Análisis Convexo

Jean-Pierre Crouzeix
LIMOS, Université Blaise Pascal
Clermont-Ferrand, France

Eladio Ocaña Anaya
IMCA & FC, Universidad Nacional de Ingeniería
Lima, Perú

Prefacio

El análisis convexo es uno de los fundamentos de la teoría de la optimización. Este es esencialmente geométrico pero, a menudo, abordado desde un punto de vista analítico. Las propiedades topológicas de los conjuntos convexos se derivan de su geometría. Los teoremas fundamentales de la proyección sobre un conjunto convexo y la separación de conjuntos convexos, son geométricos. Las propiedades de continuidad y convexidad de las funciones se derivan de la convexidad de sus epígrafos. Un enfoque puramente analítico enmascara el porqué de los conceptos propios del análisis convexo. Este texto privilegia el enfoque geométrico.

Este texto es una edición revisada y ampliada de la Monografía del IMCA N° 33 escritas por J.-P. Crouzeix, E. Ocaña y W. Sosa. Se han añadido dos capítulos, uno dedicado al análisis convexo matricial y el otro a ejercicios resueltos correspondientes a los temas desarrollados en el texto.

Lima, Mayo del 2017

Jean-Pierre Crouzeix

Eladio Ocaña Anaya

Contenido

Capítulo 1

Convexidad

1.1 Conjuntos convexos

Definición 1.1 (Conjunto convexo) *Un conjunto* $C \subset \mathbb{R}^n$ *es* **convexo** *si para todo* $x, y \in C$ *y todo* $\lambda \in [0, 1]$, *se cumple*

$$tx + (1-t)y \in C.$$

Note que la convexidad es una propiedad unidimensional y no requiere ninguna propiedad topológica.

Definición 1.2 (Combinación convexa) *Sea* $C \subset \mathbb{R}^n$. *Un elemento* $x \in \mathbb{R}^n$ *es una* **combinación convexa** *de* C, *si existen* $p \in \mathbb{N}$, $\{t_i\}_{i=1}^p \subset [0, 1]$ *y* $\{x_i\}_{i=1}^p \subset C$ *tales que*

$$\sum_{i=1}^p t_i = 1 \quad y \quad x = \sum_{i=1}^p t_i x_i.$$

Lema 1.1 *El conjunto de todas las combinaciones convexas de un conjunto, es convexo.*

Demostración : Sean $x = \sum_{i=1}^p \lambda_i x_i$ y $z = \sum_{j=1}^q \mu_j z_j$ combinaciones convexas de un cierto conjunto. Si $t \in [0, 1]$, entonces

$$tx + (1-t)z = \sum_{i=1}^p t\lambda_i x_i + \sum_{j=1}^q (1-t)\mu_j z_j$$

también es una combinación convexa del conjunto. ■

Proposición 1.1 *Un conjunto es convexo si y solo si contiene a todas sus combinaciones convexas.*

Demostración : i) Si C contiene a todas sus combinaciones convexas, entonces este es convexo por el Lema 1.1.

ii) Recíprocamente, sean C convexo y x una combinación convexa de C. Procederemos por inducción, con respecto a p, para demostrar que $x \in C$. Claramente este se cumple para $p = 1$ y 2. Por la hipótesis inductiva, asumamos que se cumple para $p \geq 2$. Mostraremos que también se cumple para $p + 1$. Sea $x = \sum_{i=1}^{p+1} t_i x_i$, donde $\{x_i\}_{i=1}^{p+1} \subset C$ y $\{t_i\}_{i=1}^{p+1} \subset [0, 1]$, con $\sum_{i=1}^{p+1} t_i = 1$. Entonces,

$$x = ty + (1 - t)x_{p+1} \;\; \text{con} \;\; t = \sum_{i=1}^{p} t_i \;\; \text{y} \;\; y = \sum_{i=1}^{p} \frac{t_i}{t} x_i.$$

Por la hipótesis inductiva, $y \in C$, y por lo tanto, por la convexidad de C, $x \in C$. ■

Proposición 1.2 *Sea $\{C_i\}_{i\in I}$ una familia de conjuntos convexos. Entonces $\cap_{i\in I} C_i$ es convexo.*

Demostración : Queda como ejercicio para el lector. ■

Definición 1.3 (Cápsula convexa) *La* **cápsula convexa** *de un conjunto S, denotada por $\operatorname{co}(S)$, es la intersección de todos los conjuntos convexos que contienen a S.*

Por la Proposición 1.2, $\operatorname{co}(S)$ es el conjunto convexo más pequeño que contiene a S.

La siguiente proposición nos da otra forma para describir a la cápsula convexa de un conjunto.

Proposición 1.3 *La cápsula convexa* $\mathrm{co}\,(S)$ *es el conjunto de todas las combinaciones convexas de* S*.*

Demostración : Sea D el conjunto de todas las combinaciones convexas de S. Entonces D es convexo y contiene a S, de donde $\mathrm{co}\,(S) \subset D$. Para la inclusión contraria, sea $z \in D$. Entonces z es una combinación convexa de S y por lo tanto también de $\mathrm{co}\,(S)$. Por la Proposición 1.1, $z \in \mathrm{co}\,(S)$. ■

Corolario 1.1 *Si* $S_1 \subset S_2$*, entonces* $\mathrm{co}\,(S_1) \subset \mathrm{co}\,(S_2)$*.*

Demostración : Queda como ejercicio para el lector. ■

1.1.1 Conos de $\mathbb{R}^n$

Definición 1.4 (Cono) $K \subset \mathbb{R}^n$ *es un* **cono***, si para todo* $x \in K$ *y todo* $\lambda > 0$*, se cumple* $\lambda x \in K$*.*

— *El cono* K *es un* **cono sin punta***, si* $0 \notin K$ *; en caso contrario, este es un* **cono con punta***.*

— *El cono* K *es un* **cono convexo***, si este es además convexo.*

Proposición 1.4 **a)** *Si* K *es un cono, entonces este es convexo si y solo si* $K + K \subset K$*.*

b) *Si* $\{K_i\}_{i\in I}$ *es una familia de conos de* $\mathbb{R}^n$*, entonces* $\cap_{i\in I} K_i$ *es un cono.*

Demostración : Queda como ejercicio para el lector. ■

Definición 1.5 (Cápsula cónica) *La* **cápsula cónica** *de un conjunto* $S \subset \mathbb{R}^n$*, denotada por* $\mathrm{cono}(S)$*, es la intersección de todos los conos que contienen a* S*.*

La siguiente proposición se deduce inmediatamente de la definición.

Proposición 1.5 *Sea* $S \subset \mathbb{R}^n$*. Se cumplen :*

a) $\mathrm{cono}(S) = \{\lambda x : x \in S,\ \lambda > 0\}$.

b) *Si* S *es convexo, entonces* $\mathrm{cono}(S)$ *es convexo.*

Demostración : Queda como ejercicio para el lector. ■

1.1.2 Subespacio afín

$H \subset \mathbb{R}^n$ es un subespacio afín (de $\mathbb{R}^n$) si existe $a \in H$ tal que

$$H - a = \{h - a : h \in H\}$$

es un subespacio vectorial. Se deduce que H es convexo y que

$$H - a = H - b \quad \text{para todo } \; a, b \in H.$$

Proposición 1.6 *Si $\{H_i\}_{i \in I}$ es una familia de subespacios afines de $\mathbb{R}^n$, entonces $\cap_{i \in I} H_i$ es un subespacio afín.*

Demostración : Queda como ejercicio para el lector. ■

Definición 1.6 (Subespacio afín generado) *El subespacio afín generado por $S \subset \mathbb{R}^n$, denotado por* $\operatorname{aff}(S)$*, es la intersección de todos los subespacios afines que contienen a S.*

Proposición 1.7 *Sean $S \subset \mathbb{R}^n$ y $a \in S$. Se cumple :*

i) $\operatorname{aff}(S) = \{a + \sum_{i=1}^{p} t_i (b_i - a) : p \in \mathbb{N},\ \{t_i\}_{i=1}^{p} \subset \mathbb{R},\ \{b_i\}_{i=1}^{p} \subset S\}$.

ii) *Si $S_1 \subset S_2$, entonces* $\operatorname{aff}(S_1) \subset \operatorname{aff}(S_2)$.

iii) $\operatorname{aff}(S) = \operatorname{aff}(\operatorname{co}(S))$.

Demostración : Queda como ejercicio para el lector. ■

Teorema 1.1 *Sean $S \subset \mathbb{R}^n$ y $x_0 \in S$. Si $x \in \operatorname{co}(\mathrm{S})$, entonces existen $p \in \mathbb{N}$, $x_i \in S$ y $t_i > 0$ ($i = 1, \cdots, p$), tales que los p vectores $(x_i - x_0)$ son linealmente independientes, y*

$$x = x_0 + \sum_{i=1}^{p} t_i (x_i - x_0), \quad \sum_{i=1}^{p} t_i \leq 1.$$

Demostración : Por la Proposición 1.3, existen $t_0, \cdots, t_p \geq 0$ y $x_1, \cdots, x_p \in S$ tales que

$$x = t_0 x_0 + \sum_{i=1}^{p} t_i x_i, \quad \sum_{i=0}^{p} t_i = 1,$$

de donde

$$x = x_0 + \sum_{i=1}^{p} t_i (x_i - x_0), \qquad \sum_{i=1}^{p} t_i \leq 1. \tag{1.1}$$

Para completar la demostración, sigamos el siguiente proceso de eliminación :

(1) Eliminar todos los t_i y $(x_i - x_0)$ nulos.

(2) Si los vectores $(x_i - x_0)$ son linealmente independientes, acabó.

(3) Caso contrario, existen $\lambda_1, \cdots, \lambda_p$, no todos nulos, tales que

$$0 = \sum_{i=1}^{p} \lambda_i (x_i - x_0). \tag{1.2}$$

Sin perdida de generalidad, podemos suponer que

$$\sum_{i=1}^{p} \lambda_i \geq 0. \tag{1.3}$$

Sea $i \in \{1, \cdots, p\}$ tal que

$$\frac{\lambda_i}{t_i} = \max\{\frac{\lambda_j}{t_j} : j \in \{1, \cdots, p\}\}.$$

Después de una eventual reenumeración de los subindices, podemos asumir $i = p$. Luego, de (1.3), $\lambda_p > 0$; y de las expresiones (1.1) y (1.2),

$$(x_p - x_0) = -\sum_{i=1}^{p-1} \frac{\lambda_i}{\lambda_p}(x_i - x_0) \quad \text{ y } \quad x = x_0 + \sum_{i=1}^{p-1} (t_i - \frac{t_p \lambda_i}{\lambda_p})(x_i - x_0).$$

Además,

- $t'_i := t_i - \frac{t_p \lambda_i}{\lambda_p} \geq 0$ para todo $i = 1, \cdots, p-1,$ y
- $\sum_{i=1}^{p-1} t'_i = \sum_{i=1}^{p} t_i - \frac{t_p}{\lambda_p} \left[\lambda_p + \sum_{i=1}^{p-1} \lambda_i \right] \leq \sum_{i=1}^{p} t_i \leq 1,$

de donde,

$$x = x_0 + \sum_{i=1}^{p-1} t_i'(x_i - x_0), \quad \sum_{i=1}^{p-1} t_i' \leq 1,$$

con $x_i \in S$ y $t_i \geq 0$, para todo $i = 1, \cdots, p-1$. Retornar a la fase **(1)** del proceso.

En cada etapa se elimina al menos un término en (1.1). Por lo tanto, el proceso termina después de un número finito de etapas. ■

Corolario 1.2 (Teorema de Carathéodory) *Sea $\emptyset \neq S \subset \mathbb{R}^n$. La cápsula convexa de S es el conjunto de todas las combinaciones convexas de a lo más $n+1$ elementos de S.*

Corolario 1.3 *La cápsula convexa de un conjunto compacto, es compacto.*

Demostración : Sea $S \subset \mathbb{R}^n$ compacto. Consideremos

$$T = \{t = (t_0, \cdots, t_n) \in \mathbb{R}^{n+1} : \{t_i\}_{i=0}^{n} \subset [0,1], \ \sum_{i=0}^{n} t_i = 1\},$$

y $f : \mathbb{R}^{n+1} \times (\mathbb{R}^n)^{n+1} \to \mathbb{R}^n$, definida por

$$f(t, x_0, x_1, \cdots, x_n) = \sum_{i=0}^{n} t_i x_i \,.$$

Siendo T y S compactos, y f continua, $\operatorname{co}(S) = f(T \times S^{n+1})$ es compacto. ■

1.2 Propiedades topológicas

1.2.1 Propiedades relativas a los conjuntos abiertos

Definición 1.7 (Interior relativo) *Sea $S \subset \mathbb{R}^n$. El interior relativo de S, denotado por $\operatorname{ri}(S)$, es el interior de S con respecto al subespacio $\operatorname{aff}(S)$, esto es,*

$$\operatorname{ri}(S) = \bigcup_{U \in \mathcal{U}_{rel},\, U \subset S} U \,,$$

donde $\mathcal{U}_{rel} := \{U = W \cap \operatorname{aff}(S) : W$ *es abierto de* $\mathbb{R}^n\}$.

