego para $\varepsilon > 0$ existen $x_1, ..., x_n$ tales que

$$A \subseteq B(x_1, \varepsilon) \cup ... \cup B(x_n, \varepsilon) = \{x_1, ..., x_n\} + B(0, \varepsilon) \subseteq \text{co}(\{x_1, ..., x_n\}) + B(0, \varepsilon),$$

como $B(0, \varepsilon)$ es convexo se tiene que

$$\text{co}(A) \subseteq \text{co}(\{x_1, ..., x_n\}) + B(0, \varepsilon).$$

Notemos que, por la parte anterior, $\mathrm{co}(\{x_1,...,x_n\})$ es compacto, luego es totalmente acotado, y por lo tanto existen $y_1,....,y_m \in \mathrm{co}(\{x_1,...,x_n\}) \subseteq \mathrm{co}(A)$ tales que

$$\mathrm{co}(\{x_1,...,x_n\}) \subseteq \{y_1,...,y_m\} + B(0,\varepsilon).$$

Finalmente, obtenemos

$$\mathrm{co}(A) \subseteq \{y_1,...,y_m\} + B(0,2\varepsilon),$$

de donde se tiene que $\mathrm{co}(A)$ es totalmente acotado.

Supongamos que V es un espacio de Banach, luego ser totalmente acotado es equivalente a ser pre-compacto. Por lo tanto si $A \subseteq K$, con K compacto, entonces A es totalmente acotado y por lo tanto $\mathrm{co}(A)$ también lo es. Finalmente, como $\mathrm{co}(A)$ es totalmente acotado, entonces $\overline{\mathrm{co}}(A)$ es compacto pues $\mathrm{co}(A)$ es pre-compacto.

Problema 1.4

(a) Directo.

(b) Notemos que por la propiedad (i), $c < +\infty$. Supongamos que $c > -\infty$, si no, por la parte anterior la igualdad es directa. Sea $A_y \subseteq A$ como en el enunciado, por (i) se tiene que A_y es cerrado, luego por la compacidad de A si la familia $\{A_y\}$ satisface la P.I.F. (propiedad de la intersección finita) entonces $\bigcap_{y\in B} A_y \neq \emptyset$. Luego existiría $\bar{x} \in A_y$ para todo $y \in B$, es decir, $c \leq \phi(\bar{x}, y)$, para todo $y \in B$. Entonces, tomando ínfimo sobre los $y \in B$ obtendríamos el resultado buscado, pues

$$c = \inf_{y\in B} \sup_{x\in A} \phi(x,y) \leq \inf_{y\in B} \phi(\bar{x},y) \leq \sup_{x\in A} \inf_{y\in B} \phi(x,y).$$

(c) Sea f como en el enunciado. Notemos que $f(A)$ es no vacío, veamos que $f(A)$ es convexo y compacto, y que además $f(A) \cap \mathbb{R}^n_+ = \emptyset$, con esto la existencia de ℓ y α La compacidad de $f(A)$ viene de la compacidad de A y de la continuidad de $\phi(\cdot, y)$. Para probar la convexidad, definamos $\Phi(x) = (\phi(x,y_1),...,\phi(x,y_n))$ y consideremos $x_1, x_2 \in A$ y $\lambda \in [0,1]$, luego

$$\lambda f(x_1) + (1-\lambda) f(x_2) = \lambda\Phi(x_1) + (1-\lambda)\Phi(x_2) - \mathbf{1}c,$$

pero por (ii.a), se tiene que $\lambda\Phi(x_1) + (1-\lambda)\Phi(x_2) = \Phi(\lambda x_1 + (1-\lambda)x_2)$. Luego por la convexidad de A se tiene que $\lambda f(x_1) + (1-\lambda) f(x_2) \in f(A)$. Por otra parte, no existe $x \in A$ tal que $f(x) \geq 0$, pues si no $\bigcap_{i=1}^{n} A_{y_i} \neq \emptyset$. Entonces, por Hanh-Banach, existe $\ell \in \mathbb{R}^n \setminus \{0\}$ y $\alpha \in \mathbb{R}$ tal que

$$(*) \qquad \langle \ell, u\rangle < \alpha \leq \langle \ell, v\rangle \qquad \forall u \in f(A), \forall v \in \mathbb{R}^n_+$$

Luego evaluando $(*)$ en $v = 0$ tenemos que $\alpha \leq 0$, por lo que podemos suponer que $\alpha = 0$, además, evaluando $(*)$ en $v = e_i$, los vectores canónicos de $\mathbb{R}^n$, tenemos que $\ell_i \geq 0$. Por lo tanto, sin perdida de generalidad podemos suponer que $\sum \ell_i = 1$. Ahora sea $u \in f(A)$, luego existe $x \in A$ tal que $u = f(x)$, entonces

$$\langle \ell, u\rangle = \langle \ell, \Phi(x)\rangle - \langle \ell, \mathbf{1}c\rangle = \sum_{i=1}^{n} \ell_i \phi(x, y_i) - c = \phi\left(x, \sum_{i=1}^{n} \ell_i y_i\right) - c,$$

donde la última igualdad es cierta por (ii.b). Definamos $\bar{y} = \sum \ell_i y_i$, luego tenemos que $\langle \ell, u\rangle = \phi(x, \bar{y}) - c$ y por $(*)$ se tiene que $\phi(x, \bar{y}) < c$. Finalmente, tomando supremo en $x \in A$ obtenemos que

$$c = \inf_{y\in B} \sup_{x\in A} \phi(x, y) \leq \sup_{x\in A} \phi(x, \bar{y}) < c.$$

Lo que es una contradicción, por lo tanto la familia $\{A_y\}$ tiene la P.I.F.

Problema 1.5

(a) Sea $y \notin C_\infty^\circ$, luego existe $d \in C_\infty \setminus \{0\}$ tal que $\langle d, y\rangle > 0$. Por definición del cono de recesión, existen $\{t_k\} \subseteq \mathbb{R}^+$ y $\{x_k\} \subseteq C$ tales

$$\frac{1}{t_k} x_k \to d \text{ y } t_k \to +\infty,$$

luego

$$\langle x_k, y\rangle = t_k \langle \frac{1}{t_k} x_k, y\rangle \to +\infty.$$

Por lo tanto $y \notin \operatorname{dom} \sigma_C$.

(b) Sea $y \notin \operatorname{dom}_C$. Luego existe $x_k \subseteq C$ tal que $\langle x_k, y\rangle \to +\infty$. Sea $d_k = \frac{x_k}{\|x_k\|}$, sin perdida de generalidad podemos suponer que $d_k \to d \neq 0$, más aún, por definición, $d \in C_\infty$ además para k suficientemente grande se tiene que $\frac{1}{\|x_k\|}\langle x_k, y\rangle \geq 0$. Por lo tanto, para todo $\varepsilon > 0$ tenemos que

$$\langle d, y + \varepsilon d\rangle \geq \varepsilon \|d\|^2 > 0,$$

de donde se concluye que $y + \varepsilon d \notin C_\infty^\circ$, es decir, $y \notin \operatorname{int}(C_\infty^\circ)$.

(c) Notemos que si $A \subseteq B$ entonces $B^\circ \subseteq A^\circ$. Además, es fácil verificar que $C_\infty \subseteq C_\infty^{\circ\circ}$. Luego por la parte (a) se tiene que

$$\operatorname{dom} \sigma_C \subseteq C_\infty^\circ \Longrightarrow C_\infty \subseteq C_\infty^{\circ\circ} \subseteq (\operatorname{dom} \sigma_C)^\circ.$$

Para la otra inclusón consideremos $d \in (\operatorname{dom} \sigma_C)^\circ$, $t > 0$ y $\bar{x} \in C$. Como $td \in (\operatorname{dom} \sigma_C)^\circ$, tenemos que para cualquier $y \in \operatorname{dom} \sigma_C$ se tiene

$$\langle \bar{x} + td, y\rangle \leq \langle \bar{x}, y\rangle \leq \sigma_C(y).$$

Por otra parte, si $y \notin \operatorname{dom} \sigma_C$ el lado derecho es infinito, por lo que la desigualdad sigue siendo cierta. Luego, $\bar{x} + td \in \overline{C}$ para todo $t > 0$, en efecto, como C es convexo, por el Teorema de Hanh-Banach existe $z \in \mathbb{R}^n \setminus \{0\}$ y $\alpha \in \mathbb{R}$ tales que

$$\langle \bar{x} + td, z\rangle > \alpha \geq \langle y, z\rangle \quad \forall y \in \overline{C},$$

tomando supremo sobre los $y \in \overline{C}$ obtenemos una contradicción. Finalmente, usando las propiedades elementales del cono de recesión concluimos el resultado pedido.

Problema 1.6

(a) Consideremos la función $f(x) := d(x, \phi(x))$. Es fácil ver que f es continua e inferiormente acotada por 0. Luego por el PVE con $\varepsilon \in (0, 1-k)$ y $\lambda = 1$ existe $\bar{x} \in X$ tal que

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

Supongamos que $\phi(\bar{x}) \neq \bar{x}$, luego como ϕ es una contracción direccional, existe $z \neq \bar{x}$ con $z \in [\bar{x}, \phi(\bar{x})]$, esto implica que

$$d(\bar{x}, z) + d(z, \phi(\bar{x})) = d(\bar{x}, \phi(\bar{x})) = f(\bar{x}),$$

luego,

$$\begin{aligned} d(\bar{x}, z) &= f(\bar{x}) - d(z, \phi(\bar{x})) \\ &\leq f(z) + \varepsilon d(\bar{x}, z) - d(z, \phi(\bar{x})) \\ &= d(z, \phi(z)) + \varepsilon d(\bar{x}, z) - d(z, \phi(\bar{x})) \\ &\leq d(z, \phi(\bar{x})) + d(\phi(\bar{x}), \phi(z)) + \varepsilon d(\bar{x}, z) - d(z, \phi(\bar{x})) \\ &\leq k d(\bar{x}, z) + \varepsilon d(\bar{x}, z) \\ &= (k + \varepsilon) d(\bar{x}, z), \end{aligned}$$

lo que es una contradicción, por lo tanto $\phi(\bar{x}) = \bar{x}$.

(b) Consideremos el espacio $X \times X$ dotado de la métrica $\rho((x_1, y_1)(x_2, y_2)) = d(x_1, x_2) + d(y_1, y_2)$. Es claro que $(X \times X, \rho)$ es un espacio métrico completo. Tomemos $\varepsilon \in (0, \frac{1}{2})$ y consideremos la función $g : X \times X \to \mathbb{R} \cup \{+\infty\}$ dada por

$$g(x, y) = f(x) - (1 - \varepsilon) d(x, y) + \delta_{\mathrm{Gr}(F)}(x, y) \quad \forall (x, y) \in X \times X.$$

Ahora, como f es s.c.i. y $\mathrm{Gr}(F)$ es cerrado tenemos que g es s.c.i., más aún, g es inferiormente acotada por $\inf f$, en efecto, sean $(x, y) \in \mathrm{Gr}(F)$

$$\begin{aligned} g(x, y) &= f(x) - (1 - \varepsilon) d(x, y) + \delta_{\mathrm{Gr}(F)}(x, y) \\ &= f(x) - (1 - \varepsilon) d(x, y) \\ &\geq f(y) + \varepsilon d(x, y) \\ &\geq \inf f. \end{aligned}$$

Además es claro que g es propia pues f lo es, luego por PVE existe $(\bar{x}, \bar{y}) \in \mathrm{Gr}(F)$ tal que

$$g(\bar{x}, \bar{y}) \leq g(x, y) + \varepsilon \rho((x, y), (\bar{x}, \bar{y})) \quad \forall (x, y) \in X \times X.$$

Sean $(x, y) \in X \times X$, luego re-escribiendo la desigualdad anterior tenemos

$$f(\bar{x}) - (1 - \varepsilon) d(\bar{x}, \bar{y}) \leq f(x) - (1 - \varepsilon) d(x, y) + \varepsilon [d(\bar{x}, x) + d(\bar{y}, y)].$$

Sea $z \in F(\bar{y})$ y tomemos $(x, y) = (\bar{y}, z)$, luego obtenemos

$$f(\bar{x}) - (1 - \varepsilon) d(\bar{x}, \bar{y}) \leq f(\bar{y}) - (1 - \varepsilon) d(\bar{y}, z) + \varepsilon [d(\bar{x}, \bar{y}) + d(\bar{y}, z)],$$

sigue que

$$f(\bar{x}) + d(\bar{x}, \bar{y}) \leq f(\bar{y}) - (1 - 2\varepsilon) d(\bar{y}, z),$$

usando la condición de crecimiento de f, se obtiene que $(1 - 2\varepsilon) d(\bar{y}, z) \leq 0$, por lo que necesariamente $z = \bar{y}$. Luego $\bar{y}$ es un punto fijo de F.