Observe que la inclusión $S_1 \subset S_2$ no implica necesariamente que $\operatorname{ri}(S_1) \subset \operatorname{ri}(S_2)$. Consideremos por ejemplo $S_1 = [0,1] \times \{0\}$ y $S_2 = [0,1] \times [0,1]$, ambos en $\mathbb{R}^2$. Entonces, $\operatorname{ri}(S_1) = (0,1) \times \{0\}$ y $\operatorname{ri}(S_2) = (0,1) \times (0,1)$.

No obstante, si $\operatorname{aff}(S_1) = \operatorname{aff}(S_2)$, entonces $\operatorname{ri}(S_1) \subset \operatorname{ri}(S_2)$ cuando $S_1 \subset S_2$.

Una propiedad topológica fundamental de los conjuntos convexos se da en el siguiente teorema.

Teorema 1.2 *Sea* $C \subset \mathbb{R}^n$ *convexo no vacío. Entonces* $\operatorname{ri}(C)$ *es no vacío.*

Demostración : Asumamos, sin pérdida de generalidad, que C no se reduce a un sólo elemento. Sean $x_0, x_1, \cdots, x_p \in C$ tales que $\{x_i - x_0\}_{i=1}^p$ es una base del subespacio vectorial $\operatorname{aff}(C) - x_0$, y consideremos

$$\begin{aligned} \widetilde{T} &:= \left\{ \sum_{i=0}^{p} t_i x_i : \{t_i\}_{i=0}^p \subset]0,1[\,,\ \sum_{i=0}^{p} t_i = 1 \right\} \\ &= x_0 + \left\{ \sum_{i=1}^{p} t_i (x_i - x_0) : \{t_i\}_{i=1}^p \subset]0,1[\,,\ \sum_{i=1}^{p} t_i < 1 \right\}. \end{aligned}$$

$\widetilde{T}$ está contenido en C y es abierto no vacío de $\operatorname{aff}(C)$, por ser la imagen del conjunto abierto no vacío (de $\mathbb{R}^p$) $\{t = (t_1, \cdots, t_p) \in \mathbb{R}^p_{++} : \sum_{i=1}^p t_i < 1\}$, por medio de la aplicación afín biyectiva

$$(t_1, \cdots, t_p) \in \mathbb{R}^p \to x_0 + \sum_{i=1}^{p} t_i (x_i - x_0).$$

La demostración queda establecida. ■

Corolario 1.4 *Si* $C \subset \mathbb{R}^n$ *es convexo, entonces* $\operatorname{aff}(C) = \operatorname{aff}(\operatorname{ri}(C))$.

Demostración : Con las mismas notaciones de la demostración del teorema precedente, $\operatorname{aff}(\widetilde{T}) = \operatorname{aff}(C)$ y por lo tanto, $\operatorname{aff}(\operatorname{ri}(C)) = \operatorname{aff}(C)$. ■

Proposición 1.8 *Si C es convexo, entonces* $\mathrm{ri}\,(C)$ *es convexo.*

Demostración : Sean $x,\ y \in \mathrm{ri}\,(C)$ y $t \in [0,1]$. Existe $r > 0$ tal que $B(x,r) \cap \mathrm{aff}\,(C) \subset C$ y $B(y,r) \cap \mathrm{aff}\,(C) \subset C$. Se deduce que $B(tx+(1-t)y, r) \cap \mathrm{aff}\,(C) \subset C$ y por lo tanto, $tx+(1-t)y \in \mathrm{ri}\,(C)$. ■

1.2.2 Propiedades relativas a los conjuntos cerrados

Proposición 1.9 *Si $S \subset \mathbb{R}^n$, entonces* $\mathrm{aff}\,(\overline{S}) = \mathrm{aff}\,(S)$.

Demostración : Queda como ejercicio para el lector. ■

Proposición 1.10 *Si $C \subset \mathbb{R}^n$ es convexo, entonces $\overline{C}$ es convexo.*

Demostración : Sean $a,\ b \in \overline{C}$, $t \in [0,1]$ y $\epsilon > 0$. Existen a' y $b' \in C$ tales que $||a-a'|| \leq \epsilon$ y $||b-b'|| \leq \epsilon$, de donde

$$\|a' + t(b'-a') - a - t(b-a)\| \leq (1-t)\|a'-a\| + t\|b-b'\| \leq \epsilon.$$

Se deduce que $a + t(b-a) \in \overline{C}$. ■

Proposición 1.11 *Sea $C \subset \mathbb{R}^n$ convexo no vacío. Si $\bar{x} \in \mathrm{ri}\,(C)$ y $\bar{y} \in \overline{C}$, entonces*

$$\bar{x} + t(\bar{y}-\bar{x}) \in \mathrm{ri}\,(C), \quad \textit{para todo } t \in [0,1[\,.$$

Demostración : Siendo $\mathrm{aff}\,(\mathrm{ri}\,(C)) = \mathrm{aff}\,(C) = \mathrm{aff}\,(\overline{C})$, podemos asumir, sin pérdida de generalidad, que $\mathrm{aff}\,(C) = \mathbb{R}^n$. Sean $t \in]0,1[\,$, $\bar{z} = \bar{x} + t(\bar{y}-\bar{x})$ (asumamos $\bar{x} \neq \bar{y}$), V una vecindad convexa abierta de $\bar{x}$ tal que $\bar{z} \notin V \subset C$, y

$$T := \{y = \bar{z} + r(\bar{z}-x) : r > 0,\ x \in V\} = \bar{z} + \bigcup_{r>0} r(\bar{z}-V).$$

Se cumple $C \cap T \neq \emptyset$ debido a que T es un conjunto abierto conteniendo a $\bar{y}$. Sean $\hat{y} \in T \cap C$ y

$$W := \{z = \hat{y} + t(x-\hat{y}) : t \in]0,1[\,,\ x \in V\} = \hat{y} + \bigcup_{t \in]0,1[} t(V-\hat{y}).$$

W es convexo abierto contenido en C y, debido a que $\bar{z} \in W$, se cumple que $\bar{z} \in \operatorname{int}(C)$. ■

Proposición 1.12 *Sea $C \subset \mathbb{R}^n$ convexo. Entonces*

$$\operatorname{ri}(C) = \operatorname{ri}(\overline{C}) \quad y \quad \overline{\operatorname{ri}(C)} = \overline{C}.$$

Demostración : i) Puesto que $C \subset \overline{C}$ y $\operatorname{aff}(C) = \operatorname{aff}(\overline{C})$, se cumple $\operatorname{ri}(C) \subset \operatorname{ri}(\overline{C})$. Para verificar la inclusión contraria, sean $\bar{z} \in \operatorname{ri}(\overline{C})$ y $\bar{x} \in \operatorname{ri}(C)$. Entonces existen $\bar{y} \in \overline{C}$ y $\bar{t} \in]0,1[$ tales que $\bar{z} = \bar{x} + \bar{t}(\bar{y} - \bar{x})$, de donde, por la Proposición 1.11, $\bar{z} \in \operatorname{ri}(C)$.

ii) Sabemos que $\overline{\operatorname{ri}(\mathrm{C})} \subset \overline{C}$. Para verificar la inclusión contraria, sea $\bar{y} \in \overline{C}$. Para $\bar{x} \in \operatorname{ri}(C)$, se cumple $z_k := \bar{x} + (1 - 1/k)(\bar{y} - \bar{x}) \in \operatorname{ri}(C)$ para todo $k \in \mathbb{N}$. Haciendo $k \to \infty$, obtenemos $\bar{y} \in \overline{\operatorname{ri}(C)}$. ■

1.3 Conos Asintóticos

Sean $C \subset \mathbb{R}^n$ convexo no vacío y $a \in C$. Consideremos el conjunto

$$C_\infty(a) := \{d \in \mathbb{R}^n : a + \lambda d \in C \text{ para todo } \lambda > 0\}.$$

Este es un cono convexo con punta ($0 \in C_\infty(a)$).

En general, $C_\infty(a)$ depende de a, como se muestra por ejemplo considerando $C = \mathbb{R}^2_{++} \cup \{(0,0)\}$. Se tiene

$$C_\infty((0,0)) = C \quad \text{y} \quad C_\infty((1,1)) = \mathbb{R}^2_+.$$

Proposición 1.13 *Sean $C \subset \mathbb{R}^n$ convexo no vacío y $a \in C$. Si C es cerrado, entonces $C_\infty(a)$ es cerrado.*

Demostración : Sean $\lambda > 0$ y $\{d_k\}$ una sucesión en $C_\infty(a)$ tal que $d_k \to d$. Tenemos $a + \lambda d_k \in C$ para todo k, de donde, haciendo $k \to \infty$, $a + \lambda d \in C$. Por lo tanto, $d \in C_\infty(a)$. ■

Definición 1.8 *Dado $C \subset \mathbb{R}^n$ convexo no vacío, el* **cono de recesión** *de C (o cono asintótico de C), denotado por C_∞, es el conjunto*

$$C_\infty := \bigcap_{a \in C} C_\infty(a).$$

Los elementos de C_∞ son llamados **direcciones de recesión** *de C.*

Teorema 1.3 *Sea $C \subset \mathbb{R}^n$ convexo cerrado no vacío. Entonces*

$$C_\infty(a) = C_\infty(b) \quad \text{para todo } a, b \in C.$$

Se deduce que $C_\infty = C_\infty(a)$ para todo $a \in C$. Por la Proposición 1.13, C_∞ es cerrado.

Demostración : Sean $a, b \in C$. Es suficiente demostrar que $C_\infty(a) \subset C_\infty(b)$. Sean $d \in C_\infty(a)$ y $\lambda > 0$. Debido a que $a + k\lambda d \in C$ para todo $k \in \mathbb{N}$, tenemos $(1 - \frac{1}{k})b + \frac{1}{k}a + \lambda d = (1 - \frac{1}{k})b + \frac{1}{k}(a + k\lambda d) \in C$. Haciendo $k \to \infty$, obtenemos $b + \lambda d \in C$, de donde $d \in C_\infty(b)$. ∎

Proposición 1.14 *Sea $C \subset \mathbb{R}^n$ convexo cerrado no vacío. Entonces, C es acotado si y solo si $C_\infty = \{0\}$.*

Demostración : Trivialmente se cumple $C_\infty = \{0\}$ cuando C es acotado. Para demostrar la recíproca, asumamos que C no es acotado. Fijemos $a \in C$. Entonces, para todo $k \in \mathbb{N}$, existe $x_k \in C$ tal que $\|x_k - a\| \geq k$, de donde, para $d_k = \frac{x_k - a}{\|x_k - a\|}$, tenemos $a + \frac{t\|x_k - a\|}{\lambda}(\lambda d_k) = a + t(x_k - a) \in C$ para todo $t \in [0, 1]$, lo que implica que $a + \lambda d_k \in C$ al tomar $t = t_k = \frac{\lambda}{\|x_k - a\|}$ para k suficientemente grande. Debido a que $\|d_k\| = 1$, tenemos $a + \lambda d \in C$, siendo $d(\neq 0)$, un valor de adherencia de la sucesión $\{d_k\}$. Así, C_∞ no se reduce a cero. ∎

Definición 1.9 (Espacio de linealidad) *Sea $C \subset \mathbb{R}^n$ convexo cerrado no vacío. El* **espacio de linealidad** *de C, es la intersección*

$$L := C_\infty \cap (-C_\infty).$$

Este es un subespacio vectorial de $\mathbb{R}^n$ *y se cumple*

$$C = L + (C \cap L^\perp),$$

donde $L^\perp$ *es el subespacio ortogonal de* L*, esto es,*

$$L^\perp = \{y \in \mathbb{R}^n : \langle x, y \rangle = 0 \text{ para todo } x \in L\}.$$

También,

$$C_\infty = L + (C_\infty \cap L^\perp).$$

Proposición 1.15 *Si* $\{C_i\}$ *es una familia de conjuntos convexos cerrados con intersección no vacía, entonces* $(\cap C_i)_\infty = \cap (C_i)_\infty$.

Demostración : Queda como ejercicio para el lector. ∎

1.4 Conjuntos convexos a través de las aplicaciones lineales

Sean $A : \mathbb{R}^n \to \mathbb{R}^p$ lineal, $C \subset \mathbb{R}^n$ y $D \subset \mathbb{R}^p$. Los conjuntos $A(C) := \{y \in \mathbb{R}^p : \exists\, x \in C \text{ con } y = Ax\}$ y $A^{-1}(D) := \{x \in \mathbb{R}^n : Ax \in D\}$ satisfacen las siguientes propiedades :

— Por la continuidad de A : $A(\overline{C}) \subset \overline{A(C)}$ y $A^{-1}(D)$ es abierto (cerrado) cuando D es abierto (cerrado).

— Por la linealidad de A : $A(C)$ es convexo cuando C es convexo y $A^{-1}(D)$ es convexo cuando D es convexo.

Proposición 1.16 *Sean* $C \subset \mathbb{R}^n$ *convexo no vacío y* $A : \mathbb{R}^n \to \mathbb{R}^p$ *lineal. Entonces*

$$A(\mathrm{ri}\,(C)) = \mathrm{ri}\,(A(C)) = \mathrm{ri}\,(A(\overline{C})).$$

Demostración : Se cumple

$$A(\mathrm{ri}\,(C)) \subset A(C) \subset A(\overline{C}) = A(\overline{\mathrm{ri}\,(C)}) \subset \overline{A(\mathrm{ri}\,(C))} \subset \overline{A(C)},$$

de donde $\overline{A(\mathrm{ri}\,(C))} = \overline{A(C)} = \overline{A(\overline{C})}$ y, por la Proposición 1.12,

$$\mathrm{ri}\,(A(C)) = \mathrm{ri}\,(A(\overline{C})) = \mathrm{ri}\,(A(\mathrm{ri}\,(C))) \subset A(\mathrm{ri}\,(C)).$$

Para mostrar la inclusión contraria de la última relación, sea $\bar{y} = A\bar{x} \in A(\mathrm{ri}\,(C))$, con $\bar{x} \in \mathrm{ri}\,(C)$. Para $\hat{x} \in C$, existen $\tilde{x} \in \mathrm{ri}\,(C)$ y $t \in]0,1[$ tales que $\bar{x} = t\hat{x} + (1-t)\tilde{x}$ y por lo tanto $\bar{y} = tA\hat{x} + (1-t)A\tilde{x}$. Tomando $\hat{x} \in C$ tal que $A\hat{x} \in \mathrm{ri}\,(A(C))$, se deduce que $\bar{y} \in \mathrm{ri}\,(A(C))$. ■

Comentario: Si $C \subset \mathbb{R}^n$ es cerrado y $A : \mathbb{R}^n \to \mathbb{R}^p$ lineal, la imagen $A(C)$ no necesariamente es cerrado. Consideremos por ejemplo :

$$C = \{(x,y) \in \mathbb{R}^2_{++} : xy \geq 1\} \quad \text{y} \quad A : \mathbb{R}^2 \to \mathbb{R} \ \text{ tal que } \ A(x,y) = x.$$

Se verifica fácilmente que $A(C) =]0, +\infty[$, que no es cerrado.

Sabemos que $A(C)$ es cerrado (de hecho compacto) si C es compacto. En el siguiente teorema debilitamos la condición de compacidad sobre C.

Teorema 1.4 *Sean $A : \mathbb{R}^n \to \mathbb{R}^p$ lineal y $C \subset \mathbb{R}^n$ convexo no vacío. Si la siguiente condición se cumple :*

$$Ad = 0 \ \ y \ \ d \in (\overline{C})_\infty \ \Rightarrow \ -d \in (\overline{C})_\infty,$$

entonces

a) $\overline{A(C)} = A(\overline{C})$ *; y*

b) $A\left((\overline{C})_\infty\right) = \left(\overline{A(C)}\right)_\infty$.

Demostración : Siempre se cumple $A(C) \subset A(\overline{C}) \subset \overline{A(C)} \subset \overline{A(\overline{C})}$. Sean $D = \overline{C}$ y $\widetilde{D} = D \cap L^\perp$, donde $L^\perp$ el subespacio ortogonal de $L = D_\infty \cap (-D_\infty) \cap \{d : Ad = 0\}$. Tenemos que $\widetilde{D}$ es cerrado, $D = \widetilde{D} + L$ y $A(D) = A(\widetilde{D})$. Además,

$$\begin{aligned}
\widetilde{D}_\infty \cap \{d : Ad = 0\} &= (L^\perp)_\infty \cap D_\infty \cap \{d : Ad = 0\} \\
&= L^\perp \cap D_\infty \cap (-D_\infty) \cap \{d : Ad = 0\} \\
&= \{0\}.
\end{aligned}$$

En lo que sigue asumiremos, sin pérdida de generalidad, que $0 \in D$.

Para la parte **a)**, es suficiente mostrar que $A(D)$ es cerrado. Sean $\bar{y} \in \overline{A(D)} = \overline{A(\widetilde{D})}$ y $\{x_k\}$ una sucesión en $\widetilde{D}$ tal que $Ax_k \to \bar{y}$. Consideremos el conjunto

$$K = \widetilde{D} \cap \{x : \|Ax - \bar{y}\| \leq 1\},$$

que es convexo cerrado no vacío. Este también es acotado debido a que $\{x : \|Ax - \bar{y}\| \leq 1\}_\infty = \{d : Ad = 0\}$ y $K_\infty = (\widetilde{D})_\infty \cap \{d : Ad = 0\} = \{0\}$.

Se deduce que $\{x_k\}$ es acotada y por lo tanto admite un valor de adherencia, $\bar{x} \in \widetilde{D}$, lo cual implica que $\bar{y} = A\bar{x} \in A(\widetilde{D}) = A(D)$.

b) Debido a que $0 \in D$, se cumple $D_\infty \subset D$ y por lo tanto $A(D_\infty) \subset A(D)$. Siendo $A(D_\infty)$ un cono y $0 \in A(D)$, la última inclusión implica que $A(D_\infty) \subset (A(D))_\infty$. Recíprocamente, sea $v \in (A(D))_\infty = (A(\widetilde{D}))_\infty$, con $v \neq 0$. Tenemos $\lambda v \in A(\widetilde{D})$ para todo $\lambda > 0$ debido a que $0 \in \widetilde{D}$. Consideremos

$$K = \{d \in \widetilde{D} : Ad = v\} = \widetilde{D} \cap \{d : Ad = v\},$$

que es convexo cerrado no vacío. Este también es acotado porque $K_\infty = \widetilde{D}_\infty \cap \{d : Ad = 0\} = \{0\}$. De otro lado, debido a que $nv \in A(\widetilde{D})$ (para todo $n \in \mathbb{N}$), existe $d_n \in K$ tal que $nd_n \in \widetilde{D}$. La sucesión $\{d_n\}$ siendo acotada, admite un valor de adherencia, $d \in K$. Mostremos que $d \in \widetilde{D}_\infty$. Si eso no sucede, entonces existe $\lambda > 0$ tal que $\lambda d \notin \widetilde{D}$. De otro lado, como 0 y nd_n están en $\widetilde{D}$, $\lambda d_n \in \widetilde{D}$ para todo $n > \lambda$ y por lo tanto, haciendo $n \to \infty$, $\lambda d \in \widetilde{D}$. Una contradicción. ■

Note que la suma de dos conjuntos convexos cerrados no siempre es cerrado. Consideremos por ejemplo en $\mathbb{R}^2$:

$$C = \{(x, y) : x > 0,\ y > 0,\ xy \geq 1\} \quad \text{y} \quad D = \{(x, y) : y = 0,\ x \leq 0\}.$$

La suma

$$C + D = \{(x, y) : y > 0\},$$

es abierto.

La condición dada en el siguiente resultado es una adaptación a la del teorema precedente con el fin de que la suma de dos conjuntos convexos cerrados sea cerrado.

Corolario 1.5 *Sean C y D convexos de $\mathbb{R}^n$ satisfaciendo la condición*

$$d \in (\overline{C})_\infty \cap (-\overline{D})_\infty \;\Rightarrow\; -d \in (\overline{C})_\infty \cap (-\overline{D})_\infty\,.$$

Entonces

$$\overline{C+D} = \overline{C} + \overline{D} \quad \textit{y} \quad (\overline{C+D})_\infty = (\overline{C})_\infty + (\overline{D})_\infty\,.$$

Demostración : Sea $A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ definida por $A(x,y) = x + y$, la suma en $\mathbb{R}^n$. Entonces $A(C \times D) = C + D$ y, teniendo en cuenta que $\overline{C \times D} = \overline{C} \times \overline{D}$ y $(\overline{C} \times \overline{D})_\infty = (\overline{C})_\infty \times (\overline{D})_\infty$, la condición

$$Ad = 0 \quad \text{y} \quad d \in (\overline{C} \times \overline{D})_\infty \;\Rightarrow\; -d \in (\overline{C} \times \overline{D})_\infty$$

es equivalente a

$$d_1 \in (\overline{C})_\infty \;\text{ y }\; -d_1 \in (\overline{D})_\infty \;\Rightarrow\; -d_1 \in (\overline{C})_\infty \;\text{ y }\; d_1 \in (\overline{D})_\infty.$$

Aplicar el Teorema 1.4. ■

Proposición 1.17 (Corolario de la Proposición 1.16) *Si C y D son convexos de $\mathbb{R}^n$, entonces* $\operatorname{ri}(C) + \operatorname{ri}(D) = \operatorname{ri}(C+D)$.

Demostración : Debido a que $C \times D$ es convexo y $\operatorname{ri}(C \times D) = \operatorname{ri}(C) \times \operatorname{ri}(D)$, el resultado se obtiene aplicando la Proposición 1.16 a la aplicación suma en $\mathbb{R}^n$ ($A(x,y) = x + y$, para todo $x, y \in \mathbb{R}^n$) y al conjunto $C \times D$. ■

Capítulo 2

Funciones semicontinuas y funciones convexas

Dada una función $f : S \to \mathbb{R}$, con $S \subset \mathbb{R}^n$, siempre se puede extender esta a una función $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ definida por

$$g(x) = \begin{cases} f(x) & \text{si } x \in S, \\ +\infty & \text{si } x \notin S. \end{cases}$$

Asociada a una función $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ y $\lambda \in \mathbb{R}$, consideremos los siguientes conjuntos :

$\operatorname{dom}(f) := \{x \in \mathbb{R}^n : f(x) < +\infty\}$	dominio de f
$\operatorname{epi}(f) := \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \lambda\}$	epígrafo de f
$\widetilde{\operatorname{epi}}(f) := \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(x) < \lambda\}$	epígrafo estricto de f
$S_\lambda(f) := \{x \in \mathbb{R}^n : f(x) \leq \lambda\}$	subnivel λ de f
$\widetilde{S}_\lambda(f) := \{x \in \mathbb{R}^n : f(x) < \lambda\}$	subnivel estricto λ de f

Se cumplen las siguientes relaciones :

- $\operatorname{dom}(f) = \bigcup_{\lambda \in \mathbb{R}} S_\lambda(f) = \bigcup_{\lambda \in \mathbb{R}} \widetilde{S}_\lambda(f)$;

- $\mathrm{dom}\,(f) = \mathrm{proj}_{\mathbb{R}^n}(\mathrm{epi}\,(f)) = \mathrm{proj}_{\mathbb{R}^n}(\widetilde{\mathrm{epi}}\,(f))$;
- $\lambda < \mu \quad \Rightarrow \quad \widetilde{S}_\lambda(f) \subset S_\lambda(f) \subset \widetilde{S}_\mu(f) \subset S_\mu(f)$;
- $\lambda < \mu$ y $(x,\lambda) \in \mathrm{epi}\,(f) \quad \Rightarrow \quad (x,\mu) \in \mathrm{epi}\,(f)$;
- $f(x) = \inf[\lambda : (x,\lambda) \in \mathrm{epi}\,(f)] = \inf[\lambda : x \in S_\lambda(f)]$;
- $f(x) = \inf[\lambda : (x,\lambda) \in \widetilde{\mathrm{epi}}\,(f)] = \inf[\lambda : x \in \widetilde{S}_\lambda(f)]$;
- $f_1 \le f_2 \;\Leftrightarrow\; \mathrm{epi}\,(f_1) \supset \mathrm{epi}\,(f_2) \Leftrightarrow S_\lambda(f_1) \supset S_\lambda(f_2)$ para todo $\lambda \in \mathbb{R}$;
- $f_1 \le f_2 \;\Leftrightarrow\; \widetilde{\mathrm{epi}}\,(f_1) \supset \widetilde{\mathrm{epi}}\,(f_2) \;\Leftrightarrow\; \widetilde{S}_\lambda(f_1) \supset \widetilde{S}_\lambda(f_2)$ para todo $\lambda \in \mathbb{R}$;
- $S_\lambda(f) \times \{\lambda\} = \mathrm{epi}\,(f) \bigcap [\mathbb{R}^n \times \{\lambda\}]$ para todo $\lambda \in \mathbb{R}$;
- $\widetilde{S}_\lambda(f) \times \{\lambda\} = \widetilde{\mathrm{epi}}\,(f) \bigcap [\mathbb{R}^n \times \{\lambda\}]$ para todo $\lambda \in \mathbb{R}$.

2.1 Funciones semicontinuas

Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es

— **semicontinua inferior (sci)** en x_0 si para todo $\lambda < f(x_0)$ existe una vecindad V de x_0 tal que $\lambda < f(x)$ para todo $x \in V$;

— **semicontinua superior (scs)** en x_0 si $-f$ es sci en x_0, es decir, si para todo $\lambda > f(x_0)$ existe una vecindad V de x_0 tal que $\lambda > f(x)$ para todo $x \in V$;

— **sci (scs)**, si es sci (scs) en todo $x \in \mathbb{R}^n$.

Teorema 2.1 *f sci $\iff$ $\mathrm{epi}\,(f)$ cerrado $\iff$ $S_\lambda(f)$ cerrado para todo $\lambda \in \mathbb{R}$.*

Demostración : a) f sci $\Rightarrow$ $\mathrm{epi}\,(f)$ cerrado : Mostraremos que $[\mathrm{epi}\,(f)]^c$ es abierto. Sea $(x_0,\lambda_0) \in [\mathrm{epi}\,(f)]^c$. Entonces $\lambda_0 < f(x_0)$, de donde, por la semicontinuidad inferior de f, existe una vecindad V de x_0 tal que $\mu := \frac{1}{2}(f(x_0)+\lambda_0) < f(x)$ para todo $x \in V$. Se deduce que $V \times]-\infty,\mu[$ es una vecindad de (x_0,λ_0) que no intersecta a $\mathrm{epi}\,(f)$.

b) $\mathrm{epi}\,(f)$ cerrado $\Rightarrow S_\lambda(f)$ cerrado : Se deduce del hecho que

$$S_\lambda(f) \times \{\lambda\} = \mathrm{epi}\,(f) \cap [\mathbb{R}^n \times \{\lambda\}].$$

c) $S_\lambda(f)$ cerrado para todo $\lambda \in \mathbb{R} \Rightarrow f$ sci : Si $f(x_0) = -\infty$, entonces f es sci en x_0 por definición. Si $f(x_0) > -\infty$, entonces para todo $\lambda \in \mathbb{R}$ con

$\lambda < f(x_0)$, se cumple $x_0 \notin S_\lambda(f)$ y por lo tanto, por ser $S_\lambda(f)$ cerrado, existe una vecindad V de x_0 tal que $x \notin S_\lambda(f)$ para todo $x \in V$. Es decir, $\lambda < f(x)$ para todo $x \in V$. ■

Las siguientes propiedades se deducen inmediatamente de la definición.