A.2. Fundamentos de Análisis Convexo

Problema 2.2

Probaremos: (a) ⇔ (b) y (b) ⇔ (c)

$\boxed{\text{(a)} \Rightarrow \text{(b)}}$ Supongamos que f es coerciva y que la afirmación no es cierta. Recordemos que

$$\liminf_{\|x\|\to+\infty} \frac{f(x)}{\|x\|} = \lim_{r\to+\infty} \inf_{\|x\|\geq r} \frac{f(x)}{\|x\|} = \sup_{r>0} \inf_{\|x\|\geq r} \frac{f(x)}{\|x\|}$$

entonces para cualquier $r \geq 0$ se tiene que $\inf\left\{\frac{f(x)}{\|x\|} : \ \|x\| \geq r\right\} \leq 0$. Luego podemos tomar una sucesión $\{x_k\} \subseteq X$ tal que $\frac{\|x_k\|}{k} \to +\infty$ y $\frac{f(x_k)}{\|x_k\|} \leq \frac{1}{k}$. Sea $x \in \operatorname{dom} f$ y consideremos $y_k = x + \frac{k}{\|x_k\|}(x_k - x) \in \operatorname{dom} f$ para k suficientemente grande tal que $\frac{k}{\|x_k\|} < 1$. Sigue que

$$\|y_k - x\| \geq k - \frac{k}{\|x_k\|}\|x\| \to +\infty.$$

En particular, $\{y_k\}$ no es acotada, pero por convexidad se tiene que

$$f(y_k) \leq \underbrace{(1 - \frac{k}{\|x_k\|})}_{\leq 1} f(x) + \underbrace{\frac{k}{\|x_k\|} f(x_k)}_{\leq 1} \leq f(x) + 1 =: \nu < +\infty,$$

es decir, $\{y_k\} \subseteq \Gamma_\nu(f)$, lo que es una contradicción, pues f es coerciva.

$\boxed{\text{(a)} \Leftarrow \text{(b)}}$ Supongamos que $\liminf\limits_{\|x\|\to+\infty} \frac{f(x)}{\|x\|} > 0$. Sea $r, \varepsilon > 0$ tales que si $\|x\| > r$ entonces $f(x) \geq \varepsilon\|x\|$. Tomemos $\nu \in \mathbb{R}$, luego es claro que si $x \in \Gamma_\nu(f)$ entonces $\|x\| \leq \max\{r, \frac{\nu}{\varepsilon}\}$, lo que implica que f es coerciva.

$\boxed{\text{(b)} \Leftrightarrow \text{(c)}}$ Notemos que si $L = \liminf\limits_{\|x\|\to+\infty} \frac{f(x)}{\|x\|} > 0$ entonces para cualquier $\delta \in (0, L)$ existe $r_0 > 0$ tal que $\inf\left\{\frac{f(x)}{\|x\|} : \ \|x\| \geq r_0\right\} \geq L - \delta > 0$. Además, como f es inferiormente acotada en $B(0, r_0)$ se tiene que, podemos escoger $\alpha > 0$ tal que $\inf\{f(x) : \ \|x\| \leq r_0\} \geq (L - \delta)r_0 - \alpha$. Luego (b) es equivalente a decir que existen $\alpha, \varepsilon > 0$ tales que para todo $x \in X$ se tiene $f(x) \geq \varepsilon\|x\| - \alpha$.

Por otra parte, si suponemos (b) entonces

$$f^*(x^*) = \sup_{x\in X}\{\langle x^*, x\rangle - f(x)\} \leq \sup_{x\in X}\{\langle x^*, x\rangle - \varepsilon\|x\|\} + \alpha = \delta_{B(0,\varepsilon)}(x^*) + \alpha.$$

Luego, f^* es uniformemente acotada en una vecindad del origen, entonces $0 \in \operatorname{int}(\operatorname{dom} f^*)$. Notemos que como $f^* \in \Gamma_0(X^*)$ también se tiene la continuidad en el origen.

Recíprocamente, si $0 \in \operatorname{int}(\operatorname{dom} f^*)$ entonces f^* es continua en el origen y por lo tanto uniformemente acotada en una vecindad de éste. Por lo tanto, existe $\varepsilon, \alpha > 0$ tales que $f^*(x^*) \leq \delta_{B(0,\varepsilon)}(x^*) + \alpha$. Esto implica que

$$f(x) = f^{**}(x) = \sup_{x^*\in X^*}\{\langle x^*, x\rangle - f^*(x^*)\} \leq \sup_{\|x^*\|\leq\varepsilon}\{\langle x^*, x\rangle\} - \alpha = \varepsilon\|x\| - \alpha.$$

Por lo tanto las tres afirmaciones son equivalentes.

Ahora bien, supongamos que f es coerciva, luego por el Teorema de Moreau se tiene que $0 \in \operatorname{int}(\operatorname{dom} f^*)$ y por lo tanto existe una vecindad V del origen donde f^* es continua. Sea $x_0^* \in V$ y definamos $f_0(x) = f(x) - \langle x_0^*, x\rangle$. Por las reglas de caĺculo de la conjugada de Fenchel se tiene que $f_0^*(x^*) = f^*(x^* + x_0^*)$. Además, como f^* es continua en x_0^* se tiene que f_0^* es continua en el origen y por lo tanto $0 \in \operatorname{int}(\operatorname{dom} f^*)$. Nuevamente, por el Teorema de Moreau, se tiene que f_0 coerciva y por lo tanto el problema $\min\{f_0(x) : \ x \in X\}$ tiene solución, lo que implica que existe $x_0 \in X$ tal que

$$\inf_{x\in X}\{f(x) - \langle x_0^*, x\rangle\} = f(x_0) - \langle x_0^*, x_0\rangle,$$

pero $-f^*(x_0^*) = \inf_{x\in X}\{f(x) - \langle x_0^*, x\rangle\}$. Luego

$$f^*(x_0^*) + f(x_0) = \langle x_0^*, x_0\rangle,$$

de donde se tiene que $x_0^* \in \partial f(x_0) \subseteq \bigcup_{x\in X} \partial f(x)$. Finalmente, $V \subseteq \bigcup_{x\in X} \partial f(x)$
Problema 2.4

(a) Sea $B = \frac{1}{n}\operatorname{tr}(A)I_n$, notemos que $\|B\|^2 = \frac{1}{n^2}\operatorname{tr}(A)^2\|I_n\|^2 = \frac{1}{n}\operatorname{tr}(A)^2$, pues $\|I_n\|^2 = n$. Además, notemos f se puede escribir de una manera más simple:

$$\begin{aligned} f(A) &= \frac{1}{n}\left[\|A\|^2 - \frac{1}{n}\operatorname{tr}(A)^2\right] \\ &= \frac{1}{n}\left[\|A\|^2 - \frac{2}{n}\operatorname{tr}(A)^2 + \frac{1}{n}\operatorname{tr}(A)^2\right] \\ &= \frac{1}{n}\left[\|A\|^2 - 2\left\langle A, \frac{1}{n}\operatorname{tr}(A)I_n\right\rangle + \frac{1}{n}\operatorname{tr}(A)^2\right] \\ &= \frac{1}{n}\left[\|A\|^2 - 2\langle A, B\rangle + \|B\|^2\right] \\ &= \frac{1}{n}\|A - B\|^2. \end{aligned}$$

Consideremos la aplicación $A \mapsto A - \frac{1}{n}\operatorname{tr}(A)I_n$, es sencillo ver que esta aplicación es lineal. Más aún, como la función $U \mapsto \frac{1}{n}\|U\|^2$ es convexa (y diferenciable), se tiene que f también es convexa (y diferenciable).

(b) Sea $S \in \mathcal{S}^n$ y consideremos la función $\phi_S(A) := \langle S, A\rangle - f(A)$, luego ϕ es cóncava y diferenciable, luego ϕ alcanza un máximo en $A \in \mathcal{S}^n$ si y sólo la $\nabla\phi_S(A) = 0$. Calculando obtenemos que

$$(*) \qquad \nabla\phi_S(A) = S - \frac{2}{n}(A - \frac{1}{n}\operatorname{tr}(A)I_n) = 0$$

Supongamos que ϕ_S alcanza su máximo en A, luego A satisface la ecuación $(*)$ y tomando traza en esta ecuación se obtiene que $\operatorname{tr}(S) = 0$. Supongamos ahora que $\operatorname{tr}(S) \neq 0$ y consideremos $A_k = k\operatorname{tr}(S)I_n$, entonces se tiene que $\phi_S(A_k) = k\operatorname{tr}(S)^2 \to +\infty$, si $k \to +\infty$, sigue que $f^*(S) = +\infty$. Por lo tanto, $\operatorname{dom} f^*$ es como en el enunciado.

Finalmente, es fácil ver que si $S \in \mathcal{S}^n$ y $\operatorname{tr}(S) = 0$, en particular, $A = \frac{n}{2}S$ resuelve la ecuación $(*)$ y por lo tanto

$$f^*(S) = \frac{n}{2}\operatorname{tr}(S^2) - f(\frac{n}{2}S) = \frac{n}{4}\operatorname{tr}(S^2)$$

Problema 2.6

(a) Probaremos: (1) $\Leftrightarrow$ (2) y (2) $\Leftrightarrow$ (3)

$\boxed{(1) \Rightarrow (2)}$ Supongamos que f es $\beta-$fuertemente convexa. Sea $x, y \in \mathbb{R}^n$ y $\lambda \in (0,1)$. Luego reescribiendo la desigualdad de la definición obtenemos

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

dividiendo por $(1-\lambda)$ y haciendo $\lambda \to 1$ obtenemos que

$$(*) \qquad f(x) + \frac{1}{2}\beta\|x-y\|^2 + f'(x; y-x) \leq f(y).$$

Por otra parte, como f es continua, se tiene que para todo $x, d \in \mathbb{R}^n$

$$\langle x^*, d\rangle \leq f'(x; d) \quad \forall x^* \in \partial f(x).$$

Tomando $d = y - x$ y reemplazando en $(*)$, obtenemos el resultado.

$\boxed{(1) \Leftarrow (2)}$ Supongamos que para todo $x^* \in \partial f(x)$, $y \in \mathbb{R}^n$

$$(**) \qquad f(x) + \langle x^*, y-x\rangle + \frac{\beta}{2}\|y-x\|^2 \leq f(y).$$

Sean $u, v \in \mathbb{R}^n$ y $\lambda \in (0,1)$. Tomando en $(**)$, $x = \lambda u + (1-\lambda)v$, $y = v$, obtenemos que

$$\text{(A.1)} \qquad f(\lambda u + (1-\lambda)v) + \lambda\langle x^*, v-u\rangle + \frac{\beta\lambda^2}{2}\|v-u\|^2 \leq f(v).$$

Haciendo lo mismo pero con $y = u$ obtenemos

$$\text{(A.2)} \qquad f(\lambda u + (1-\lambda)v) - (1-\lambda)\langle x^*, v-u\rangle + \frac{\beta(1-\lambda)^2}{2}\|v-u\|^2 \leq f(u).$$

Finalmente, notando que $(1-\lambda)\lambda^2 + \lambda(1-\lambda)^2 = \lambda(1-\lambda)$ y calculando $(1-\lambda)$(A.1)$+\lambda$(A.2) obtenemos que f es $\beta-$fuertemente convexa.

$\boxed{(2) \Rightarrow (3)}$ Sean $x, y \in \mathbb{R}^n$, luego sabemos que

$$f(x) + \langle x^*, y-x\rangle + \frac{\beta}{2}\|y-x\|^2 \leq f(y), \quad x^* \in \partial f(x).$$

Intercambiando x por y obtenemos

$$f(y) + \langle y^*, x-y\rangle + \frac{\beta}{2}\|y-x\|^2 \leq f(x), \quad y^* \in \partial f(y).$$

El resultado viene de sumar estas dos expresiones.