Proposición 2.1 *Sean* $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$.
— *Si* f *y* g *son sci en* x_0*, entonces* $f+g$, $\min(f,g)$ *y* kf*, para todo* $k > 0$*, son sci en* x_0.
— *Si* $\{f_i\}_{i\in I}$ *es una familia de funciones sci en* x_0*, entonces la función* $\sup_{i\in I} f_i$ *definida por* $(\sup_{i\in I} f_i)(x) = \sup_{i\in I} f_i(x)$ *para todo* x*, es sci en* x_0.

Proposición 2.2 *Sean* $C \subset \mathbb{R}^n$ *compacto no vacío y* f *sci en* C*. Entonces existe* $\bar{x} \in C$ *tal que* $f(\bar{x}) \leq f(x)$ *para todo* $x \in C$.

Demostración : Asumamos $f \not\equiv +\infty$. Para todo $\lambda > \alpha = \inf[f(x) : x \in C]$, los conjuntos $K_\lambda := \{x \in C : f(x) \leq \lambda\}$ son compactos no vacíos y satisfacen $K_\mu \supset K_\lambda$ cuando $\mu > \lambda > \alpha$. Se deduce que $K = \cap_{\lambda>\alpha} K_\lambda$ es no vacío y por lo tanto, para todo $\bar{x} \in K$, se cumple $f(\bar{x}) \leq f(x)$ para todo $x \in C$. ■

Sean $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$ y $\varphi : X \times Y \to \overline{\mathbb{R}}$. La función $h : X \to \overline{\mathbb{R}}$ definida por

$$h(x) = \inf_{y\in Y} \varphi(x,y) \quad \text{para todo } x \in X,$$

es llamada **función marginal**.

Proposición 2.3 *Si* Y *es compacto no vacío y* φ *sci en* $(\bar{x}, y)$ *para todo* $y \in Y$*, entonces* h *es sci en* $\bar{x}$ *y*

$$S(\bar{x}) := \{y \in Y \;:\; h(\bar{x}) = \varphi(\bar{x}, y)\}$$

es compacto no vacío.

Demostración : La compacidad de Y y la semicontinuidad inferior de $\varphi(\bar{x}, \cdot)$ implican que $S(\bar{x})$ es compacto no vacío. Sean $\lambda < h(\bar{x})$ y μ tales que $\lambda < \mu < h(\bar{x})$. Para $y \in Y$, tenemos $\mu < \varphi(\bar{x}, y)$ y por lo tanto, existe una vecindad abierta V_y de $\bar{x}$, y una vecindad abierta W_y de y, tales que $\mu < \varphi(x,z)$ para todo $(x,z) \in V_y \times W_y$. Siendo Y compacto y $Y = \cup_{y\in Y} W_y$, existe un subconjunto finito J de Y tal que $Y = \cup_{y\in J} W_y$. El conjunto $V = \cap_{y\in J} V_y$ es una vecindad abierta de $\bar{x}$ y para todo $(x,y) \in V \times Y$, se cumple $\varphi(x,y) > \mu > \lambda$, de donde, $h(x) \geq \mu > \lambda$. ■

2.1.1 Regularización sci

Asociada a $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, consideremos la familia

$$\Im = \{g : \mathbb{R}^n \to \overline{\mathbb{R}} : g \text{ es sci y } g(x) \leq f(x) \ \text{ para todo } \ x \in \mathbb{R}^n\}.$$

Note que $\Im$ es no vacío pues contiene a la función idénticamente $-\infty$.

La **regularización sci** de f, denotada por $\bar{f}$ (o $\mathrm{cl}\,(f)$), es la función $\sup_{g\in\Im} g$ definida por $(\sup_{g\in\Im} g)(x) = \sup_{g\in\Im} g(x)$ para todo x. Esta es la función sci más grande por debajo de f.

Proposición 2.4 $\mathrm{epi}\,(\bar{f}) = \overline{\mathrm{epi}\,(f)}$ *y* $S_\lambda(\bar{f}) = \bigcap_{\mu>\lambda} \overline{S_\mu(f)}$.

Demostración : Por definición, $\mathrm{epi}\,(\bar{f}) \supset \mathrm{epi}\,(f)$, de donde $\mathrm{epi}\,(\bar{f}) \supset \overline{\mathrm{epi}\,(f)}$. Para verificar la inclusión contraria, consideremos la función g definida por $g(x) = \inf[\lambda : (x,\lambda) \in \overline{\mathrm{epi}\,(f)}]$ para todo $x \in \mathbb{R}$. Esta función está en $\Im$ debido a que $\mathrm{epi}\,(g) = \overline{\mathrm{epi}\,(f)}$. Se deduce que $g \leq \bar{f}$, de donde $\mathrm{epi}\,(\bar{f}) \subset \mathrm{epi}\,(g) = \overline{\mathrm{epi}\,(f)}$.

La segunda igualdad se deduce de la primera, los detalles quedan como ejercicio para el lector. ■

Proposición 2.5 *f es sci en x_0 si y solo si $f(x_0) = \bar{f}(x_0)$.*

Demostración : Por definición, $\bar{f} \leq f$. Si $\bar{f}(x_0) < f(x_0)$, elegimos λ tal que $\bar{f}(x_0) < \lambda < f(x_0)$, de donde, $\mathrm{epi}\,(f) \cap (V \times]-\infty,\lambda[) \neq \emptyset$ para toda vecindad V de x_0. Se deduce que para toda vecindad V de x_0 existe $x \in V$ tal que $f(x) < \lambda$, lo que equivale a decir que f no es sci en x_0. Para mostrar la recíproca, sea $\lambda < f(x_0)$. Debido a que $\bar{f}$ es sci y $\bar{f}(x_0) = f(x_0)$, existe una vecindad V de x_0 tal que para todo $x \in V$ se cumple $\lambda < \bar{f}(x)$, de donde, debido a que $\bar{f} \leq f$, la sci de f en x_0. ■

Comentario: El siguiente ejemplo muestra que no siempre se cumple $S_\lambda(\bar{f}) = \overline{S_\lambda(f)}$. Sea

$$f(x) = \begin{cases} x^2 & \text{si } x \neq 0, \\ 1 & \text{si } x = 0. \end{cases}$$

Se verifica fácilmente que

$$\overline{S_0(f)} = \emptyset \quad \text{y} \quad S_0(\bar{f}) = \{0\}.$$

2.2 Funciones convexas

En todo lo que sigue haremos uso de la siguiente convención :

$$\begin{array}{ccccc} \infty + \alpha & = & \alpha + \infty & = & +\infty \quad \forall\, \alpha \in \mathbb{R}, \\ -\infty + \alpha & = & \alpha - \infty & = & -\infty \quad \forall\, \alpha \in \mathbb{R}, \\ (-\infty) + \infty & = & \infty + (-\infty) & = & +\infty. \end{array}$$

Definición 2.1 *Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es* **convexa** *si su epígrafo es convexo.*

Como ejemplo, la función $f : \mathbb{R} \to \overline{\mathbb{R}}$ definida por

$$f(x) = \begin{cases} +\infty & \text{si } x < 0, \\ 1 & \text{si } x = 0, \\ -\infty & \text{si } x > 0, \end{cases}$$

es convexa.

Las siguientes propiedades se deducen inmediatamente de la definición :

i) f es convexa si y solo si $\widetilde{\mathrm{epi}}\,(f)$ es convexo.

ii) f es convexa si y solo si para todo x, $y \in \mathbb{R}^n$ y para todo $t \in]0,1[$, se cumple

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

iii) f es convexa si y solo si para todo $x \in \mathbb{R}^n$ y para todo $d \in \mathbb{R}^n$, la función de una variable real $f_{x,d}$ definida por $f_{x,d}(t) = f(x + td)$, es convexa.

iv) Si f es convexa, entonces $\mathrm{dom}\,(f) = \mathrm{proj}_{\mathbb{R}^n}(\mathrm{epi}\,(f))$ es convexo ; y para todo $\lambda \in \mathbb{R}$, $S_\lambda(f)$ es convexo.

v) Si f es convexa, entonces $\bar{f}$ es convexa.

vi) Si f es convexa y $\lambda > 0$, entonces λf es convexa.

vii) Si $f : \mathbb{R}^n \to \mathbb{R}$ es convexa y $k : \mathbb{R} \to \overline{\mathbb{R}}$ convexa creciente, entonces la composición $k \circ f$ es convexa.

viii) Si f y g son convexas, entonces $f + g$ es convexa.

ix) Si $\{f_i\}_{i \in I}$ es una familia de funciones convexas, entonces la función $\sup_{i \in I} f_i$ es convexa.

Proposición 2.6 *Si $\varphi : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$ es convexa, entonces la función marginal h, definida en la sección precedente, es convexa.*

Demostración : Se deduce del hecho que

$$h(x) < \lambda \Longleftrightarrow \exists\, y \text{ tal que } \varphi(x,y) < \lambda,$$

que es equivalente a $\widetilde{\mathrm{epi}}\,(h) = \mathrm{proj}_{\mathbb{R}^n \times \mathbb{R}}\, \widetilde{\mathrm{epi}}\,(\varphi)$. ■

La siguiente proposición es fundamental.

Proposición 2.7 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa. Si existe x_0 tal que $f(x_0) = -\infty$, entonces $f(x) = -\infty$ para todo $x \in \mathrm{ri}\,(\mathrm{dom}\,(f))$.*

Demostración : Sea $x \in \mathrm{ri}\,(\mathrm{dom}\,(f))$. Entonces existen $t \in]0,1[$ e $y \in \mathrm{dom}\,(f)$ tales que $x = tx_0 + (1-t)y$, de donde $f(x) \leq tf(x_0) + (1-t)f(y)$. Por lo tanto, $f(x) = -\infty$. ■

De esta proposición se deduce que f no puede tomar valor finito fuera de la frontera relativa de su dominio. Afín de evitar tal patología, trabajaremos esencialmente con las funciones propias.

Definición 2.2 *Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es **propia**, si su dominio es no vacío y $f(x) > -\infty$ para todo $x \in \mathbb{R}^n$.*

2.2.1 Funciones convexas de una variable real

Hemos visto que una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es convexa si y solo si para todo $x, d \in \mathbb{R}^n$, la función $f_{x,d} : \mathbb{R} \to \overline{\mathbb{R}}$ definida por $f_{x,d}(t) = f(x + td)$, es convexa. Por lo tanto, similar al caso de conjuntos, la convexidad de funciones es una propiedad unidimensional.

En esta subsección estudiaremos la convexidad de funciones de una sóla variable para luego extenderla, en la siguiente sección, a funciones de varias variables.

Sabemos que si $\theta : \mathbb{R} \to \overline{\mathbb{R}}$ es convexa, entonces $\operatorname{dom}(\theta)$ es convexo y por lo tanto, un intervalo.

Teorema 2.2 *Sea $I \subset \mathbb{R}$ un intervalo no vacío y $\theta : I \to \mathbb{R}$. Cada una de las tres condiciones siguientes es equivalente a la convexidad de θ :*

$$\frac{\theta(c) - \theta(a)}{c - a} \leq \frac{\theta(c) - \theta(b)}{c - b}, \quad \textit{para todo } a, b, c \in I \textit{ tales que } a < b < c. \tag{2.1}$$

$$\frac{\theta(b) - \theta(a)}{b - a} \leq \frac{\theta(c) - \theta(a)}{c - a}, \quad \textit{para todo } a, b, c \in I \textit{ tales que } a < b < c. \tag{2.2}$$

$$\frac{\theta(b) - \theta(a)}{b - a} \leq \frac{\theta(c) - \theta(b)}{c - b}, \quad \textit{para todo } a, b, c \in I \textit{ tales que } a < b < c. \tag{2.3}$$

Demostración : Sabemos que θ es convexa si y solo si para todo $a, b, c \in I$, con $a < b < c$, se cumple

$$\theta(b) \leq t\theta(a) + (1 - t)\theta(c),$$

donde $t = (c-b)/(c-a)$. Esta desigualdad es equivalente a cada una de las tres relaciones siguientes :

$$\begin{aligned} t\,[\,\theta(c)-\theta(a)] &\leq [\,\theta(c)-\theta(b)]. \\ \theta(b)-\theta(a) &\leq (1-t)\,[\,\theta(c)-\theta(a)]. \\ t\,[\theta(b)-\theta(a)] &\leq (1-t)\,[\,\theta(c)-\theta(b)]. \end{aligned}$$

Reemplazar el valor de t en cada una de estas relaciones. ■

Se deduce el siguiente resultado.