$\boxed{(2) \Leftarrow (3)}$ Sean $x, y \in \mathbb{R}^n$ y $\lambda \in (0,1)$. Consideremos la función $\phi(\lambda) = f(x + \lambda(y-x))$, es claro que ϕ es convexa y finita, por lo tanto es subdiferenciable y su derivada direccional está bien definida, más aún,

$$\begin{aligned}
\phi'(\lambda;1) &= \inf_{t>0} \left\{ \frac{f(x+(\lambda+t)(y-x)) - f(x+\lambda(y-x))}{t} \right\} \\
&= \inf_{t>0} \left\{ \frac{f(x+\lambda(y-x)+t(y-x)) - f(x+\lambda(y-x))}{t} \right\} \\
&= f'(x+\lambda(y-x); y-x) \\
&\geq \langle z^*, y-x \rangle, \quad \forall z^* \in \partial f(x+\lambda(y-x)) = (1-\lambda)\partial f(x) + \lambda \partial f(y) \\
&\geq \langle x^*, y-x \rangle + \lambda \langle y^* - x^*, y-x \rangle, \quad \forall x^* \in \partial f(x), y^* \in f(y) \\
&\geq \langle x^*, y-x \rangle + \lambda\beta \|x-y\|^2, \quad \forall x^* \in \partial f(x).
\end{aligned}$$

Finalmente, tenemos que

$$\begin{aligned}
f(x) - f(y) &= \phi(1) - \phi(0) \\
&\geq \int_0^1 \phi'(\lambda;1) d\lambda \\
&\geq \int_0^1 [\langle x^*, y-x \rangle + \lambda\beta\|x-y\|^2] d\lambda, \quad \forall x^* \in \partial f(x) \\
&= \langle x^*, y-x \rangle + \frac{\beta}{2}\|x-y\|^2, \quad \forall x^* \in \partial f(x).
\end{aligned}$$

(b) Supongamos que f es $\beta-$fuertemente convexa, luego como f es finita se tiene que f es continua y por lo tanto $\operatorname{dom} \partial f = \mathbb{R}^n$. Sea $x \in \mathbb{R}^n$ y tomemos $x^* \in \partial f(x)$. Reemplacemos en (a.2) con $y = x + td$, $t > 0$ y $d \in \mathbb{R}^n \setminus \{0\}$, luego

$$f(x+td) \geq f(x) + t\langle x^*, d\} + \frac{\beta}{2} t^2 \|d\|.$$

Esto implica que $f_\infty(d) = +\infty$, para todo $d \in \mathbb{R}^n \setminus \{0\}$. En particular, f es coercitiva y por el Teorema de Moreau (ejercicio 2.2) se tiene que $\operatorname{int}(\operatorname{dom} f^*) \neq \emptyset$. Ahora bien, notemos que

$f_\infty = \sigma_{\mathrm{dom}\, f^*}$, en efecto, como $f = f^{**}$ se tiene que

$$\begin{aligned}
f'_\infty(d) &= \sup_{t>0} \left\{ \frac{f(x+td) - f(x)}{t} \right\} \quad \forall x \in \mathbb{R}^n \\
&= \sup_{x\in\mathbb{R}^n} \sup_{t>0} \left\{ \frac{f(x+td) - f(x)}{t} \right\} \\
&= \sup_{x\in\mathbb{R}^n} \sup_{t>0} \left\{ \frac{f^{**}(x+td) - f(x)}{t} \right\} \\
&= \sup_{x\in\mathbb{R}^n} \sup_{t>0} \sup_{x^*\in\mathrm{dom}\, f^*} \left\{ \frac{\langle x+td, x^*\rangle - f^*(x^*) - f(x)}{t} \right\} \\
&= \sup_{x^*\in\mathrm{dom}\, f^*} \sup_{t>0} \sup_{x\in\mathbb{R}^n} \left\{ \langle d, x^*\rangle + \frac{\langle x, x^*\rangle - f^*(x^*) - f(x)}{t} \right\} \\
&= \sup_{x^*\in\mathrm{dom}\, f^*} \sup_{t>0} \left\{ \langle d, x^*\rangle + \frac{\sup_{x\in\mathbb{R}^n}\{\langle x, x^*\rangle - f(x)\} - f^*(x^*)}{t} \right\} \\
&= \sup_{x^*\in\mathrm{dom}\, f^*} \{\langle d, x^*\rangle\}
\end{aligned}$$

Luego $\sigma_{\mathrm{dom}\, f^*}(d) = +\infty$, para todo $d \in \mathbb{R}^n \setminus \{0\}$. Sigue que $\mathrm{dom}\, f^*$ debe ser denso en $\mathbb{R}^n$. Notemos que en un e.v.n. un convexo cualquiera K de interior no vacío satisface que $\mathrm{int}(\overline{K}) = \mathrm{int}(K)$, por lo tanto $\mathrm{dom}\, f^* = \mathbb{R}^n$.

Sea $x \in \mathbb{R}^n$ luego existe $x^* \in \partial f(x)$, y por la fórmula de reciprocidad se tiene que $x \in \partial f^*(x^*)$. Luego para ver la diferenciabilidad de f^* basta ver que $\partial f^*(x^*)$ es un singletón, lo que es directo de (a.3). Finalmente, usando (a.3) una vez más se obtiene que ∇f^* es $\frac{1}{\beta}$−Lipschitz.

Por otra parte, asumamos que $\mathrm{dom}\, f^* = \mathbb{R}^n$, que f^* es diferenciable y que ∇f^* es Lipschitz de constante $\frac{1}{\beta}$ y consideremos la función diferenciable $\varphi(t) := f^*(u + t(v-u))$, luego

$$\varphi'(t) = \langle \nabla f^*(u + t(v-u)), v - u\rangle,$$

pero como ∇f^* es Lipschitz de constante $\frac{1}{\beta}$ se tiene que

$$\begin{aligned}
\varphi'(t) - \varphi'(0) &= \langle \nabla f^*(u + t(v-u)) - \nabla f^*(u), v - u\rangle \\
&\leq \frac{|t|}{\beta}\|v-u\|^2,
\end{aligned}$$

por lo tanto

$$\begin{aligned}
f^*(v) - f^*(u) &= \varphi(1) - \varphi(0) \\
&= \int_0^1 \varphi'(t)dt \\
&\leq \int_0^1 [\varphi'(0) + \frac{t}{\beta}\|v-u\|^2]dt \\
&= \langle \nabla f^*(u), v - u\rangle + \frac{1}{2\beta}\|v-u\|^2.
\end{aligned}$$

Notemos que la desigualdad anterior es válida para cualquier $u, v \in \operatorname{dom} f^* = \mathbb{R}^n$, y por lo tanto, fijando $y \in \operatorname{int}(\operatorname{dom} f)$ podemos tomar $u \in \partial f(y)$ y se tiene que $f^*(u) = \langle y, u\rangle - f(y)$, sigue que

$$\begin{aligned}
f(x) &= \sup_{v\in\mathbb{R}^n} \{\langle x, v\rangle - f^*(v)\} \\
&\geq \sup_{v\in\mathbb{R}^n} \{\langle x, v\rangle - f^*(u) - \langle \nabla f^*(u), v-u\rangle - \frac{1}{2\beta}\|v-u\|^2\} \\
&= \sup_{v\in\mathbb{R}^n} \{\langle x, v\rangle - \langle y, u\rangle + f(y) - \langle \nabla f^*(u), v-u\rangle - \frac{1}{2\beta}\|v-u\|^2\} \\
&= f(y) + \sup_{v\in\mathbb{R}^n} \{\langle x - \nabla f^*(u), v\rangle + \langle \nabla f^*(u) - y, u\rangle - \frac{1}{2\beta}\|v-u\|^2\},
\end{aligned}$$

pero como $u \in \partial f(y)$ se tiene que $y \in \partial f^*(u)$ y dado que f^* es diferenciable se tiene que $\partial f^*(u) = \{\nabla f^*(u)\}$ y por lo tanto $\langle \nabla f^*(u) - y, u\rangle = 0$. Esto último implica que

$$f(x) = f(y) + \sup_{v\in\mathbb{R}^n} \{\langle x - y, v\rangle - \frac{1}{2\beta}\|v-u\|^2\}.$$

Es fácil ver que en $v = u + \beta(x-y)$ se alcanza el supremo de la última desigualdad, luego se tiene que

$$f(x) \geq f(y) + \langle u, x-y\rangle + \frac{\beta}{2}\|x-y\|^2, \quad \forall u \in \partial f(y).$$

Problema 2.8

(a) Asumamos que ∂f es un síngleton en su dominio y que existe un convexo de $\operatorname{ran}\partial f$ donde f^* no es estrictamente convexa. Luego, existen $x^* \neq y^*$, $[x^*, y^*] \subseteq \operatorname{ran}\partial f$ y $\lambda \in (0,1)$ tales que

$$f(z^*) = \lambda f^*(x^*) + (1-\lambda) f^*(y^*),$$

con $z^* = \lambda x^* + (1-\lambda)y^*$. Como $z^* \in [x^*, y^*] \subseteq \operatorname{ran}\partial f$ se tiene que existe $z \in X$ tal que $z^* \in \partial f(z)$, es decir, $f^*(z^*) + f(z) = \langle z^*, z\rangle$, pero

$$0 = f^*(z^*) + f(z) - \langle z^*, z\rangle = \lambda[\underbrace{f^*(x^*) + f(z) - \langle x^*, z\rangle}_{\geq 0}] + (1-\lambda)[\underbrace{f^*(y^*) + f(z) - \langle y^*, z\rangle}_{\geq 0}],$$

ésto permite deducir, dado que $\lambda \in (0,1)$, que $f^*(x^*) + f(z) = \langle x^*, z\rangle$ y $f^*(y^*) + f(z) = \langle y^*, z\rangle$, es decir $x^*, y^* \in \partial f(z)$, lo que es imposible. Por lo tanto, f^* es estrictamente convexa en subconjuntos convexo de $\operatorname{ran}\partial f$.

Por otra parte, asumamos que f^* es estrictamente convexa en subconjuntos convexo de $\operatorname{ran}\partial f$. Sean $x_1^*, x_2^* \in \partial f(x)$ para algún $x \in \operatorname{dom}\partial f$. Luego, por la igualdad de Fenchel sabemos que $f^*(x_i^*) + f(x) = \langle x_i^*, x\rangle$, $i = 1, 2$. Sea $\lambda \in [0,1]$, luego

$$\begin{aligned}
f(x) + \lambda f^*(x_1^*) + (1-\lambda) f^*(x_2^*) &= \langle \lambda x_1^* + (1-\lambda)x_2^*, x\rangle \\
&\leq f(x) + f^*(\lambda x_1^* + (1-\lambda)x_2^*) \\
&\leq f(x) + \lambda f^*(x_1^*) + (1-\lambda) f^*(x_2^*).
\end{aligned}$$

Por lo tanto, todas las desigualdades anteriores son en realidad igualdades, sigue que

$$f^*(\lambda x_1^* + (1-\lambda)x_2^*) = \lambda f^*(x_1^*) + (1-\lambda)f^*(x_2^*), \quad \forall \lambda \in [0,1]$$

y dado que f^* es estrictamente convexa, necesariamente $x_1^* = x_2^*$.

La conclusión es directa, basta notar que f no es esencialmente por vacuidad ya que $\operatorname{int}(\operatorname{dom} f) \subseteq \operatorname{dom} \partial f$ y que $\operatorname{ran} \partial f = \operatorname{dom} \partial f^*$.

(b) Dado que f es esencialmente suave, por la parte anterior f^* es esencialmente convexa estricta. Además, dado que $\operatorname{int}(\operatorname{dom} f^*) \cap \operatorname{int}(\operatorname{dom} g^*) \neq \emptyset$ es fácil ver que $f \square g \in \Gamma_0(X)$, luego por ejercicio 2.5 se tiene que

$$f \square g = (f^* + g^*)^* \quad ,$$

Además, por el Teorema de Moreau-Rockafellar tenemos que

$$\partial(f^* + g^*)(x^*) = \partial f^*(x^*) + \partial g^*(x^*) \quad \forall x^*.$$

Esto último implica que

$$\operatorname{dom} \partial(f^* + g^*) \subseteq \operatorname{dom} \partial f^*.$$

luego como f^* es esencialmente convexa estricta, $f^* + g^*$ también lo es. Finalmente, por la parte (a) se tiene que $(f^* + g^*)^*$ es esencialmente suave, de donde se concluye.

(c) Notemos que dado que f es esencialmente suave, $\partial f(x)$ es un síngleton para todo $x \in \operatorname{dom} \partial f$, y dado que $\operatorname{int}(\operatorname{dom} f) \subseteq \operatorname{dom} \partial f$ se tiene que f es diferenciable en $\operatorname{int}(\operatorname{dom} f)$. Sea $\{x_n\} \subseteq \operatorname{int}(\operatorname{dom} f)$ tal que $x_n \to x \in \partial \operatorname{dom} f$ y supongamos que $\{\nabla f(x_n)\}$ es acotada. Luego, por el Teorema de Banach-Alaoglu, podemos suponer (pasando a una subsucesión) que existe $x^* \in X^*$ tal que $\nabla f(x_n) \stackrel{*}{\rightharpoonup} x^*$. Por otra parte, por definición tenemos que

$$f(x_n) + \langle \nabla f(x_n), y - x_n \rangle \leq f(y) \quad \forall y \in X.$$

Finalmente, haciendo $n \to \infty$ en esta última desigualdad y, usando la s.c.i de f y un argumento de convergencia débil-fuerte, tenemos que

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

es decir, $x^* \in \partial f(x)$, pero dado que $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$ y $x \in \operatorname{dom} \partial f \setminus \operatorname{int}(\operatorname{dom} f)$ se tiene que $\partial f(x)$ es no acotado (ver problema 2.12), lo que contradice el hecho que $\partial f(x)$ sea un síngleton.

Problema 2.9

(a) Notemos que $\operatorname{dist}_C(x) = (\|\cdot\| \square \delta_C)(x)$ y como la función distancia es continua (de hecho Lipschitz), tenemos que

$$\operatorname{dist}_C(x) = \operatorname{dist}_C^{**}(x) = (\|\cdot\| \square \delta_C)^{**}(x) = (\|\cdot\|^* + \delta_C^*)^*(x).$$

No es difícil ver que $\|\cdot\|^*(x^*) = \delta_{B^*}(x^*)$ con B^* la bola unitaria en X^* y que $\delta_C^*(x^*) = \sigma_C(x^*)$. Luego,

$$\operatorname{dist}_C(x) = \sup_{x^* \in X^*} \{\langle x, x^* \rangle - \delta_{B^*}(x^*) - \sigma_C(x^*)\} = \sup_{x^* \in B^*} \{\langle x, x^* \rangle - \sigma_C(x^*)\} = \sup_{x^* \in B^*} \inf_{y \in C} \{\langle x - y, x^* \rangle\}.$$

Notemos que el supremo es alcanzado, pues si consideramos $\varphi(x^*) = \text{ínf}_{y\in C}\{\langle x-y, x^*\rangle\}$, es fácil ver que $-\varphi$ es s.c.i. para la topología fuerte, luego como $-\varphi$ es también convexa se tiene que $-\varphi$ es s.c.i. para la topología débil y como B^* es compacto débil (pues estamos en un Hilbert), se concluye que el supremo es en realidad un máximo.

(b) Sea $x \notin C$ luego sabemos que existe $\bar{x} \in C$ tal que $\text{dist}_C(x) = \|x - \bar{x}\|$, tal elemento es la proyección de x sobre X y usualmente es denotado por $\text{Proy}_C(x)$. En este caso, como C es cerrado se tiene que $x \neq \text{Proy}_C(x)$. Ahora, como la inf-convolución que define la función distancia es exacta se tiene que (ver problema 2.5),

$$\partial\text{dist}_C(x) = \partial\|\cdot\|(x - \text{Proy}_C(x)) \cap \partial\delta_C(\text{Proy}_C(x)).$$

Como $x \neq \text{Proy}_C(x)$ se tiene que $\|\cdot\|$ es diferenciable en $x - \text{Proy}_C(x)$ pues X es un espacio de Hilbert, y además, su gradiente viene dado por

$$\nabla\|\cdot\|(x - \text{Proy}_C(x)) = \frac{x - \text{Proy}_C(x)}{\|x - \text{Proy}_C(x)\|}.$$

Calculemos ahora $\partial\delta_C(\text{Proy}_C(x)$, por definición se tiene que

$$\begin{aligned}
x^* \in \partial\delta_C(\text{Proy}_C(x)) &\Leftrightarrow \langle y - \text{Proy}_C(x), x^*\rangle + \delta_C(\text{Proy}_C(x)) \leq \delta_C(y), \quad \forall y \in X \\
&\Leftrightarrow \langle y - \text{Proy}_C(x), x^*\rangle \leq 0, \quad \forall y \in C. \\
&\Leftrightarrow x^* \in N_C(\text{Proy}_C(x)).
\end{aligned}$$

Luego, si $x \notin C$ entonces

$$\partial\text{dist}_C(x) = \left\{\frac{x - \bar{x}}{\|x - \bar{x}\|}\right\},$$

donde $\bar{x}$ es la unica solución de la desigualdad variacional

$$\langle y - \bar{x}, x - \bar{x}\rangle \leq 0, \quad \forall y \in C.$$

Por otro lado, si $x \in C$, de todas formas tenemos que la inf-convolución que define la función distancia es exacta, pero ahora $\text{dist}_C(x) = \|0\| + \delta_C(x) = 0$ y por lo tanto

$$\partial\text{dist}_C(x) = \partial\|\cdot\|(0) \cap \partial\delta_C(x).$$

Calculemos $\partial\|\cdot\|(0)$. Notemos que como $\|\cdot\|$ es finita y continua en 0 se tiene que este subdiferencial en no vacío. Sigue que

$$\begin{aligned}
x^* \in \partial\|\cdot\|(0) &\Leftrightarrow \langle y - 0, x^*\rangle + \|0\| \leq \|y\|, \quad \forall y \in X \\
&\Leftrightarrow \langle y, x^*\rangle + \leq \|y\|, \quad \forall y \in X \\
&\Leftrightarrow \|x^*\| \leq 1 \\
&\Leftrightarrow x^* \in B^*.
\end{aligned}$$

Por lo tanto, en este caso se tiene que

$$\partial\text{dist}_C(x) = N_C(x) \cap B^*.$$

Problema 2.12

(a) Consideremos f como en la indicación. Primero notemos que f es supremo de funciones lineales afines, luego f es convexa y s.c.i., y además $f > -\infty$.

Veamos que $\operatorname{dom} T \subseteq \operatorname{dom} f$. Sea $(\bar{x}, \bar{x}^*) \in \operatorname{Gr}(T)$, tomemos $k \in \mathbf{N}$, $\{(x_i, x_i^*)\}_{i=1}^k \subseteq \operatorname{Gr}(T)$, fijemos $n = k+1$ y $(x_n, x_n^*) = (\bar{x}, \bar{x}^*)$. Dado que T es ciclícamente monótono, sigue que

$$\sum_{i=0}^{k-1} \langle x_{i+1} - x_i, x_i^* \rangle + \langle \bar{x} - x_k, x_k^* \rangle \leq \langle \bar{x} - x_0, \bar{x}^* \rangle,$$

tomando supremo se obtiene que $f(x) \leq \langle \bar{x} - x_0, \bar{x}^* \rangle$ por lo tanto $\bar{x} \in \operatorname{dom} f$. En particular, $f(x_0) \leq 0$. Además, tomando $n = 1$ y $(x_1, x_1^*) = (x_0, x_0^*)$ se tiene que $f(x_0) \geq 0$.

Mostremos ahora que si $(\bar{x}, \bar{x}^*) \in \operatorname{Gr}(T)$ entonces $\bar{x}^* \in \partial f(\bar{x})$. Sea $x \in X$ y tomemos $\mu < f(\bar{x}) + \langle x - \bar{x}, \bar{x}^* \rangle$. Luego, existe $k \in \mathbf{N}$ y $\{(x_i, x_i^*)\}_{i=0}^k \subseteq \operatorname{Gr}(T)$ tales que

$$\mu - \langle x - \bar{x}, \bar{x}^* \rangle < \langle \bar{x} - x_k, x_k^* \rangle + \sum_{i=0}^{k-1} \langle x_{i+1} - x_i, x_i^* \rangle.$$

Tomemos $n = k+1$ y $(x_n, x_n^*) = (\bar{x}, \bar{x}^*)$, sigue que

$$\mu < \langle x - x_n, x_n^* \rangle + \sum_{i=0}^{n-1} \langle x_{i+1} - x_i, x_i^* \rangle \leq f(x).$$

Haciendo $\mu \to f(\bar{x}) + \langle x - \bar{x}, \bar{x}^* \rangle$ se concluye $\operatorname{Gr}(T) \subseteq \operatorname{Gr}(\partial f)$. Finalmente, dado que $f \in \Gamma_0(X)$ se tiene que ∂f es un operador maximal monótono, luego como T también lo es se concluye que $\partial f = T$.

(b) Sea $x \in \operatorname{dom} T$, y sea $z^* \in N_{\operatorname{dom} T}(x)$, luego $\langle z^*, y - x \rangle \leq 0$, para todo $y \in \operatorname{dom} T$. Tomemos $x^* \in T(x)$ y $t \geq 0$, sigue que $\forall y \in \operatorname{dom} T$, $\forall y^* \in T(y)$

$$(*) \qquad \langle (x^* + tz^*) - y^*, x - y \rangle = \langle x^* - y^*, x - y \rangle + t \langle z^*, x - y \rangle \geq 0.$$

Como T es maximal monótono, necesariamente $x^* + tz^* \in T(x)$, y como esto es cierto para todo $t \geq 0$, concluimos que $z^* \in T(x)_\infty$.

Recíprocamente, sea $z^* \in (T(x))_\infty$, $y \in \operatorname{dom} T$, $y^* \in T(y)$ y $t \geq 0$. Podemos reescribir nuevamente $(*)$ y haciendo $t \to \infty$ se obtiene que $\langle z^*, x - y \rangle \geq 0$ y por lo tanto $z^* \in N_{\operatorname{dom} T}(x)$.

(a) Basta notar que por la observación 2.4.2 tenemos que ∂f es un operador maximal monótono. Además por el Teorema de Brondsted-Rockafellar se tiene que $\overline{\operatorname{dom} \partial f} = \operatorname{dom} f$, luego $N_{\operatorname{dom} \partial f}(x) = N_{\operatorname{dom} f}(x)$ para todo $x \in \operatorname{dom} \partial f$. El resultado se obtiene al aplicar lo anteriormente demostrado.

(b) Sea $x \in \operatorname{dom} f \setminus \operatorname{int}(\operatorname{dom} f)$, luego por el Teorema de Hanh-Banach, existe $z^* \in X \setminus \{0\}$ tal que

$$\langle z^*, x \rangle \geq \langle z^*, y \rangle, \quad \forall y \in \operatorname{dom} f.$$

Por lo tanto, se tiene que $z^* \in N_{\operatorname{dom} f}(x)$ y por (a) se tiene que $(\partial f(x))_\infty \neq \emptyset$ y por lo tanto $\partial f(x))$ es no acotado.

Problema 2.14

(a) Es claro que la función $(t,x,y)\mapsto f(x-y)$ es convexa, pues f lo es. Luego, para ver la convexidad de φ basta ver la convexidad de la función $h(t,x,y)=t\theta\left(\frac{y}{t}\right)$. Sean $(t_1,x_1,y_1),(t_2,x_2,y_2)\in(0,\infty)\times\mathbb{R}^n\times\mathbb{R}^n$ y $\lambda\in[0,1]$, sigue que

$$\begin{aligned}
h(\lambda(t_1,x_1,y_1)+(1-\lambda)(t_2,x_2,y_2)) &= (\lambda t_1+(1-\lambda)t_2)\theta\left(\frac{\lambda y_1+(1-\lambda)y_2}{\lambda t_1+(1-\lambda)t_2}\right)\\
&= (\lambda t_1+(1-\lambda)t_2)\theta\left(\frac{\lambda t_1}{\lambda t_1+(1-\lambda)t_2}\frac{1}{t_1}y_1+\frac{(1-\lambda)t_2}{\lambda t_1+(1-\lambda)t_2}\frac{1}{t_2}y_2\right)\\
&\leq \lambda t_1\theta\left(\frac{1}{t_1}y_1\right)+(1-\lambda)t_2\theta\left(\frac{1}{t_2}y_2\right)\\
&= \lambda h(t_1,x_1,y_1)+(1-\lambda)h(t_2,x_2,y_2).
\end{aligned}$$

Luego φ es convexa y como u es una función marginal de φ se concluye que u es convexa en si $t>0$. Notemos que por la condición de creciemiento de θ, para $x\in\mathbb{R}^n$ fijo se tiene que $u(\cdot,x)$ es continua en $[0,\infty]$, por lo que se puede concluir la convexidad de u para todo $t\in\mathbb{R}$. Por otra parte, como θ es finita y f es propia se tiene que u también en finita, lo que tiene como consecuencia la continuidad de u en todo $[0,\infty]\times\mathbb{R}^n$.