Teorema 2.3 *Sean $I \subset \mathbb{R}$ un intervalo no vacío, $\theta : I \to \mathbb{R}$ convexa, y $a, b, c \in I$ con $a < b < c$. Entonces θ es continua en b y admite las derivadas por la izquierda y por la derecha en ese punto. Además*

$$\frac{\theta(a)-\theta(b)}{a-b} \leq \theta'_-(b) \leq \theta'_+(b) \leq \frac{\theta(c)-\theta(b)}{c-b}. \tag{2.4}$$

Demostración : Sean y_1, y_2, x_1, x_2 tales que $a < y_1 < y_2 < b < x_2 < x_1 < c$. Se obtiene la siguiente cadena de desigualdades :

$$\frac{\theta(a)-\theta(b)}{a-b} \leq \frac{\theta(y_1)-\theta(b)}{y_1-b} \leq \frac{\theta(y_2)-\theta(b)}{y_2-b} \leq \cdots$$

$$\cdots \leq \frac{\theta(x_2)-\theta(b)}{x_2-b} \leq \frac{\theta(x_1)-\theta(b)}{x_1-b} \leq \frac{\theta(c)-\theta(b)}{c-b}.$$

Hacer $y \to b_-$ y $x \to b_+$. ■

Corolario 2.1 *Toda función convexa de una variable real es continua en el interior de su dominio.*

Corolario 2.2 *Sean $I \subset \mathbb{R}$ un intervalo no vacío y $\theta : I \to \mathbb{R}$ convexa. Entonces, para todo $a, b, c \in \operatorname{int}(I)$, con $a < b < c$, se cumple :*

$$\theta'_+(a) \leq \frac{\theta(b)-\theta(a)}{b-a} \leq \theta'_-(b) \leq \theta'_+(b) \leq \frac{\theta(c)-\theta(b)}{c-b} \leq \theta'_-(c).$$

Se deduce que si θ es derivable, entonces las dos derivadas laterales coinciden, obteniéndose así las siguientes caracterizaciones (de primer orden) de convexidad.

Proposición 2.8 *Sean $I \subset \mathbb{R}$ un intervalo abierto no vacío y $\theta : I \to \mathbb{R}$ derivable. Las tres condiciones siguientes son equivalentes :*

(a) *θ es convexa en I.*

(b) *$(t-s)(\theta'(t)-\theta'(s)) \geq 0$ para todo $t, s \in I$.*

(c) *$\theta(t) \geq \theta(s) + (t-s)\theta'(s)$ para todo $t, s \in I$.*

Demostración : Si θ es convexa, (b) y (c) se obtienen a partir del Corolario 2.2. Asumamos (b), entonces para $a, b, c \in I$ con $a < b < c$, existen r y s, con $a < r < b < s < c$, tales que

$$\frac{\theta(b)-\theta(a)}{b-a} = \theta'(r) \leq \theta'(s) = \frac{\theta(c)-\theta(b)}{c-b},$$

de donde, por el Teorema 2.2, se obtiene la convexidad de θ. Finalmente, asumamos (c). Entonces, para todo $t \in I$, $\theta(t) = \sup_{s \in I}[\theta(s) + (t-s)\theta'(s)]$. Se deduce que θ es convexa por ser el supremo de funciones afines. ∎

Como consecuencia inmediata de esta proposición tenemos la siguiente caracterización (de segundo orden) de convexidad.

Proposición 2.9 *Sean $I \subset \mathbb{R}$ un intervalo abierto no vacío y $\theta : I \to \mathbb{R}$ dos veces derivable. Entonces θ es convexa si y solo si $\theta''(t) \geq 0$ para todo $t \in I$.*

2.2.2 Funciones convexas de varias variables

En esta parte extenderemos los resultados de la subsección precedente a funciones de varias variables.

Proposición 2.10 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa propia. Entonces f es sci en $\mathrm{ri}\,(\mathrm{dom}\,(f))$ y continua en $\mathrm{int}\,(\mathrm{dom}\,(f))$.*

Demostración : a) Sea $x_0 \in \mathrm{ri}\,(\mathrm{dom}\,(f))$. Si f no es sci en x_0, entonces $\bar{f}(x_0) < f(x_0)$. Sea $\bar{\lambda}$ tal que $\bar{f}(x_0) < \bar{\lambda} < f(x_0)$. Tenemos, $x_0 \in S_{\bar{\lambda}}(\bar{f}) = \cap_{\lambda > \bar{\lambda}} \overline{S_\lambda(f)}$. Sea $S_\lambda(f)$, con $\lambda \in]\bar{\lambda}, f(x_0)[\,$, y $\bar{x} \in \mathrm{ri}\,(S_\lambda(f))$ (tal $\bar{x}$ existe debido a que $\overline{S_\lambda(f)} \neq \emptyset$). Consideremos la función convexa $\theta : \mathbb{R} \to \overline{\mathbb{R}}$ definida por $\theta(t) = f(x_0 + t(\bar{x} - x_0))$. Tenemos $0 \in \mathrm{int}\,(\mathrm{dom}\,(\theta))$ (ya que $x_0 \in \mathrm{ri}\,(\mathrm{dom}\,(f))$ y $\bar{x} \in \mathrm{dom}\,(f))$ y por consiguiente, θ es continua en 0. De otro lado, para todo $t \in]0, 1[\,$, $x_0 + t(\bar{x} - x_0) \in S_\lambda(f)$ y por consiguiente $f(x_0 + t(\bar{x} - x_0)) = \theta(t) \leq \lambda$. Por la continuidad de θ en 0, tenemos $f(x_0) = \theta(0) \leq \lambda$, en contradicción con $\lambda \in]\bar{\lambda}, f(x_0)[\,$. Por lo tanto $\bar{f}(x) = f(x)$ para todo $x \in \mathrm{ri}\,(\mathrm{dom}\,(f))$.

b) Por la parte a), es suficiente mostrar que f es scs en $\mathrm{int}\,(\mathrm{dom}\,(f))$. Sean $x_0 \in \mathrm{int}\,(\mathrm{dom}\,(f))$ y $\lambda > f(x_0)$. Para cada $i = 1, \cdots, n$, consideremos la función $\theta_i : \mathbb{R} \to \overline{\mathbb{R}}$ definida por $\theta_i(t) = f(x_0 + te_i)$, donde $\{e_1, \cdots, e_n\}$ es la base canónica de $\mathbb{R}^n$. Siendo θ_i convexa, con $0 \in \mathrm{int}\,(\mathrm{dom}\,(\theta_i))$, esta es continua en 0. Debido a que $\theta_i(0) = f(x_0) < \lambda$, existe $t_i > 0$ tal que $f(x_0 + te_i) = \theta_i(t) < \lambda$ para todo $t \in [-t_i, t_i]$. Por la convexidad de $\widetilde{S}_\lambda(f)$ se deduce que $\mathrm{co}\,(\{x_0 \pm t_i e_i\}) \subset \widetilde{S}_\lambda(f)$. ■

Proposición 2.11 *Si f es convexa propia, entonces $\bar{f}$ es convexa propia.*

Demostración : Sabemos que si $A : \mathbb{R}^p \to \mathbb{R}^q$ es continua y $C \subset \mathbb{R}^p$, entonces $A(\overline{C}) \subset \overline{A(C)}$. Haciendo $C = \mathrm{epi}\,(f)$ y $A = \mathrm{proj}_{\mathbb{R}^n}$, tenemos $\mathrm{dom}\,(f) \subset \mathrm{dom}\,(\bar{f}) \subset \overline{\mathrm{dom}\,(f)}$ y por lo tanto $\mathrm{ri}\,(\mathrm{dom}\,(f)) = \mathrm{ri}\,(\mathrm{dom}\,(\bar{f}))$. Si $\bar{f}(x_0) = -\infty$, entonces $f(x) = \bar{f}(x) = -\infty$ para todo $x \in \mathrm{ri}\,(\mathrm{dom}\,(f))$. Una contradicción con el hecho que f es propia. ■

Proposición 2.12 *Sean $C \subset \mathbb{R}^n$ convexo abierto no vacío y $f : C \to \mathbb{R}$ diferenciable. Las tres propiedades siguientes son equivalentes :*

- **(a)** *f es convexa en C.*
- **(b)** *$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ para todo $x, y \in C$.*
- **(c)** *$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$ para todo $x, y \in C$.*

Proposición 2.13 *Sean C un convexo abierto no vacío de $\mathbb{R}^n$ y $f : C \to \mathbb{R}$ dos veces diferenciable. Entonces f es convexa si y solo si la matriz Hessiana $\nabla^2 f(x)$ es semi-definida positiva en todo $x \in C$.*

Comentario: Por la Proposición 2.10, una función convexa es continua en el interior de su dominio. El siguiente ejemplo muestra que una función convexa, aún siendo semicontinua inferior, esta puede no ser continua en la frontera de su dominio.

Sea $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ defnida por

$$f(x,y) = \begin{cases} \frac{x^2}{y} & \text{si } \; y > 0, \\ +\infty & \text{en otro caso.} \end{cases}$$

Para todo $(x,y) \in \mathbb{R}\times \,]0,\infty[\,$,

$$\nabla^2 f(x,y) = \frac{2}{y^3} \begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix}$$

es semidefinida positiva, de donde la convexidad de f en todo $\mathbb{R}\times]0,+\infty[\,$. De otro lado, para cada $\mu \in \mathbb{R}$,

$$S_\mu(f) = \begin{cases} \{(x,y) : y > 0, \; x^2 - \mu y \leq 0\} & \text{si } \; \mu \geq 0, \\ \\ \emptyset & \text{si } \; \mu < 0, \end{cases}$$

y por lo tanto, por la Proposición 2.4, para cada $\lambda \in \mathbb{R}$,

$$S_\lambda(\bar{f}) = \bigcap_{\mu > \lambda} \overline{S_\mu(f)} = \begin{cases} \{(x,y) : y \geq 0, \; x^2 - \lambda y \leq 0\} & \text{si } \; \lambda \geq 0, \\ \\ \emptyset & \text{si } \; \lambda < 0. \end{cases}$$

Se deduce que

$$\bar{f}(x,y) = \inf[\lambda : (x,y) \in S_\lambda(\bar{f})] = \begin{cases} \frac{x^2}{y} & \text{si } \; y > 0, \\ 0 & \text{si } \; x = y = 0, \\ +\infty & \text{en otro caso.} \end{cases}$$

Siendo $\bar{f}$ sci, este no es scs en $(0,0)$ debido a que $\bar{f}(\frac{1}{k}, \frac{1}{k^2}) = 1$ para todo $k \in \mathbb{N}$.

Proposición 2.14 *Sea f convexa propia. Entonces*

$$\operatorname{aff}(S_\lambda(f)) = \operatorname{aff}(\operatorname{dom}(f)) \quad \textit{para todo} \;\; \lambda > \inf_x f(x).$$

Demostración : En general $\operatorname{aff}(S_\lambda(f)) \subset \operatorname{aff}(\operatorname{dom}(f))$. Con el fin de mostrar la inclusión contraria, mostraremos que existe un abierto no vacío de $\operatorname{aff}(\operatorname{dom}(f))$ contenido en $\operatorname{aff}(S_\lambda(f))$. En efecto, sean $x_0 \in \operatorname{dom}(f)$ tal que $f(x_0) < \lambda$ y $x_1 \in \operatorname{ri}(\operatorname{dom}(f))$. Debido a que $f(x_t = (x_0 + t(x_1 - x_0)) \leq f(x_0) + t[f(x_1) - f(x_0)]$ para todo $t \in [0,1]$, existe $\bar{t} \in]0,1[$ tal que $f(x_{\bar{t}}) < \lambda$. Siendo $x_{\bar{t}}$ un elemento de $\operatorname{ri}(\operatorname{dom}(f))$, la restricción de f sobre $\operatorname{aff}(\operatorname{dom}(f))$ es continua en $x_{\bar{t}}$ y en consecuencia, $S_\lambda(f)$ contiene una vecindad de $x_{\bar{t}}$ en la topología de $\operatorname{aff}(\operatorname{dom}(f))$. ∎

2.3 Funciones asintóticas

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces $F = \operatorname{epi}(f)$ es convexo cerrado no vacío y por lo tanto F_∞ es un cono convexo cerrado, con $(0,1) \in F_\infty$. Además, se cumple la siguiente relación :

$$(d, \lambda) \in F_\infty \;\; \text{y} \;\; \mu > \lambda \Rightarrow \; (d, \mu) \in F_\infty.$$

Se deduce que F_∞ es el epígrafo de una función convexa sci, llamada la función asintótica (o función de recesión) de f, denotada por f_∞ (o por $f0^+$).

Note que $f_\infty(d) > -\infty$ para todo $d \in \mathbb{R}^n$ debido a que f es propia. Por lo tanto, f_∞ es convexa, sci y propia.

Proposición 2.15 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Se cumplen*

$$f_\infty(0) = 0 \;\; \textit{y} \;\; f_\infty(kd) = k f_\infty(d) \;\; \textit{para todo} \; d \in \mathbb{R}^n \; \textit{y} \;\; k > 0.$$

Demostración : a) Mostremos primeramente que $f_\infty(0) = 0$. La desigualdad $f_\infty(0) \leq 0$ se deduce del hecho que $(0,0) \in (\operatorname{epi}(f))_\infty$. Para mostrar la desigualdad contraria, sea $a \in \operatorname{dom}(f)$. Entonces, por la desigualdad anterior, $(a + \lambda 0, f(a) + \lambda f_\infty(0)) = (a, f(a)) + \lambda(0, f_\infty(0)) \in \operatorname{epi}(f)$ para todo

$\lambda > 0$, de donde $f(a) \leq f(a) + \lambda f_\infty(0)$ para todo $\lambda > 0$. Se deduce que $f_\infty(0) \geq 0$.

b) Para mostrar la segunda igualdad, sean $d \in \mathbb{R}^n$ y $k > 0$. i) Si $f_\infty(d) = \infty$, entonces $(d, \lambda) \notin \operatorname{epi}(f_\infty)$ para todo $\lambda \in \mathbb{R}$ lo que equivale (por ser $\operatorname{epi}(f_\infty)$ un cono) a $(kd, \xi) \notin \operatorname{epi}(f_\infty)$ para todo $\xi \in \mathbb{R}$. Se deduce que $f_\infty(kd) = \infty$. ii) Si $f_\infty(d) \in \mathbb{R}$, entonces $(d, f_\infty(d)) \in (\operatorname{epi}(f))_\infty$ y por lo tanto $(kd, kf_\infty(d)) \in (\operatorname{epi}(f))_\infty$. Se deduce que $f_\infty(kd) \leq kf_\infty(d)$. Para mostrar la desigualdad contraria, por la desigualdad anterior tenemos $(kd, f_\infty(kd)) \in (\operatorname{epi}(f))_\infty$ lo que es equivalente a $(d, \frac{1}{k} f_\infty(kd)) \in (\operatorname{epi}(f))_\infty$. Se deduce que $f_\infty(d) \leq \frac{1}{k} f_\infty(kd)$. ■

Proposición 2.16 *Sean* $a \in \operatorname{dom}(f)$ *y* $d \in \mathbb{R}^n$. *Entonces*

$$\begin{aligned} f_\infty(d) &= \sup_{k>0} \frac{f(a+kd) - f(a)}{k}, \\ &= \lim_{k\to+\infty} \frac{f(a+kd) - f(a)}{k}, \\ &= \lim_{k\to+\infty} \frac{f(a+kd)}{k}. \end{aligned}$$

Demostración : Sea $d \in \mathbb{R}^n$. Distinguimos dos casos : i) $f_\infty(d) = \infty$: Para todo $\alpha \in \mathbb{R}$, tenemos $(d, \alpha) \notin \operatorname{epi}(f_\infty)$, lo que implica que para algún $\bar{k} > 0$,

$$\alpha < \frac{f(a+\bar{k}d) - f(a)}{\bar{k}} \leq \sup_{k>0} \frac{f(a+kd) - f(a)}{k}.$$

Por lo tanto,

$$f_\infty(d) = \infty = \sup_{k>0} \frac{f(a+kd) - f(a)}{k}.$$

ii) $f_\infty(d) \in \mathbb{R}$: Para todo $k > 0$, tenemos $(a + kd, f(a) + kf_\infty(d)) \in \operatorname{epi}(f)$, lo que implica que

$$\frac{f(a+kd) - f(a)}{k} \leq f_\infty(d).$$

Por lo tanto,

$$\sup_{k>0} \frac{f(a+kd) - f(a)}{k} \leq f_\infty(d).$$

Para demostrar la desigualdad contraria, sea $m \geq \sup_{k>0} \frac{f(a+kd)-f(a)}{k}$. Tenemos, para todo $k > 0$, $(a+kd, f(a)+km) \in \mathrm{epi}\,(f)$, lo que equivale a decir que $(d, m) \in (\mathrm{epi}\,(f))_\infty$ o $f_\infty(d) \leq m$.

Se deduce por lo tanto que

$$\sup_{k>0} \frac{f(a+kd)-f(a)}{k} \geq f_\infty(d).$$

Finalmente, las dos últimas igualdades de la proposición se verifican debido a que la función $]0, \infty[\, \ni k \to \frac{f(a+kd)-f(a)}{k}$, es creciente. ■

Ejemplo 2.1 *1) Sea $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = e^x$. Entonces*

$$f_\infty(d) = \begin{cases} 0 & \text{si } d \leq 0, \\ +\infty & \text{en otro caso.} \end{cases}$$

2) Sea $f : \mathbb{R}^n \to \mathbb{R}$, definida por $f(x) = \frac{1}{2}\langle x, Ax\rangle - \langle b, x\rangle$ con A simétrica semidefinida positiva. Entonces

$$f_\infty(d) = \begin{cases} -\langle b, d\rangle & \text{si } Ad = 0, \\ +\infty & \text{en otro caso.} \end{cases}$$

3) Sea $f : \mathbb{R}^n \to \mathbb{R}$, definida por $f(x) = \|x\|$. Entonces

$$f_\infty(d) = \|d\| \quad \text{para todo } d \in \mathbb{R}^n.$$

Teorema 2.4 *Sea f convexa sci y propia. Entonces $S_0(f_\infty) = [S_\lambda(f)]_\infty$ para todo $\lambda \in \mathbb{R}$ tal que $S_\lambda(f) \neq \emptyset$.*

Demostración : Sea $a \in S_\lambda(f)$. Tenemos $d \in S_0(f_\infty)$ si y solo si $(d, 0) \in \mathrm{epi}\,(f_\infty)$ si y solo si $(a+kd, \lambda) \in \mathrm{epi}\,(f)$ para todo $k > 0$ si y solo si $a + kd \in S_\lambda(f)$ para todo $k > 0$ si y solo si $d \in [S_\lambda(f)]_\infty$. ■

Definición 2.3 *Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es* **ínfimo compacto** *(inf-compacto), si para todo $\lambda \in \mathbb{R}$, el conjunto $S_\lambda(f)$ es compacto.*

Se deduce de la definición que inf-compacto implica semicontinuidad inferior.

La propiedad fundamental de las funciones inf-compactos es la siguiente

Proposición 2.17 *Asuma f inf-compacto y sean $m = \inf_{x\in\mathbb{R}^n} f(x)$ y $S = \{x \in \mathbb{R}^n : m = f(x)\}$. Si $m < \infty$, entonces S es compacto no vacío.*

Demostración : Note que

$$S = \bigcap_{\lambda > m} S_\lambda(f) = \bigcap_{i=1}^{\infty} S_{\lambda_i}(f),$$

donde $\lambda_1 > \lambda_2 > \cdots$, con $\lambda_i \to m$. Se deduce que S es compacto no vacío por ser una intersección de compactos encajados no vacíos. ∎

Proposición 2.18 *Sea f convexa sci propia. Entonces*

a) *f es inf-compacto si y solo si $S_0(f_\infty) = \{0\}$,*

b) *f es inf-compacto si y solo si existe $\lambda \in \mathbb{R}$ tal que $S_\lambda(f)$ es compacto no vacío.*

Demostración : Aplicar el Teorema 2.4 y el hecho que un conjunto convexo cerrado C es compacto si y solo si $C_\infty = \{0\}$. ∎

La siguiente proposición es una caracterización de la inf-compacidad para funciones no necesariamente convexas.

Proposición 2.19 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ sci. Entonces f es inf-compacto si y solo si para toda sucesión (x_k) en $\mathbb{R}^n$, con $\|x_k\| \to +\infty$, implica $f(x_k) \to +\infty$.*

Demostración : Por hipótesis, $S_\lambda(f)$ es cerrado para todo λ. Si f no es inf-compacto, entonces existe λ tal que $S_\lambda(f)$ no es acotado, lo que implica que existe una sucesión $\{x_k\}$ en $\mathbb{R}^n$, con $\|x_k\| \to +\infty$, tal que $f(x_k) \leq \lambda$ para todo k. Recíprocamente, si existe una sucesión $\{x_k\}$ en $\mathbb{R}^n$ tal que $\|x_k\| \to +\infty$ y $\lim f(x_k) < \lambda$, con $\lambda \in \mathbb{R}$, entonces $S_\lambda(f)$ no es acotado. ∎

Proposición 2.20 *Sean $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$, con f inf-compacto y g sci acotada inferiormente (es decir, existe $\beta > -\infty$ tal que $g(x) \geq \beta$ para todo $x \in \mathbb{R}^n$). Entonces $f + g$ es inf-compacto.*

Demostración : Se deduce del hecho que $f(x) + g(x) \leq \lambda$ implica $f(x) \leq \lambda - \beta$. ■

Definición 2.4 *Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es **estrictamente convexa** si para todo $t \in]0, 1[$ y para todo $a, b \in \mathbb{R}^n$, con $a \neq b$, se cmple*

$$f(ta + (1-t)b) < tf(a) + (1-t)f(b).$$

Cuando f es diferenciable tenemos las siguientes caracterizaciones de la estricta convexidad :

Proposición 2.21 *Sean $C \subset \mathbb{R}^n$ convexo abierto no vacío y $f : C \to \mathbb{R}$ diferenciable. Las siguientes tres afirmaciones son equivalentes :*

(a) *f es estrictamente convexa ;*

(b) *$\langle \nabla f(x) - \nabla f(y), x - y \rangle > 0$ para todo $x, y \in C$, $x \neq y$;*

(c) *$f(y) > f(x) + \langle \nabla f(x), y - x \rangle$ para todo $x, y \in C$, $x \neq y$.*

Se deduce la siguiente condición suficiente de segundo orden. Esta no necesaria : considere por ejemplo la función $\theta(t) = t^4$.

Proposición 2.22 *Sea $C \subset \mathbb{R}^n$ convexo abierto no vacío y $f : C \to \mathbb{R}$ dos veces diferenciable. Si para todo $x \in C$ la matriz $\nabla^2 f(x)$ es definida positiva, entonces f es estrictamente convexa.*

Definición 2.5 *Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es **fuertemente convexa, de coeficiente** $\alpha > 0$, si para todo $t \in [0, 1]$ y para todo $x, y \in \mathbb{R}^n$,*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\alpha}{2} t(1-t)\|x - y\|^2.$$

Observe que fuerte convexidad implica estricta convexidad.

La siguiente caracterización es bastante manejada.

Proposición 2.23 *Sea $a \in \mathbb{R}^n$ fijo arbitrario. Entonces f es fuertemente convexa de coeficiente $\alpha > 0$ si y solo si la función g definida por*

$$g(x) = f(x) - \frac{\alpha}{2}\|x - a\|^2 \quad \textit{para todo } x \in \mathbb{R}^n,$$

es convexa.

Demostración : Para $x, y \in \mathbb{R}^n$ y $t \in [0, 1]$, sean

$$\begin{aligned} u &= tf(x) + (1-t)f(y) - f(tx + (1-t)y); \\ v &= tg(x) + (1-t)g(y) - g(tx + (1-t)y); \\ w &= \frac{\alpha}{2}[\,t\|x-a\|^2 + (1-t)\|y-a\|^2 - \|tx + (1-t)y - a\|^2\,]. \end{aligned}$$

Es fácil comprobar que $u = v + w$, y

$$w = \frac{\alpha}{2}t(1-t)\|x - y\|^2.$$

Se deduce el resultado. ■

De esta proposición derivamos inmediatamente las siguientes caracterizaciones de fuerte convexidad de primer y segundo orden, aplicando las Proposiciones 2.12 y 2.13 a la función g, con $a = 0$. Tenemos

$$\nabla f(x) = \nabla g(x) + \alpha x \quad \text{y} \quad \nabla^2 f(x) = \nabla^2 g(x) + \alpha I.$$

Proposición 2.24 *Sean C convexo abierto no vacío de $\mathbb{R}^n$ y $f : C \to \mathbb{R}$ diferenciable. Las siguientes tres afirmaciones son equivalentes :*

- **(a)** *f es fuertemente convexa de coeficiente $\alpha > 0$;*
- **(b)** *$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \alpha \|x - y\|^2$ para todo $x, y \in C$;*
- **(c)** *$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$ para todo $x, y \in C$.*

Proposición 2.25 *Sean C convexo abierto no vacío de $\mathbb{R}^n$ y $f : C \to \mathbb{R}$ dos veces diferenciable. Entonces f es fuertemente convexa de coeficiente $\alpha > 0$ si y solo si para todo $x \in C$, los valores propios de la matriz Hessiana $\nabla^2 f(x)$, son mayores o iguales que α.*

El siguiente resultado se refiere a la existencia de solución óptima de un problema de optimización (minimización).

Teorema 2.5 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Si f es fuertemente convexa, entonces f es inf-compacto.*

Demostración : Sean α el coeficiente de fuerte convexidad de f y $\bar{x} \in \operatorname{dom}(f)$. La función $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ definida por $g(x) = f(x) - \frac{\alpha}{2}\|x - \bar{x}\|^2$ es convexa sci propia. Para $0 \neq d \in \mathbb{R}^n$, tenemos

$$f_\infty(d) = \sup_{t>0} \frac{g(\bar{x} + td) - g(\bar{x})}{t} + \frac{\alpha t}{2}\|d\|^2$$

y, por la convexidad de g,

$$\frac{g(\bar{x} + td) - g(\bar{x})}{t} \geq \frac{g(\bar{x} + d) - g(\bar{x})}{1} \qquad \text{para todo } \ t \geq 1.$$

Se deduce que $f_\infty(d) = \infty$. ∎

Finalizamos esta sección con una aplicación al principio variacional de Ekeland.