(b) Es claro que para $(t,x)\in[0,\infty]\times\mathbb{R}^n$ dado, la función $\varphi(t,x,\cdot)$ es s.c.i., veamos que esta función también es coerciva. Como $f\in\Gamma_0(\mathbb{R}^n)$, existen $x^*\in\mathbb{R}^n$ y $\alpha\in\mathbb{R}$ tal que $f(\cdot)\geq\langle x^*,\cdot\rangle+\alpha$, entonces

$$\varphi(t,x,y)\geq\langle x^*,x-y\rangle+\alpha+t\theta\left(\frac{y}{t}\right)\geq-\|x^*\|(\|x\|+\|y\|)+\alpha+t\theta\left(\frac{y}{t}\right).$$

Ahora por la condición de crecimiento de θ tenemos que para todo $M>0$ existe $R>0$ tal que $\theta(u)\geq M\|y\|$ para todo $\|y\|>R$. Tomando $M=2\|x^*\|$ se tiene que

$$\varphi(t,x,y)\geq\|x^*\|(\|y\|-\|x\|)+\alpha,\quad\forall\|y\|>R.$$

Luego, haciendo $\|y\|\to\infty$ se tiene que φ es coerciva y por lo tanto en ínfimo en la definición es un mínimo y se alcanza en algún punto $\bar{y}=\bar{y}(t,x)$.

Para probar la equivalencia entre los subdiferenciales basta escribir la definción en cada caso y usar el hecho que $u(t,x)=\varphi(t,x,\bar{y})\leq\varphi(t,x,y)$, para todo $y\in\mathbb{R}^n$.

(c) Sean $(t^*,x^*,y^*)\in(0,\infty)\times\mathbb{R}^n\times\mathbb{R}^n$, por definición tenemos que

$$\begin{aligned}
\varphi^*(t^*,x^*,y^*) &= \sup\{t^*t+\langle x^*,x\rangle+\langle y^*,y\rangle-\varphi(t,x,y):\ t\in\mathbb{R},\ x,y\in\mathbb{R}^n\}\\
&= \sup\left\{t^*t+\langle x^*,x\rangle+\langle y^*,y\rangle-f(x-y)-t\theta\left(\frac{y}{t}\right):\ t>0,\ x,y\in\mathbb{R}^n\right\}\\
&= \sup\left\{t^*t+\langle x^*+y^*,y\rangle-t\theta\left(\frac{y}{t}\right)+\sup_{x\in\mathbb{R}^n}\{\langle x^*,x-y\rangle-f(x-y)\}:\ t>0,\ y\in\mathbb{R}^n\right\}\\
&= \sup_{t>0}\left\{t^*t+t\sup_{y\in\mathbb{R}^n}\left\{\left\langle x^*+y^*,\frac{y}{t}\right\rangle-\theta\left(\frac{y}{t}\right)\right\}\right\}+f^*(x^*)\\
&= \sup_{t>0}\{t(t^*+\theta^*(x^*+y^*))\}+f^*(x^*).
\end{aligned}$$

Por lo tanto

$$\varphi^*(t^*,x^*,y^*) = \begin{cases} f^*(x^*) & \text{si } t^*+\theta^*(x^*+y^*) \leq 0 \\ +\infty & \text{si } t^*+\theta^*(x^*+y^*) > 0. \end{cases}$$

(d) Para $(t,x) \in (0,\infty)\times\mathbb{R}^n$ notaremos por $\bar{y} = \bar{y}(t,x)$ al punto obtenido en (b). Fijemos $t>0$, luego $u(t,x) = f\Box t\theta(\frac{\cdot}{t})(x)$ es una inf-convolución exacta y se tiene que

$$\partial u(t,\cdot)(x) = \partial f(x-\bar{x}) \cap \partial\left(t\theta\left(\frac{\cdot}{t}\right)\right)(\bar{y}).$$

Más todavía, como θ es diferenciable se tiene que $\partial\left(t\theta\left(\frac{\cdot}{t}\right)\right)(\bar{y}) = \{\nabla\theta(\frac{\bar{y}}{t})\}$. Luego para $(t^*,x^*) \in \partial u(t,x)$ es directo que $x^* = \nabla\theta(\frac{\bar{y}}{t})$ y $x^* \in \partial f(x-\bar{y})$.

Además, se tiene que

$$\begin{aligned}(t^*,x^*) \in \partial u(t,x) &\Leftrightarrow (t^*,x^*,0) \in \partial\varphi(t,x,\bar{y}) \\ &\Leftrightarrow t^*t + \langle x^*,x\rangle = \varphi(t,x,\bar{y}) + \varphi^*(t^*,x^*,0) \\ &\Leftrightarrow t\left[t^* - \theta\left(\frac{\bar{y}}{t}\right)\right] = f(x-\bar{y}) + f^*(x^*) - \langle x^*,x\rangle \\ &\quad \text{con } t^*+\theta^*(x^*) \leq 0.\end{aligned}$$

Pero como $x^* \in \partial f(x-\bar{y})$, se tiene $f(x-\bar{y}) + f^*(x^*) - \langle x^*,x\rangle = -\langle x^*,\bar{y}\rangle$, por lo que

$$(t^*,x^*) \in \partial u(t,x) \Leftrightarrow t\left[t^* - \theta\left(\frac{\bar{y}}{t}\right) + \langle x^*,\frac{\bar{y}}{t}\rangle\right] = 0 \quad \text{con } t^*+\theta^*(x^*) \leq 0.$$

Finalmente, como $x^* \in \partial\theta(\frac{\bar{y}}{t})$ se tiene que $\theta^*(x^*) = \langle x^*,\frac{\bar{y}}{t}\rangle - \theta\left(\frac{\bar{y}}{t}\right)$. De donde obtenemos que $t^*+\theta^*(x^*) = 0$.

(e) Sea $(t^*,x^*) \in \partial u(t,x)$, por lo anterior sabemos que x^* está únicamente determinado y como $t^*+\theta^*(x^*) = 0$, t^* también es único. Luego, para cada $(t,x) \in (0,\infty)\times\mathbb{R}^n$, $\partial u(t,x)$ es un síngleton. Por lo tanto, u debido a la continuidad de u se tiene que u es diferenciable y $t^* = \frac{\partial u}{\partial t}(t,x)$ y $x^* = \nabla_x u(t,x)$, de donde se obtiene la ecuación.

(f) Es fácil ver que $\theta(z) = \frac{1}{2}\|z\|^2$, luego la soluciíon entregada por la fórmula de Lax-Oleinik es

$$u(t,x) = \inf_{y\in\mathbb{R}^n}\left\{f(x-y) + \frac{1}{2t}\|y\|^2\right\} = \inf_{y\in\mathbb{R}^n}\left\{f(y) + \frac{1}{2t}\|x-y\|^2\right\} = f_t(x),$$

donde $f_t(x)$ es la Regularizada de Moreau-Yosida (ver problema 2.7).

A.3. Dualidad en Optimización Convexa

Problema 3.1

(a) Consideremos, para $v = (v_1,...,v_m) \in \mathbb{R}^m$, una función del estilo $f(v) = \frac{1}{m}\sum_{i=1}^m v_i + \delta_C(v)$, con C algún convexo cerrado no vacío. Luego, es claro que $f \in \Gamma_0(\mathbb{R}^m)$, y esto sigue siendo cierto si en particular tomamos $C = \{v \in \mathbb{R}^m : \; v_1 = v_2 = ... = v_m\}$.

Consideremos ahora $g(v) = \sum_{i=1}^{m} \theta(v_i)$, con θ como en el enunciado. No es difícil verificar que $g \in \Gamma_0(\mathbb{R}^m)$ y que además g es estricamente convexa y diferenciable.

Calculemos $f \square g$

$$\begin{aligned}
(f \square g)(y_1, ..., y_m) &= \operatorname{ínf} \left\{ \frac{1}{m} \sum_{i=1}^{m} v_i + \delta_C(v) + \sum_{i=1}^{m} \theta(y_i - v_i) : \ (v_1, ..., v_m) \in \mathbb{R}^m \right\} \\
&= \operatorname{ínf} \left\{ \frac{1}{m} \sum_{i=1}^{m} v_i + \sum_{i=1}^{m} \theta(y_i - v_i) : \ (v_1, ..., v_m) \in \mathbb{R}^m \cap C \right\} \\
&= \operatorname{ínf} \left\{ v + \sum_{i=1}^{m} \theta(y_i - v) : \ v \in \mathbb{R} \right\} \\
&= M(y_1, ..., y_m)
\end{aligned}$$

Luego, usando las propiedades de la inf-convolución (problema 2.5) se tiene que M es convexa.

Por otro lado, para $y_1, ..., y_m \in \mathbb{R}$ consideremos la función $\phi(v) = v + \sum_{i=1}^{m} \theta(y_i - v)$, luego como θ es positiva se tiene que $\phi(v) \geq v$. Sigue que $\phi(v) \to +\infty$ si $v \to \infty$. Notemos también que, por la misma razón anterior se tiene que $\phi(v) \geq v + \theta(y_1 - v)$. Consideremos $v < 0$, luego

$$\frac{\phi(v)}{-v} \geq -1 + \frac{\theta(y_1 - v) - \theta(y_1)}{-v},$$

y haciendo $v \to -\infty$ se tiene que

$$\lim_{v \to -\infty} \frac{\phi(v)}{-v} \geq -1 + \theta_\infty(1) = \infty,$$

en particular, $\phi(v) > 0$ para $v < 0$ suficientemente grande. Ahora bien, como $\frac{\phi(v)}{-v} < \phi(v)$ para $v < 0$ suficientemente grande se tiene que ϕ es coerciva y por lo tanto tiene mínimo. La unicidad viene de la estricta convexidad de ϕ. Finalmente, la diferenciabilidad de M viene de las propiedades de la inf-convolución (problema 2.5).

(b) Notemos que

$$M_r(y) \leq v + r \sum_{i=1}^{m} \theta\left(\frac{1}{r}(y_i - v)\right) \quad \forall v \in \mathbb{R}.$$

Sea $y_{i_0} = \operatorname{máx} y_i$, luego como θ es creciente se tiene que

$$\theta\left(\frac{1}{r}(y_i - v)\right) \leq \theta\left(\frac{1}{r}(y_{i_0} - v)\right), \quad \forall i = 1, ..., m.$$

Sigue que $\forall v \in \mathbb{R}$

$$M_r(y) \leq v + rm\theta\left(\frac{1}{r}(y_{i_0} - v)\right) \leq r\left[m\theta\left(\frac{1}{r}(y_{i_0} - v)\right) - \frac{1}{r}(y_{i_0} - v)\right] + y_{i_0}.$$

Luego, tomando ínfimo sobre $v \in \mathbb{R}$ se tiene que

$$M_r(y) \leq -rm\theta^*\left(\frac{1}{m}\right) + y_{i_0}.$$

Por otra parte, se tiene que

$$\begin{aligned}
M_r(y) &= \min_{v\in\mathbb{R}} \left\{ v + r\theta\left(\frac{1}{r}(y_{i_0} - v)\right) + \underbrace{r\sum_{i=1,\ i\neq i_0}^{m} \theta\left(\frac{1}{r}(y_{i_0} - v)\right)}_{\geq 0} \right\} \\
&\geq \min_{v\in\mathbb{R}} \left\{ v + r\theta\left(\frac{1}{r}(y_{i_0} - v)\right) \right\} \\
&= r\min_{v\in\mathbb{R}} \left\{ \theta\left(\frac{1}{r}(y_{i_0} - v)\right) - \frac{1}{r}(y_{i_0} - v) \right\} + y_{i_0} \\
&= -r\theta^*(1) + y_{i_0}.
\end{aligned}$$

(c) Sea $L = \max\{|\theta^*(1)|, m|\theta(1/m)|\} > 0$. Notemos por L bien podría no ser un real, (si $1 \notin \operatorname{dom}\theta^*$ o $1/m \notin \operatorname{dom}\theta^*$). Sin embargo, debido a la características de θ, esta constante L es finita. Específicamente, esto viene del hecho que $[0,\infty) = \operatorname{dom}\theta^*$. En efecto, primero notemos que $0 \in \operatorname{dom}\theta^*$, pues recordemos que $\theta^*(0) = -\inf\{\theta(x) : \ x \in \mathbb{R}\}$, y dado que θ es finita y positiva se tiene que el ínfimo es finito. Por otra parte, como $\theta_\infty(d) = \sigma_{\operatorname{dom}\theta^*}(d)$ (ver solución de problema 2.6) y por hipótesis $\theta_\infty(1) = +\infty$, se tiene que $\sup\{y : \ y \in \operatorname{dom}\theta^*\} = +\infty$. Luego como θ^* es una función convexa, su dominio también lo es, por lo tanto se concluye que $[0,\infty) \subseteq \operatorname{dom}\theta^*$, y en particular, $1, 1/m \in \operatorname{dom}\theta^*$ y $L \in [0,\infty)$. Más aún, dado que θ es creciente, $\operatorname{ran}\partial\theta \subseteq [0,+\infty)$, y como $\operatorname{dom}\theta^* \subseteq \operatorname{ran}\partial\theta^*$, la afirmación anterior es cierta.