Teorema 2.6 *Sea $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable tal que $-\infty < m := \inf f(x)$. Para $\epsilon > 0$, sea x_ϵ tal que $f(x_\epsilon) \leq m+\epsilon$. Entonces existe $\bar{x}$ tal que $\|\bar{x}-x_\epsilon\| \leq \sqrt{\epsilon}$, $\|\nabla f(\bar{x})\| \leq 2\sqrt{\epsilon}$ y $f(\bar{x}) \leq m + \epsilon$.*

Demostración : Consideremos $g(x) = f(x) + \frac{\alpha}{2}\|x - x_\epsilon\|^2$, con $\alpha > 0$. Como f es sci y la función $x \to \frac{\alpha}{2}\|x - x_\epsilon\|^2$ inf-compacto, entonces g es inf-compacto y por lo tanto existe $\bar{x}$ tal que $g(\bar{x}) \leq g(x)$ para todo x. Se sigue que $\nabla f(\bar{x})+\alpha(\bar{x}-x_\epsilon) = \nabla g(\bar{x}) = 0$ y en consecuencia $\|\nabla f(\bar{x})\| = \alpha\|\bar{x}-x_\epsilon\|$.

De otro lado,

$$m + \frac{\alpha}{2}\|\bar{x} - x_\epsilon\|^2 \leq f(\bar{x}) + \frac{\alpha}{2}\|\bar{x} - x_\epsilon\|^2 = g(\bar{x}) \leq g(x_\epsilon) = f(x_\epsilon) \leq m + \epsilon.$$

Se deduce por lo tanto que

$$\|\bar{x} - x_\epsilon\|^2 \leq \frac{2\epsilon}{\alpha} \quad \text{ y } \quad f(\bar{x}) \leq m + \epsilon.$$

Considerar $\alpha = 2$. ∎

Capítulo 3

Conjugación y Subdiferencial

3.1 Teoremas de separación

$H \subset \mathbb{R}^n$ es un **hiperplano** si este es de la forma

$$H = H(a, \alpha) = \{x \in \mathbb{R}^n : \langle a, x \rangle = \alpha\},$$

donde $0 \neq a \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$. Se deduce que H es un subespacio afín de dimensión $n-1$ y por lo tanto, divide a $\mathbb{R}^n$ en dos semiespacios cerrados :

$$E_1 = \{x \in \mathbb{R}^n : \langle a, x \rangle \geq \alpha\} \quad \text{y} \quad E_2 = \{x \in \mathbb{R}^n : \langle a, x \rangle \leq \alpha\}.$$

Sean $S_1, S_2 \subset \mathbb{R}^n$ no vacíos. Se dice que el hiperplano $H = H(a, \alpha)$

— **separa** a S_1 y S_2 si

$$\langle a, x_1 \rangle \leq \alpha \leq \langle a, x_2 \rangle \ \text{ para todo } \ (x_1, x_2) \in S_1 \times S_2.$$

— **separa propiamente** a S_1 y S_2 si existe separación y además existen $x_1 \in S_1$, $x_2 \in S_2$ tales que

$$\langle a, x_2 - x_1 \rangle > 0,$$

ie, $S_1 \cup S_2$ no está contenido en H.

— **separa estrictamente** a S_1 y S_2 si

$$\langle a, x_1 \rangle < \alpha < \langle a, x_2 \rangle \ \text{ para todo } \ x_1 \in S_1 \text{ y } x_2 \in S_2.$$

— **separa fuertemente** a S_1 y S_2 si existen α_1, $\alpha_2 \in \mathbb{R}$ tales que

$$\langle a, x_1 \rangle \leq \alpha_1 < \alpha < \alpha_2 \leq \langle a, x_2 \rangle \quad \text{para todo } \; x_1 \in S_1 \text{ y } x_2 \in S_2.$$

Estudiaremos ahora la separación de los conjuntos convexos con intersección vacía. Note que

$$S_1 \cap S_2 = \emptyset \;\text{ si y solo si }\; 0 \notin S_2 - S_1.$$

Los primeros resultados de separación caracterizan la separación del origen con un conjunto convexo.

Teorema 3.1 *Sea $C \subset \mathbb{R}^n$ convexo cerrado no vacío tal que $0 \notin C$. Entonces existen $a \in \mathbb{R}^n$ y $\alpha > 0$ tales que $\langle a, x \rangle \geq \alpha$ para todo $x \in C$ (separación fuerte).*

Demostración : Consideremos el problema $m = \inf[\, \|x\|^2 : x \in C\,]$ ($\|\cdot\|$ es la norma euclidiana). Siendo $x \to \|x\|^2$ fuertemente convexa sci, existe un único $\bar{x} \in C$ tal que $m = \|\bar{x}\|^2 \leq \|x\|^2$ para todo $x \in C$. El hecho de que $0 \notin C$, implica que $\bar{x} \neq 0$ y por lo tanto $m > 0$. De otro lado, para todo $x \in C$ y todo $t \in [0, 1]$, se cumple

$$\|\bar{x}\|^2 = m \leq \langle \bar{x} + t(x - \bar{x}), \bar{x} + t(x - \bar{x}) \rangle = \|\bar{x}\|^2 + 2t\langle \bar{x}, x - \bar{x} \rangle + t^2\|x - \bar{x}\|^2,$$

de donde, $\langle \bar{x}, x - \bar{x} \rangle \geq 0$ para todo $x \in C$. Haciendo $a = \bar{x}$ y $\alpha = \|\bar{x}\|^2$, se demuestra el resultado. ■

Teorema 3.2 *Sea C convexo no vacío tal que $0 \notin C$. Entonces existe $a \in \mathbb{R}^n$ tal que $\langle a, x \rangle \geq 0$ para todo $x \in C$ y $\langle a, x \rangle > 0$ para todo $x \in \mathrm{ri}\,(C)$ (separación propia).*

Demostración : Si $0 \notin \overline{C}$, aplicar el Teorema 3.1 ; en caso contrario, si $0 \in \overline{C}$, se cumple $0 \in \mathrm{aff}\,(C)$. Sea $D = C + [\mathrm{aff}\,(C)]^{\perp}$. Entonces D es convexo, $0 \notin D$ y $\mathrm{int}\,(D) = \mathrm{ri}\,(C) + [\mathrm{aff}\,(C)]^{\perp}$. Fijemos $\bar{x} \in \mathrm{ri}\,(C)$ ($\subset \mathrm{int}\,(D)$). Para todo $k \in \mathbb{N}$, $-\frac{\overline{x}}{k} \notin \overline{D}$ (en caso contrario, $0 \in D$). Aplicando el Teorema 3.1 al conjunto $\overline{D} + \frac{\bar{x}}{k}$, existen $a_k \neq 0$ y α_k tales que

$$0 < \alpha_k \leq \left\langle a_k\,, \left(x + \frac{\bar{x}}{k}\right)\right\rangle \quad \text{para todo } x \in D.$$

Sin perdida de generalidad, podemos suponer $\|a_k\| = 1$. Haciendo $x = \bar{x}$ en la desigualdad precedente, tenemos

$$0 < \alpha_k \leq \left(1 + \frac{1}{k}\right) \langle a_k, \overline{x} \rangle \leq 2\|\overline{x}\|,$$

lo que implica que la sucesión $\{(a_k, \alpha_k)\}$ es acotada. Sea (a, α) un valor de adherencia de esta sucesión. Tenemos, $\|a\| = 1$, $\alpha \geq 0$ y

$$0 \leq \alpha \leq \langle a, x \rangle \quad \text{para todo} \quad x \in D \text{ (y por lo tanto para todo } \ x \in C).$$

Esta última es equivalente a la inclusión $D \subset \{x \in \mathbb{R}^n : \alpha \leq \langle a, x \rangle\}$, lo que implica que $\operatorname{int}(D) \subset \{x \in \mathbb{R}^n : \alpha < \langle a, x \rangle\}$. Se deduce que $\langle a, x \rangle > 0$ para todo $x \in \operatorname{ri}(C)$. ■

Teorema 3.3 (Separación fuerte) *Sean C y D convexos cerrados no vacíos, satisfaciendo*

- **i)** $C \cap D = \emptyset$, *y*
- **ii)** $d \in C_\infty \cap D_\infty$ *implica* $-d \in C_\infty \cap D_\infty$.

Entonces existen $a \in \mathbb{R}^n$ y $\alpha_1, \alpha_2 \in \mathbb{R}$ tales que

$$\langle a, x \rangle \leq \alpha_1 < \alpha_2 \leq \langle a, y \rangle \quad \textit{para todo} \quad (x, y) \in C \times D.$$

Demostración : Sea $S = D - C$. Tenemos $0 \notin S$ y, por el Corolario 1.5, S es cerrado. Luego, por el Teorema 3.1, existen $a \in \mathbb{R}^n$ y $\beta > 0$ tales que

$$\beta \leq \inf_{x \in C, y \in D} \langle a, y - x \rangle = \inf_{y \in D} \langle a, y \rangle - \sup_{x \in C} \langle a, x \rangle.$$

Considerando $\alpha_1 = \sup_{x \in C} \langle a, x \rangle$ y $\alpha_2 = \inf_{y \in D} \langle a, y \rangle$, se obtiene $\alpha_2 - \alpha_1 \geq \beta > 0$. ■

Teorema 3.4 (Separación propia) *Sean C y D convexos no vacíos tales que* $\mathrm{ri}\,(C) \cap \mathrm{ri}\,(D) = \emptyset$. *Entonces, existen $a \in \mathbb{R}^n$ ($a \neq 0$) y $\alpha \in \mathbb{R}$ tales que*

$$\langle a, x\rangle \leq \alpha \leq \langle a, y\rangle \quad \textit{para todo} \quad (x, y) \in \overline{C} \times \overline{D}$$

y

$$\langle a, y - x\rangle > 0 \quad \textit{para todo} \quad (x, y) \in \mathrm{ri}\, C \times \mathrm{ri}\, D.$$

Demostración : Sea $M = \mathrm{ri}\,(D) - \mathrm{ri}\,(C) = \mathrm{ri}\,(M)$. Entonces, M es convexo, con $0 \notin M$, y por lo tanto, por el Teorema 3.2, existe $a \in \mathbb{R}^n$ $(a \neq 0)$ tal que

$$\langle a, x\rangle < \langle a, y\rangle \quad \text{para todo} \quad (x, y) \in \mathrm{ri}\, C \times \mathrm{ri}\, D.$$

Se deduce que

$$\sup_{x \in \overline{C}} \langle a, x\rangle = \sup_{x \in \mathrm{ri}\,(C)} \langle a, x\rangle \leq \inf_{x \in \mathrm{ri}\,(D)} \langle a, y\rangle = \inf_{y \in \overline{D}} \langle a, y\rangle.$$

El teorema se demuestra considerando $\alpha = \frac{1}{2}\left[\sup_{x \in \overline{C}} \langle a, x\rangle + \inf_{y \in \overline{D}} \langle a, y\rangle\right]$. ■

Comentario: En las condiciones del teorema anterior, no siempre existen $a \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ satisfaciendo

$$\langle a, x\rangle < \alpha < \langle a, y\rangle \quad \text{para todo} \quad (x, y) \in \mathrm{ri}\, C \times \mathrm{ri}\, D.$$

Consideremos por ejemplo los conjuntos en $\mathbb{R}^2$,

$$C = \{(x, y) : x > 0,\ y > 0,\ xy \geq 1\} \quad \text{y} \quad D = \{(x, y) : y = 0\}.$$

3.2 Cono polar

Sea $\emptyset \neq K \subset \mathbb{R}^n$. El **cono polar** de K, denotado por K^0, es el conjunto

$$\begin{aligned} K^0 &:= \{x^* \in \mathbb{R}^n : \langle x, x^*\rangle \leq 0 \ \text{ para todo } x \in K\} \\ &= \bigcap_{x \in K} \{x^* \in \mathbb{R}^n : \langle x, x^*\rangle \leq 0\}. \end{aligned}$$

Este es un cono convexo cerrado por ser una intersección de conos convexos cerrados. Se cumplen las siguientes propiedades :

- $K_1 \subset K_2 \;\Rightarrow\; K_2^0 \subset K_1^0$.
- $(\overline{K})^0 = K^0$.
- $(\mathrm{cono}(K))^0 = K^0$.

Existen numerosas propiedades concernientes a los conos polares tanto con la intersección, unión, suma, etc, que serán estudiadas aquí. En general, para cualquier conjunto K, se cumple $K \subset K^{00}$.

Proposición 3.1 *Si K es un cono convexo cerrado y no vacío, entonces $(K^0)^0 = K$.*

Demostración : Por la observación precedente, es suficiente demostrar que $(K^0)^0 \subset K$. Para tal efecto, sea $x \notin K$. Por los teoremas de separación, existen $a \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ tales que $\langle a, x\rangle \leq \alpha < \langle a, y\rangle$ para todo $y \in K$, de donde, $\alpha < 0$ (debido a que $0 \in K$) y $\langle -a, y\rangle \leq 0$ para todo $y \in K$ (debido a que K es un cono). Se deduce que $x \notin K^{00}$ debido a que $-a \in K^0$ y $\langle -a, x\rangle \geq -\alpha > 0$. ■

Corolario 3.1 *K^{00} es el cono convexo cerrado más pequeño conteniendo K.*

Demostración : Sea T un cono convexo cerrado satisfaciendo $K \subset T$. Tenemos, $T^0 \subset K^0$ y por lo tanto, $K^{00} \subset T^{00} = T$. Se deduce que K^{00} es el cono convexo cerrado más pequeño conteniendo K. ■

Si K es un cono convexo, entonces $K^{00} = \overline{K}$.