Sigue que, por la parte anterior, para todo $y \in \mathbb{R}^m$ se tiene que

$$|M_r(y) - \max y_i| \leq rL.$$

Luego, fijemos $x \in \mathbb{R}^n$ y tomemos $y = (f_1(x), ..., f_m(x))$ entonces se tiene que

$$|M_r(f_1(x), ..., f_m(x)) - \max\{f_1(x), ..., f_m(x)\}| \leq rL,$$

de donde se concluye que $M_r(f_1(x), ..., f_m(x)) \to \max\{f_1(x), ..., f_m(x)\}$, si $r \to 0$.

(d) Sea $h(x) = \max\{f_1(x), ..., f_m(x)\}$, luego, por la parte (b) tenemos que

$$h(x) - r\theta^*(1) \leq M_r(f_1(x), ..., f_m(x)).$$

Por otra parte, como cada f_k es convexa y finita (en particular continua por estar en $\mathbb{R}^n$) se tiene que h es convexa y s.c.i., más aún como $S(P)$ es no vacío y acotado, se tiene que todo los conjuntos de nivel de h también son acotados, y por lo tanto h es coerciva. En consecuencia, la función $x \mapsto M_r(f_1(x), ..., f_m(x))$ también lo es y por lo tanto existe al menos un punto, que denotaremos por x_r, que minimiza a esta función, es decir,

$$M_r(f_1(x_r), ..., f_m(x_r)) \leq M_r(f_1(x), ..., f_m(x)) \quad \forall x \in \mathbb{R}^n.$$

Nuevamente, por la parte (b) se tiene que, para todo $r > 0$

$$M_r(f_1(x), ..., f_m(x)) \leq h(x) - rm\theta^*(1/m) \quad \forall x \in \mathbb{R}^n,$$
$$h(x_r) - r\theta^*(1) \leq M_r(f_1(x_r), ..., f_m(x_r)).$$

Con todas estas desigualdades obtenemos que $h(x_r) \leq h(x) + r[\theta^*(1) - m\theta^*(1/m)]$. Tomando ínfimo sobre $x \in \mathbb{R}^n$ y $r > 0$ suficientemente pequeño se tiene que

$$h(x_r) \leq \inf_{x \in \mathbb{R}^n} h(x) + 1 =: \gamma \in \mathbb{R},$$

es decir, $x_r \in \Gamma_\gamma(h)$ para todo $r > 0$ y dado que h es coerciva se concluye que $\{x_r\}$ permanece acotada cuando $r \to 0$.

Por otro lado, notemos que

$$\inf_{x \in \mathbb{R}^n} h(x) \leq h(x_r) \leq h(x) + r[\theta^*(1) - m\theta^*(1/m)] \forall x \in \mathbb{R}^n,$$

luego es fácil ver que $v(P_r) \to v(P)$ cuando $r \to 0$. Más aú, si $\bar{x}$ es un punto de acumulación de x_r se tiene que existe una sucesión $r_k \to 0$ tal que $x_{r_k} \to \bar{x}$. Usando la s.c.i de h obtenemos

$$h(\bar{x}) \leq \lim_{k \to \infty} h(x_{r_k}) = \lim_{k \to \infty} v(P_{r_k}) = v(P) \leq h(\bar{x}).$$

(e) Recordemos que, dada la función deperturbación propuesta en el enunciado, el problema dual a (P_r) es de la forma

$$(D_r). \qquad \min_{y^* \in \mathbb{R}^m} \varphi^*(0, y^*)$$

Calculemos $\varphi^*(0, y^*)$. Consideremos, para $x \in \mathbb{R}^n$, $f(x) = (f_1(x), ..., f_m(x))$, luego por definición se tiene

$$\begin{aligned}
\varphi^*(0, y^*) &= \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \{\langle y^*, y\rangle - \varphi(x, y)\} \\
&= \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \left\{\langle y^*, y\rangle - rM\left(\frac{1}{r}(f(x) + y)\right)\right\} \\
&= \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \left\{\langle y^*, y\rangle - \inf_{v \in \mathbb{R}} \left\{v + r\sum_{i=1}^m \theta\left(\frac{f_i(x) + y_i - v}{r}\right)\right\}\right\} \\
&= r \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \sup_{v \in \mathbb{R}} \left\{\sum_{i=1}^m \left[y_i^* \frac{y_i}{r} - \frac{v}{rm} - \theta\left(\frac{f_i(x) + y_i - v}{r}\right)\right]\right\} \\
&= r \sup_{x \in \mathbb{R}^n} \sup_{v \in \mathbb{R}} \left\{\sum_{i=1}^m \left[\sup_{y \in \mathbb{R}} \left\{y_i^* \frac{y}{r} - \theta\left(\frac{f_i(x) + y - v}{r}\right)\right\}\right] - \frac{v}{r}\right\} \\
&= r \sup_{x \in \mathbb{R}^n} \sup_{v \in \mathbb{R}} \left\{\sum_{i=1}^m \left[\theta^*(y_i^*) - y_i^* \frac{f_i(x) - v}{r}\right] - \frac{v}{r}\right\} \\
&= r \sum_{i=1}^m \theta^*(y_i^*) + \sup_{x \in \mathbb{R}^n} \sup_{v \in \mathbb{R}} \left\{\sum_{i=1}^m [y_i^*(v - f_i(x))] - v\right\}.
\end{aligned}$$

Es decir,

$$\varphi^*(0, y^*) = r \sum_{i=1}^{m} \theta^*(y_i^*) + \sup_{v \in \mathbb{R}} \left\{ \left(\sum_{i=1}^{m} y_i^* - 1 \right) v \right\} - \inf_{x \in \mathbb{R}^n} \langle y^*, f(x) \rangle.$$

Por lo tanto,

$$\varphi^*(0, y^*) = r \sum_{i=1}^{m} \theta^*(y_i^*) - \inf_{x \in \mathbb{R}^n} \langle y^*, f(x) \rangle, \quad \text{con } \sum_{i=1}^{m} y_i^* = 1.$$

Por lo tanto, dado que $\operatorname{dom} \theta^* = [0, +\infty)$ se tiene que el problema dual es

$$(D_r) \qquad \min_{y^* \in \mathbb{R}^m} \left\{ r \sum_{i=1}^{m} \theta^*(y_i^*) - \inf_{x \in \mathbb{R}^n} \langle y^*, f(x) \rangle : \sum_{i=1}^{m} y_i^* = 1, y_i^* \geq 0 \right\}.$$

(f) Notemos que por (b) se tiene que, para cualquier $x_0 \in \mathbb{R}^n$,

$$\varphi(x_0, y) \leq \max\{f_i(x_0) + y_i\} - rm\theta^*(1/m) \leq h(x_0) + \|y\|_\infty - rm\theta^*(1/m),$$

en particular, $\varphi(x_0, \cdot)$ es finita y acotada superiormente en una vecindad del origen, luego por el teorema de dualidad se que $v(P_r) + v(D_r) = 0$ y $S(D_r) \neq \emptyset$. La unicidad viene del hecho que al ser θ diferenciable, se tiene que θ^* es estrictamente convexa (problema 2.8), luego el problema dual alcanza su mínimo en un único punto.

Problema 3.3

(a) Por definición se tiene que

$$\begin{aligned} \| \cdot \|^*(x^*) &= \sup_{x \in X} \{ \langle x^*, x \rangle - \|x\| \} \\ &\geq \lambda[\langle x^*, x \rangle - 1], \quad \forall x \in X,\ \|x\| \leq 1, \lambda > 0. \end{aligned}$$

Luego, tomando supremos sobre $x \in X$ tales que $\|x\| \leq 1$ se obtiene que

$$\| \cdot \|^*(x^*) \geq \lambda[\|x^*\| - 1], \quad \forall \lambda > 0.$$

luego haciendo $\lambda \to +\infty$ se obtiene que $\| \cdot \|^* \geq \delta_{B^*}$. Por otra parte, como

$$\langle x^*, x \rangle - \|x\| \leq \|x\| [\|x^*\| - 1],$$

se obtiene que $\| \cdot \|^* \leq \delta_{B^*}$, de donde se concluye que esta desigualdad es una igualdad.

Sea $x \in X \setminus \{0\}$ y consideremos $B = \{x^* \in X^* : \|x^*\| = 1, \langle x^*, x \rangle = \|x\|\}$, Notemos que por el Teorema de Hanh-Banach, este conjunto es no vacío. Sea $x^* \in \partial \| \cdot \|(x)$, luego por definición se tiene que

$$(*) \qquad \langle x^*, y - x \rangle \leq \|y\| - \|x\| \quad \forall y \in X,$$

en particular, evaluando en $y = 0$ e $y = 2x$ se obtiene que

$$\|x\| \leq \langle x^*, x \rangle \leq \|x\|.$$

Más aún, se tiene que $\langle x^*, \bar{x} \rangle = 1$, con $\bar{x} = \frac{x}{\|x\|}$. Luego sigue que $\|x^*\| = 1$, y por lo tanto $x^* \in B$. Conversamente, sea $x^* \in B$, luego para $y \in X$ cualquiera se tiene que

$$\langle x^*, y - x \rangle \leq \|x^*\| \|y\| - \langle x^*, x \rangle \leq \|y\| - \|x\|.$$

(b) En virtud de la parte (a), basta probar que B es igual a

$$C = \{\mu \in \mathcal{M} : \ \|\mu\| = 1, \ \mu^+ \text{ y } \mu^- \text{ están concentradas en } K_f^+ \text{ y } K_f^-, \text{ respectivamente}\}.$$

Sea $\mu \in C$, sigue que

$$\int_K f d\mu = \int_K f d\mu^+ - \int_K f d\mu^- = \int_{K_f^+} f d\mu^+ - \int_{K_f^-} f d\mu^- = \|f\|_\infty [\mu^+(K_f^+) + \mu^-(K_f^-)],$$

pero como $\mu^+(K_f^+) + \mu^-(K_f^-) = \mu^+(K) + \mu^-(K) = \|\mu\| = 1$, se tiene que $\langle \mu, f \rangle = \|f\|_\infty$ y por lo tanto $\mu \in B$.

Recíprocamente, supongamos que $\mu \in B$, luego

$$\int_K f d\mu = \|f\|_\infty,$$

pero

$$\int_K f d\mu = \int_K f d\mu^+ - \int_K f d\mu^- \leq \|f\|_\infty [\mu^+(K) + \mu^-(K)] = \|f\|_\infty,$$

lo que implica que

$$\int_K f d\mu^+ - \|f\|_\infty \mu^+(K) = \int_K f d\mu^- + \|f\|_\infty \mu^-(K).$$

Notemos que el lado izquierdo de esta última igualdad es siempre no positivo mientras que el lado derecho es siempre no negativo. Sigue que ambos términos son iguales a cero y por lo tanto

$$\int_K f d\mu^+ = \|f\|_\infty \mu^+(K) \quad \text{y} \quad \int_K f d\mu^- - \|f\|_\infty \mu^-(K).$$

Supongamos que existe $A \in \mathcal{B}(K)$ tal que $A \cap K_f^+ = \emptyset$ y $\mu^+(A) > 0$. Sigue que

$$\int_K f d\mu^+ = \int_{K \setminus K_f^+} f d\mu^+ \int_{K_f^+} f d\mu^+ < \|f\|_\infty [\mu^+(K \setminus K_f^+) + \mu(K_f^+),$$

ya que $\mu^+(K \setminus K_f^+) \geq \mu^+(A)$. La desigualdad anterior es una contradicción con la definición de B, por lo tanto no existe tal B. El mismo argumento es valido para μ^- y K_f^-, por lo que la igualdad queda establecida.

(c) Notemos que $\varphi : \mathbb{R}^n \times \mathcal{C}(K) \to \mathbb{R} \cup \{+\infty\}$ y por lo tanto $\varphi^* : \mathbb{R}^n \times \mathcal{M}(K) \to \mathbb{R} \cup \{+\infty\}$. Recordemos que el problema dual a (P) está dado por

$$(D) \qquad \min_{\mu \in \mathcal{M}} \varphi^*(0, \mu)$$

Calculemos $\varphi^*(0, \mu)$. Por definición se tiene que

$$\begin{aligned}
\varphi^*(0, \mu) &= \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathcal{C}(K)} \{\langle \mu, y \rangle - \|f + y - \textstyle\sum_{i=1}^n x_i \psi_i\|_\infty\} \\
&= \sup_{x \in \mathbb{R}^n} \{ \sup_{y \in \mathcal{C}(K)} \{\langle \mu, f + y - \textstyle\sum_{i=1}^n x_i \psi_i \rangle - \|f + y - \textstyle\sum_{i=1}^n x_i \psi_i\|_\infty\} - \langle \mu, f - \textstyle\sum_{i=1}^n x_i \psi_i \rangle\} \\
&= \sup_{x \in \mathbb{R}^n} \{\textstyle\sum_{i=1}^n x_i \langle \mu, \psi_i \rangle\} + \| \cdot \|_\infty^*(\mu) - \langle \mu, f \rangle \\
&= \sup_{x \in \mathbb{R}^n} \{\textstyle\sum_{i=1}^n x_i \langle \mu, \psi_i \rangle\} + \delta_{B^*}(\mu) - \langle \mu, f \rangle.
\end{aligned}$$

Luego es claro que

$$\sup_{x\in\mathbb{R}^n}\left\{\sum_{i=1}^n x_i\langle\mu,\psi_i\rangle\right\}\in\mathbb{R} \text{ si y sólo si } \langle\mu,\psi_i\rangle=0, \text{ para todo } i=1,...,n.$$

Por lo tanto el problema dual se puede plantear como

$$(D) \qquad \text{mín}\{\langle\mu,f\rangle:\ \mu\in B^*,\ \langle\mu,\psi_i\rangle=0,\ i=1,...,n\}$$

(d) Notemos que $\text{ínf}(P)\in[0,\|f\|_\infty]$ y además, no es difícil ver que $y\mapsto\varphi(0,y)$ es finita y continua en $y=0$ (de hecho en todo $\mathcal{C}(K)$). Luego por el Teorema de Dualidad se concluye que $v(P)+v(D)=0$ y $S(D)\neq\emptyset$, pues el problema primal es factible.

(e) Supongamos que existe $x\in\mathbb{R}^n\setminus\text{ran}(A)$, notemos que como A es lineal continua, se tiene que $\text{ran}(A)$ es convexo y cerrado. Luego, por el Teorema de Hanh-Banach, existe $\bar{x}\in\mathbb{R}^n\setminus\{0\}$ y $\alpha>0$ tal que

$$\langle\bar{x},x\rangle<\alpha\leq\left\langle\mu,\sum_{i=1}^n\bar{x}_i\psi_i\right\rangle,\quad\forall\mu\in\mathcal{M}(K).$$

No es difícil ver que dado que $\mathcal{M}(K)$ es un espacio vectorial, $\sum_{i=1}^n\bar{x}_i\psi_i=0$ y como $\psi_1,...,\psi_n$ son l.i. se tiene que $\bar{x}=0$, lo que es una contradicción, por lo tanto A es sobreyectivo.

Por otra parte, notemos que el problema dual puede escribirse como

$$(D). \qquad \underset{\mu\in\mathcal{M}(K)}{\text{mín}}\{\langle\mu,f\rangle:|\mu\|\leq 1,\ A(\mu)=0\}$$

Notemos que $A^*:\mathbb{R}^n\to\mathcal{C}(K)$ el operador adjunto está dado por $A^*(x)=\sum_{i=1}^n x_i\psi_i$ y además, como A es sobreyectivo, existe una constante $L>0$ tal que $\|A^*(x)\|_\infty\geq L\|x\|$, por lo tanto, dado que el problema primal se puede escribir

$$(P) \qquad \underset{x\in\mathbb{R}^n}{\text{mín}}\|f-A^*(x)\|_\infty.$$

se tiene que la función objetivo es coerciva y por lo tanto alcanza su mínimo, pues es claramente continua.

(f) Sea $f(x)=\|f-A^*(x)\|_\infty$ y $g(\mu)=\langle\mu,f\rangle+\delta_{B^*}(\mu)+\delta_{\ker(A)}(\mu)$. Luego las condiciones de extremalidad son $0\in\partial f(x)$ y $0\in\partial g(\mu)$. Notemos que $f=h\circ A^*$, con $h:\mathcal{C}(K)\to\mathbb{R}$ dada por $h(\phi)=\|f-\phi\|_\infty$. La función h es convexa y continua, por lo que

$$\partial f(x)=A\partial h(A^*(x))=A\partial\|\cdot\|_\infty(f-A^*(x)),$$

es decir, $x\in S(D)$ si y sólo si existe una medida $\mu\in\mathcal{M}(K)$ tal que $\|\mu\|=1$ y μ^+ y μ^- están concentradas en $K^+_{f-A^*(x)}$ y $K^-_{f-A^*(x)}$, respectivamente. Luego, es claro que si A es sobreyectiva, existe tal medida.

Por otra parte, $\partial g(\mu)=f+N_{B^*\cap\ker(A)}(\mu)$, luego $\mu\in S(D)$ si y sólo $-f\in N_{B^*\cap\ker(A)}(\mu)$, lo que es equivalente a

$$\langle\mu-\nu,f\rangle\leq 0,\quad\forall\nu\in B^*\cap\ker(A).$$

Problema 3.5

(a) Notemos que el problema (P) es equivalente al problema

$$(\tilde{P}) \qquad \min_{y\in B} \tfrac{1}{2}\|y\|^2,$$

donde $B = A(C + \ker A)$. Luego , si B es convexo y cerrado el problema $(\tilde{P})$ tiene solución y por lo tanto $S(P)$ es no vacío. La convexidad de B es directa de la convexidad de A y de la linealidad de A, por lo que resta ver que B es cerrado. Sea $\{y_n\} \subseteq B$ tal que $y_n \to y \in Y$, queremos ver que $y \in B$. Por definición, para todo $n \in \mathbf{N}$, existe $x_n \in C + \ker A$ tal que $y_n = Ax_n$. Por otra parte, como A es sobreyectiva, existe $x \in X$ tal que $y = Ax$. Ahora bien, por el Teorema de la aplicación abierta se tiene que existe una constante $L > 0$ tal que

$$\|x_n - x\| \leq L\|y_n - y\|,$$

es decir, $x_n \to x$ cuando $n \to +\infty$ y como $C + \ker A$ es cerrado se concluye que existe al menos una solución de (P).

(b) Notemos que el problema (P) se puede re-escribir como un problema irrestricto de la siguiente manera

$$(P) \qquad \min_{x\in X} f(x) + g(Ax),$$

donde $f : X \to \mathbb{R} \cup \{+\infty\}$, $g : Y \to \mathbb{R}$ están dadas por $f(x) = \delta_C(x)$ y $g(y) = \frac{1}{2}\|y\|^2$. Recordemos que, $\delta_C^*(x) = \sigma_C(x)$ y que $g^*(y) = g(y)$, luego el dual de Fenchel-Rockafellar de (P) es

$$(D) \qquad \min_{y\in Y} f^*(-A^*y) + g^*(y),$$

más aún, como $\operatorname{dom} g = Y$ e $\inf(P) \in \mathbb{R}$, por las condiciones de calificación se tiene que existe $\bar{y} \in S(D)$. Ahora bien, por el Teorema de Dualidad de Fenchel-Rockafellar, se tiene que $\bar{x} \in S(P)$ si y sólo si

$$-A^*\bar{y} \in \partial f(\bar{x}) \quad \text{y} \quad \bar{y} \in \partial g(A\bar{x}).$$

Recordemos tambieén que $\partial f(x) = N_C(x)$ y que $\partial g(y) = \{y\}$. Luego sigue que $\bar{y} = A\bar{x}$ y que

$$\langle x - \bar{x}, -A^*\bar{y}\rangle \leq 0, \quad \forall x \in C.$$

Luego, fijando $\bar{z} = A^*A\bar{x}$ y usando que A^* es inyectivo (pues A es sobreyectivo) y que por lo tanto admite una inversa por la izquierda, se obtiene el resultado.

(c) En virtud de la parte (a), basta verificar que $C+\ker A$ es cerrado. Sea $\{x_n\} \subseteq C+\ker A$ tal que $x_n \to x \in X$. Luego, existen dos sucesiones $\{u_n\} \subseteq C$ y $\{v_n\} \subseteq \ker A$ tales que $x_n = u_n + v_n$. Ahora bien, como A es sobreyectivo se tiene que existe $L > 0$ tal que $\|x - v_n\| \leq L\|Ax\|$, por lo que la sucesión $\{v_n\}$ permanece acotada, y como $\{x_n\}$ es convergente, la sucesión $\{u_n\}$ también permanece acotada. Luego, como X es reflexivo, existen $u, v \in X$ tales que $u_n \rightharpoonup u$ y $v_n \rightharpoonup v$. Más aún, por definición de C se tiene que $u \in C$. Por otro lado, como $Av_n = 0$, se tiene que

$$\langle y, Av_n\rangle = 0 \quad \text{para todo } y \in Y,$$

pero por definición del operado adjunto se tiene que

$$\langle A^*y, v_n\rangle = 0 \quad \text{para todo } y \in Y,$$

luego, pasando al límite y volviendo a usar la definición del operado adjunto se obtiene que

$$\langle y, Av\rangle = 0 \quad \text{para todo } y \in Y,$$

es decir, $Av \in Y^\perp = \{0\}$, por lo que $v \in \ker A$. Concluimos que $x \in C + \ker A$, ya que en particular, $x_n \rightharpoonup x$ y este límite cuando existe es único.

(d) Directo del Teorema Kuhn-Tucker Fuerte.

Problema 3.7

(a) Primero veamos que si $\bar{x}$ es un punto crítico de Φ entonces también es un punto crítico de Ψ.

Notemos que la función $x \mapsto \langle x, Ax\rangle$ es continua, luego por el Teorema de Moreau-Rockafellar se tiene que para todo $x \in H$

$$\partial\Phi(x) = Ax + \partial f(x) \quad \text{y} \quad \partial\Psi(x) = Ax + \partial(f \circ -A)(x).$$

Luego, sea $\bar{x}$ un punto crítico de Φ, luego $0 \in A\bar{x} + \partial f(\bar{x})$, es decir, $-A\bar{x} \in \partial f(\bar{x})$. Ahora bien, por la Fórmula de Reciprocidad de Legendre se tiene que $\bar{x} \in \partial f^*(-A\bar{x})$.

Por otra parte, sea $\tilde{A} : X \to Y$ operador lineal continuo entre dos espacio de Banach en dualidad con X^* e Y^*, respectivamente. Sea $g \in \Gamma_0(Y)$, notemos que en ausencia de hipótesis de calificación sobre g sólo es posible asegurar

$$\tilde{A}^*\partial g(Ax) \subseteq \partial(g \circ \tilde{A})(x).$$

Sigue que $-A\partial f^*(-A\bar{x}) \subseteq \partial(f^* \circ -A)(\bar{x})$ y por lo tanto $-A\bar{x} \in \partial(f^* \circ -A)(\bar{x})$, es decir, $0 \in A\bar{x} + \partial(f^* \circ -A)(\bar{x}) = \partial\Psi(\bar{x})$.

Ahora bien, como $\partial\Psi(x) = \partial\Psi(x + u)$, con $u \in \ker A$, luego cualquier $\bar{y} \in \ker A + \bar{x}$ es un punto crítico de Ψ.

(b) Notemos que como $0 \in \operatorname{int}(\operatorname{dom} f^* - \operatorname{Im} A)$, en particular se tiene que f^* es continua en $-Ax_0$ para algún $x_0 \in H$, y por lo tanto $-A\partial f^*(-Ax) = \partial(f^* \circ -A)(x)$.

Más aún, sea $\bar{y}$ un punto crítico de Ψ, es decir, $-A\bar{y} \in \partial(f^* \circ -A)(\bar{y}) = -A\partial f^*(-A\bar{y})$. Luego, existe $\bar{x} \in \partial f^*(-Ax)$ tal que $-A\bar{y} = -A\bar{x}$. En particular, $z = \bar{x} - \bar{y} \in \ker A$ y $\bar{x} \in \ker A + \bar{y}$. Aparte, ya que $\bar{x} \in \partial f^*(-A\bar{x})$, por la fórmula de reciprocidad se tiene que $-A\bar{x} \in \partial f(\bar{x})$, es decir, $\bar{x}$ es un punto crítico de Φ.

Por otro lado, por lo anterior se tiene que $-A\bar{y} \in \partial f(\bar{x})$, o equivalentemente

$$f(\bar{x}) + f^*(-A\bar{y}) = \langle \bar{x}, -A\bar{y}\rangle.$$

Por lo tanto,

$$\begin{aligned}\Phi(\bar{x}) &= \frac{1}{2}\langle \bar{x}, A\bar{x}\rangle + f(\bar{x})\\ &= \frac{1}{2}\langle z+\bar{y}, A(z+\bar{y})\rangle + \langle \bar{x}, -A\bar{y}\rangle - f^*(-A\bar{y})\\ &= \frac{1}{2}\langle z+\bar{y}, A\bar{y}\rangle + \langle z+\bar{y}, -A\bar{y}\rangle - f^*(-A\bar{y})\\ &= -\frac{1}{2}\langle A(z+\bar{y}), \bar{y}\rangle - f^*(-A\bar{y})\\ &= -\frac{1}{2}\langle A\bar{y}, \bar{y}\rangle - f^*(-A\bar{y})\\ &= -\Psi(\bar{y}).\end{aligned}$$

Problema 3.8

(a) Denotemos por B la bola unitaria cerrada de X. Sea $f(x) = \langle -z^*, x\rangle + \delta_B(x)$ y $g(x) = \delta_M$. No es difícil ver que $f^*(x^*) = \|x^* + z^*\|_*$ y que $g^*(x^*) = \delta_{M^\perp}(x^*)$. Consideremos el par primal-dual de Fenchel Rockafellar

$$(P) \qquad \min_{x\in X} f(x) + g(x),$$

$$(D) \qquad \min_{x^*\in X^*} f^*(-x^*) + g^*(x^*).$$

Es fácil ver que (D) coincide exactamente con el lado derecho de (3.2), y que además, (P) coincide con el negativo del lado izquierdo de esta misma ecuación. Además es claro que $v(P) \in [0, \|z^*\|]$, y como $\operatorname{dom} f = B$ se tiene que por el Teorema de Dualidad de Fenchel-Rockafellar, el problema (D) tiene solución, y además $v(P) + v(D) = 0$, que es exactamente la ecuación (3.2). Además como $S(D) \neq \emptyset$, en particular, el infimo se alcanza en algún $x_0^* \in M^\perp$.

Nuevamente, por el Teorema de Dualidad de Fenchel-Rockafellar, se tiene que

$$x_0 \in S(P) \quad \text{y} \quad x_0^* \in S(D) \tag{A.3}$$

es equivalente a

$$-x_0^* \in \partial f(x_0) \quad \text{y} \quad x_0^* \in \partial g(x_0). \tag{A.4}$$

Como $\partial f(x_0) = -z^* + N_B(x_0)$ y $\partial g(x_0) = N_M(x_0) = M^\perp$, esto último por ser M un s.e.v. de X. Sigue que (A.3) es equivalente a

$$\langle z^* - x_0^*, x - x_0\rangle \leq 0 \quad \forall x \in B \quad \text{y} \quad x_0^* \in M^\perp. \tag{A.5}$$

Pero,

$$\begin{aligned}\langle z^* - x_0^*, x - x_0\rangle \leq 0 \quad \forall x \in B &\Leftrightarrow \langle z^* - x_0^*, x\rangle \leq \langle z^* - x_0^*, x_0\rangle \quad \forall x \in B\\ &\Leftrightarrow \|z^* - x_0^*\| \leq \langle z^* - x_0^*, x_0\rangle \leq \|z^* - x_0^*\|\|x_0\|.\end{aligned}$$

Como $\|x_0\| \leq 1$, se tiene que $x_0 \in M$ alcanza el supremo del lado izquierdo de (3.2), es decir, $x_0 \in S(P)$ si y sólo si $\langle z^* - x_0^*, x_0\rangle = \|z^* - x_0^*\|\|x_0\|$.

(b) Sea $z^* \in D$, que por hipótesis es no vacío. y consideremos $M = \langle\{x_1, ..., x_n\}\rangle$, el e.v. generado por los vectores $\{x_1, ..., x_n\}$. Reescriendo la ecuación (3.2) en términos de este M y z^* particulares, vemos que el lado izquierdo de (3.3) no depende de la elección de z^* y además coincide con (3.2). Por otra parte, no cuesta mucho convencerse que $M^\perp = D - z^*$ y que por lo tanto

$$\inf_{x^* \in D} \|x^*\|_* = \inf_{x^* \in z^* + M^\perp} \|x^*\|_*,$$

pero

$$\inf_{x^* \in z^* + M^\perp} \|x^*\|_* = \inf_{x^* \in M^\perp} \|x^* - z^*\|_*$$

por lo tanto se concluye que

$$\inf_{x^* \in D} \|x^*\|_* = \inf_{x^* \in M^\perp} \|x^* - z^*\|_*.$$

Luego, las ecuaciones (3.3) y (3.2) coinciden. La alineación de los vectores es consecuencia de la parte anterior.

(c) Tomemos $X = C([0,T])$ y por lo tanto $X^* = NBV([0,T])$, con las respectivas normas del enunciado. Consideremos $n = 1$, $x_1(t) = T - t \in C([0,T])$ y $a_1 = h + \frac{1}{2}g^2T$. Luego es claro que en este caso α es igual al lado derecho de (3.3). Por lo tanto, la primera igual pedida es directa de la parte (b). Para ver la segunda igualdad basta ver que como $a_1 > 0$, $y^* \in \mathbb{R}$ que maximiza la función debe ser positivo y como $\|x_1(t)\|_\infty = T$, se concluye que $y^* = \frac{1}{T}$.

(d) Sean $v \in NBV([0,T])$ y $x \in C([0,T])$, en particular $v(0) = 0$ y $\int_0^T |dv(t)| < \infty$. Recordemos que

$$\int_0^T |dv(t)| = \sup\left\{\sum_{i=0}^n |v(t_{i+1}) - v(t_i)| : \ 0 = t_0 < t_1 < ... < t_n < t_{n+1} = T\right\}.$$

Definamos

$$\Gamma_+ = \{t \in [0,T] : \ x(t) = \|x\|_\infty\} \quad \text{y} \quad \Gamma_- = \{t \in [0,T] : \ x(t) = -\|x\|_\infty\}.$$

Notemos que como x es continua, ambos conjunto son cerrados y por lo tanto se pueden escribir como unión disjunta (a lo más numerable) de intervalos cerrados contenido en $[0,T]$. Luego $\int_{\Gamma_+} |dv(t)|$ y $\int_{\Gamma_-} |dv(t)|$ están bien definidas. Además, es claro que $\Gamma = \Gamma_+ \cup \Gamma_-$. sigue que

$$\begin{aligned}\int_0^T x(t)dv(t) &= \int_{\Gamma_+} \|x\|_\infty dv(t) - \int_{\Gamma_-} \|x\|_\infty dv(t) + \int_{\Gamma^c} x(t)dv(t) \\ &= \|x\|_\infty \left[\int_{\Gamma_+} dv(t) - \int_{\Gamma_-} dv(t)\right] + \int_{\Gamma^c} x(t)dv(t)\end{aligned}$$

Supongamos que v solo varía en Γ, con $v(t)$ no-decreciente si $x(t) > 0$ y no-creciente si $x(t) < 0$, luego $\int_{\Gamma_+} dv(t) = \int_{\Gamma_+} |dv(t)|$ y $-\int_{\Gamma_-} dv(t) = \int_{\Gamma_-} |dv(t)|$, y por lo tanto

$$\int_0^T x(t)dv(t) = \|x\|_\infty \left[\int_{\Gamma_+} dv(t) - \int_{\Gamma_-} dv(t)\right] = \|x\|_\infty \int_\Gamma |dv(t)| = \|x\|_\infty \int_0^T |dv(t)|,$$

por lo que v está alineado con $x \in C([0,T])$.

Conversamente, supongamos que

$$\int_0^T x(t)dv(t) = \|x\|_\infty \int_0^T |dv(t)|.$$

Supongamos también que la variación de v es no nula en Γ, como x es continua, existe un intervalo cerrado de interior no vacío $[a,b] \subseteq \Gamma^c$ tal que $\int_a^b |dv(t)| > 0$ y $|x(t)| \leq x_0 < \|x\|_\infty$ para todo $t \in [a,b]$, por lo tanto

$$\begin{aligned}\int_0^T x(t)dv(t) &= \int_0^a x(t)dv(t) + \int_a^b x(t)dv(t) + \int_0^T x(t)dv(t)\\ &\leq \|x\|_\infty \int_0^a |dv(t)| + x_0 \int_a^b |dv(t)| + \|x\|_\infty \int_b^T |dv(t)|\\ &< \|x\|_\infty \int_0^T |dv(t)|,\end{aligned}$$

lo que es una contradicción, por lo tanto, v varía solamente en Γ.

Supongamos que v no es no decreciente en el conjunto $A_+ = \{t \in [0,T] :\ x(t) > 0\}$. Sea $\tau \in \Gamma_+$, en particular, $\tau \in A_+$, por continuidad de x, existe $\delta > 0$ tal que el intervalo $I = [\text{máx}\{0, \tau - \delta\}, \text{mín}\{\tau + \delta, T\}] \subseteq A_+$ y supongamos que v es estrictamente decreciente en I, luego $\int_{\Gamma_+ \cap I} dv(t) = -\int_{\Gamma_+ \cap I} |dv(t)| < 0$, en particular,

$$\begin{aligned}\int_0^T x(t)dv(t) &= \int_{\Gamma_+ \cap I} x(t)dv(t) + \int_{(\Gamma_+ \cap I^c) \cup \Gamma_-} x(t)dv(t)\\ &= \|x\|_\infty \int_{\Gamma_+ \cap I} dv(t) + \int_{(\Gamma_+ \cap I^c) \cup \Gamma_-} x(t)dv(t)\\ &\leq \|x\|_\infty \left[-\int_{\Gamma_+ \cap I} |dv(t)| + \int_{(\Gamma_+ \cap I^c) \cup \Gamma_-} |dv(t)|\right]\\ &< \|x\|_\infty \int_0^T |dv(t)|,\end{aligned}$$

lo que es una contradicción, por lo tanto, v es no decreciente en A_+. El mismo argumento es aplicable para $A_- = \{t \in [0,T] :\ x(t) < 0\}$. Luego la equivalencia queda establecida.

Por otro lado, supongamos que $\Gamma = \{0 \leq s_1 < s_2 < ... < s_n \leq T\}$ es finito, luego v es contante en los intervalos que definen los punto de Γ, y por lo tanto v es a lo más discontinua en Γ, de donde se obtiene que v solo tiene un número finito de discontinuidad de salto.

(e) Notemos que en este caso $x_1(t) = \|x_1\|_\infty$ sólo en $t = 0$, luego, por la parte (c) y (d), $\Gamma = \{0\}$, de modo que v es constante en $(0,T]$. Además, como $x_1(t) > 0$ en $[0,T)$, por lo que v es no decreciente en $[0,T)$. Así obtenemos que $u = dv$ es un control impulsivo que solo varía en $t = 0$, de modo que puede ser representada por $v(t) = \left(\int_0^T |dv(s)|\right) \chi_{(0,T]}(t)$, pues $v(0) = 0$, donde $\chi_A(t) = 1$ si $t \in A$ y $\chi_A(t) = 0$ si $t \notin A$.

Por otra parte, es fácil ver que la función $T \mapsto \frac{1}{T}(h + \frac{1}{2}gT^2)$ es convexa y que su derivada se anula en $T = \sqrt{\frac{2h}{g}}$. Ahora bien, como v es solución óptima del problema se tiene que

$$\int_0^T |dv(t)| = \frac{1}{T}(h + \frac{1}{2}gT^2) = \sqrt{2gh}.$$

Finalmente, el control óptimo es de la forma.

$$v(t) = \begin{cases} 0 & \text{si } t = 0 \\ \sqrt{2gh} & \text{si } t \in (0, T]. \end{cases}$$

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.

[BaCo11] H.H. Bauschke, P.L. Combettes *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011.

[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.

[BoZ05] J. M. Borwein, Q. J. Zhu, *Techniques of Variational Analysis.*, CMS Books in Mathematics, Springer-Verlag, New York, 2005.

[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.