3.3 Funciones conjugadas

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces, $\mathrm{epi}\,(f) \subset \mathbb{R}^{n+1}$ es convexo cerrado y no vacío. Consideremos

$$K := \{k(x, \lambda, -1) : k \geq 0,\ (x, \lambda) \in \mathrm{epi}\,(f)\} = \bigcup_{k \geq 0} k(\mathrm{epi}\,(f) \times \{-1\}),$$

el cono convexo generado por $\mathrm{epi}\,(f) \times \{-1\} \subset (\mathbb{R}^n \times \mathbb{R}) \times \mathbb{R}$. Este cono no es cerrado porque es distinto con su clausura en el subespacio $(\mathbb{R}^n \times \mathbb{R}) \times \{0\}$.

Se cumple además que

$$\overline{K} \cap (\mathbb{R}^n \times \mathbb{R} \times \{-1\}) = \mathrm{epi}\,(f) \times \{-1\}.$$

Consideremos K^0, el cono polar de K,

$$K^0 = \{(x^*, \lambda^*, \mu^*) : \langle x, x^* \rangle + \lambda\lambda^* - \mu^* \leq 0 \ \text{ para todo } \ (x, \lambda) \in \mathrm{epi}\,(f)\}.$$

Sea $x \in \mathrm{dom}\,(f)$. Haciendo variar λ en $[f(x), +\infty[$, se debe cumplir $\lambda^* \leq 0$ cuando $(x^*, \lambda^*, \mu^*) \in K^0$. Sea

$$T = \{(x^*, \mu^*) : (x^*, -1, \mu^*) \in K^0\}.$$

Este es convexo y cerrado. Además, si $(x_1, \mu_1^*) \in T$ y $\mu_1^* \leq \mu_2^*$, entonces $(x_1, \mu_2^*) \in T$. Por lo tanto, T es el epígrafo de una función convexa sci, denotada por f^*. Se cumple

$$f^*(x^*) = \inf[\mu^* : \langle x, x^* \rangle - \lambda \leq \mu^* \ \text{ para todo } \ (x, \lambda) \in \mathrm{epi}\,(f)],$$

es decir

$$f^*(x^*) = \sup_{x \in \mathrm{dom}\,(f)} [\langle x, x^* \rangle - f(x)]. \tag{3.1}$$

Además, se puede demostrar que

$$K^0 = \{(kx^*, -k, k\mu^*) : k \geq 0, \ (x^*, \mu^*) \in \mathrm{epi}\,(f^*)\}.$$

Debido a que $\overline{K}$ y K^0 se corresponden por dualidad (es decir, $\overline{K} = (K^0)^0$), entonces f y f^* también se corresponden por dualidad (es decir, $f = (f^*)^*$).

Definición 3.1 (Función conjugada) *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. La* **función conjugada** *de f (en el sentido de Fenchel) es la función $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ definida por*

$$f^*(x^*) = \sup_{x \in \mathrm{dom}\,(f)} [\langle x, x^* \rangle - f(x)] = \sup_{x \in \mathbb{R}^n} [\langle x, x^* \rangle - f(x)].$$

Esta es convexa sci por ser el supremo de funciones convexas sci. Además, para cada $x^* \in \mathbb{R}^n$,

$$\langle x, x^* \rangle - f(x) \leq f^*(x^*) \ \text{ para todo } \ x \in \mathbb{R}^n,$$

lo que es equivalente a la desigualdad

$$\langle x, x^* \rangle - f^*(x^*) \leq f(x) \ \text{ para todo } \ x \in \mathbb{R}^n.$$

Esta última desigualdad nos dice que la función lineal afín $x \to \langle x, x^* \rangle - f^*(x^*)$ está por debajo de la función f. Además, $f^*(x^*)$ es el menor valor con esta propiedad, esto es, si

$$\langle x, x^* \rangle - \beta \leq f(x) \ \text{ para todo } \ x \in \mathbb{R}^n,$$

entonces $f^*(x^*) \leq \beta$.

Se cumplen las siguientes propiedades :

- Si existe x_0 tal que $f(x_0) = -\infty$, entonces $f^*(x^*) = +\infty$ para todo $x^* \in \mathbb{R}^n$.
- Si $f(x) = +\infty$ para todo $x \in \mathbb{R}^n$, entonces $f^*(x^*) = -\infty$ para todo $x^* \in \mathbb{R}^n$.
- Si $f(x) \leq g(x)$ para todo $x \in \mathbb{R}^n$, entonces $g^*(x^*) \leq f^*(x^*)$ para todo $x^* \in \mathbb{R}^n$.
- $f^*(x^*) + f(x) \geq \langle x, x^* \rangle$ para todo $(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n$.

Asociada a una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, consideremos las funciones $\overline{f}, f_c, f_{\overline{c}} : \mathbb{R}^n \to \overline{\mathbb{R}}$, definidas por

$$\begin{aligned} \overline{f}(x) &= \inf[\lambda : (x, \lambda) \in \overline{\text{epi}\,(f)}\,]; \\ f_c(x) &= \inf[\lambda : (x, \lambda) \in \text{co}\,(\text{epi}\,(f))]; \\ f_{\overline{c}}(x) &= \inf[\lambda : (x, \lambda) \in \overline{\text{co}}\,(\text{epi}\,(f))]. \end{aligned}$$

Estas son respectivamente la más grande función sci, la más grande función convexa, y la más grande función convexa sci, todas mayoradas por f.

Proposición 3.2 *Se cumple*

$$f^* = (\overline{f})^* = (f_c)^* = (f_{\overline{c}})^*.$$

Demostración : Para $\mu \in \mathbb{R}$, las siguientes relaciones son equivalentes :

- $\langle x, x^*\rangle - \lambda \leq \mu$ para todo $(x, \lambda) \in \mathrm{epi}\,(f)$,
- $\langle x, x^*\rangle - \lambda \leq \mu$ para todo $(x, \lambda) \in \overline{\mathrm{epi}\,(f)}$,
- $\langle x, x^*\rangle - \lambda \leq \mu$ para todo $(x, \lambda) \in \mathrm{co}\,(\mathrm{epi}\,(f))$,
- $\langle x, x^*\rangle - \lambda \leq \mu$ para todo $(x, \lambda) \in \overline{\mathrm{co}}\,(\mathrm{epi}\,(f))$.

El resultado se deduce del hecho que $f^*(x^*) = \sup[\langle x, x^*\rangle - \lambda : (x, \lambda) \in \mathrm{epi}\,(f)]$. ■

Definición 3.2 (Función biconjugada) *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. La* **función biconjugada** *de f (en el sentido de Fenchel), denotada por f^{**}, es la conjugada de la función conjugada, esto es,*

$$f^{**}(x) = \sup_{x^* \in \mathbb{R}^n} [\langle x, x^*\rangle - f^*(x^*)].$$

Se deduce que

$$f^{**}(x) = \sup_{x^*}[\langle x, x^*\rangle - \sup_{y}[\langle y, x^*\rangle - f(y)]] = \sup_{x^*} \inf_{y}[\langle x - y, x^*\rangle + f(y)].$$

En particular,

$$f^{**}(x) \leq \sup_{x^*}[\langle x - x, x^*\rangle + f(x)] = f(x). \tag{3.2}$$

Continuando, podemos definir la tercera conjugada, la cuarta conjugada, etc., de f. Como consecuencia de la siguiente proposición mostramos que la tercera y la primera conjugada coinciden.

Proposición 3.3 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci. Entonces*

$$f^{**}(x) = f(x) \text{ para todo } x \in \mathbb{R}^n.$$

Demostración : Por la relación (3.2), es suficiente mostrar que $f(x) \leq f^{**}(x)$. Sin pérdida de generalidad, podemos suponer que $\mathrm{epi}\,(f) \neq \emptyset$ (en caso contrario, $f(x) = +\infty$ para todo x y por lo tanto, $f^{**}(x) = +\infty$ para todo x). Si $f(x) = -\infty$, entonces, por la desigualdad (3.2), $f^{**}(x) = -\infty$. Caso contrario, si $f(x) > -\infty$, entonces $(x, \lambda) \notin \mathrm{epi}\,(f)$ para $\lambda < f(x)$. Por los teoremas de separación, existe $(\xi^*, \beta) \in \mathbb{R}^n \times \mathbb{R}$, con $\lambda \leq \beta \leq f(x)$, tal que la función lineal $y \to \langle \xi^*, y - x\rangle + \beta = \langle \xi^*, y\rangle - (\langle \xi^*, x\rangle - \beta)$ pasa por debajo de la función f. Luego, por la propiedad de minimalidad de $f^*(\xi^*)$, se cumple $f^*(\xi^*) \leq \langle \xi^*, x\rangle - \beta$, de donde $\beta \leq f^{**}(x)$. Haciendo $\lambda \to f(x)$, tenemos $f(x) \leq f^{**}(x)$. El resultado queda establecido. ■

Corolario 3.2 *Sea* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. *Entonces*

$$f^{***}(x^*) = f^*(x^*) \quad \textit{para todo} \;\; x^* \in \mathbb{R}^n.$$

Demostración : Se deduce de la Proposición 3.3 teniendo en cuenta que f^* es convexa sci. ■

Comentario: De las Proposiciones 3.2 y 3.3,

$$f^{**} = (f_{\overline{c}})^{**} = f_{\overline{c}},$$

de donde,

$$\mathrm{epi}\,(f^{**}) = \overline{\mathrm{co}}\,(\mathrm{epi}\,(f).$$

3.4 Función indicatriz y función soporte de un conjunto

Sea $C \subset \mathbb{R}^n$. La **función indicatriz de** C, denotada por $\delta(\cdot|C)$, es la función definida por

$$\delta(x|C) = \begin{cases} 0 & \text{si } \; x \in C, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Se cumplen las siguientes propiedades :

- C es convexo si y solo si $\delta(\cdot|C)$ es convexa ;
- C es cerrado si y solo si $\delta(\cdot|C)$ es sci ;
- $S \subset C$ si y solo si $\delta(\cdot|S) \geq \delta(\cdot|C)$ para todo $x \in \mathbb{R}^n$;

- $\delta(\cdot|\overline{C})$, $\delta(\cdot|\mathrm{co}\,(C))$ y $\delta(\cdot|\overline{\mathrm{co}}\,C)$ son respectivamente, la más grande función sci, la más grande función convexa, y la más grande función convexa sci, todas mayoradas por $\delta(\cdot|C)$.

Definición 3.3 (Función soporte) *Sea $C \subset \mathbb{R}^n$. La* **función soporte de** *C, denotada por $\delta^*(\cdot|C)$, es la función conjugada de la función indicatriz de C.*

Se cumplen las siguientes propiedades :

- $\delta^*(x^*|C) = \sup_{x\in\mathbb{R}^n}[\langle x, x^*\rangle - \delta(x|C)] = \sup_{x\in C}\langle x, x^*\rangle$;
- Si $C \neq \emptyset$, entonces $\delta^*(\cdot|C)$ es convexa sci y propia ;
- $\delta^*(\cdot|C)$ es positivamente homogénea de grado 1, esto es $\delta^*(kx^*|C) = k\delta^*(x^*|C)$ para todo $k > 0$ y para todo $x^* \in \mathbb{R}^n$.

De hecho, la tercera propiedad es una caracterización de la función soporte.

Teorema 3.5 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia positivamente homogénea de grado 1. Entonces, existe C convexo cerrado no vacío tal que $f(x^*) = \delta^*(x^*|C)$ para todo $x^* \in \mathbb{R}^n$. Además, $f^*(x) = \delta(x|C)$ para todo $x \in \mathbb{R}^n$.*

Demostración : a) Se cumple

$$\begin{aligned} f^*(x) &= \sup_{x^*}[\langle x, x^*\rangle - f(x^*)] \\ &= \sup_{x^*,k>0}[\langle x, kx^*\rangle - f(kx^*)] \\ &= \sup_{k>0}\left[k\left[\sup_{x^*}[\langle x, x^*\rangle - f(x^*)]\right]\right]. \end{aligned}$$

Puesto que f es propia, $\sup_{x^*}[\langle x, x^*\rangle - f(x^*)] > -\infty$. Por lo tanto

$$f^*(x) = \begin{cases} 0 & \text{si } \sup_{x^*}[\langle x, x^*\rangle - f(x^*)] \leq 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Sea $C = S_0(f^*)$. Este es convexo cerrado no vacío y $f^*(x) = \delta(x|C)$ para todo $x \in \mathbb{R}^n$.

b) Debido a que f es convexa sci propia, $f = f^{**}$ y por lo tanto, $f(x^*) = \delta^*(x^*|C)$ para todo $x^* \in \mathbb{R}^n$. ■

Definición 3.4 (Cono barrera) *Sea $C \subset \mathbb{R}^n$. El* **cono barrera de** C *es el conjunto K definido por*

$$K = \{x^* : \delta^*(x^*|C) < +\infty\} = \mathrm{dom}\,(\delta^*(\cdot|C)).$$

Se deduce que K es un cono convexo debido a que este es el dominio de una función convexa positivamente homogénea. En general no se cumple que K sea cerrado aún cuando C es convexo cerrado. Considere por ejemplo $C = \{(x,y) \in \mathbb{R}^2 : y + e^x \leq 0\}$. Entonces, $K = \mathbb{R}^2_{++} \cup (\{0\} \times \mathbb{R}_+)$.

Note que si K es el cono barrera de C, entonces K también es el cono barrera de $\mathrm{co}\,(C)$, $\overline{C}$ y $\overline{\mathrm{co}}\,(C)$.

Teorema 3.6 *Sean $C \subset \mathbb{R}^n$ convexo cerrado no vacío y K su cono barrera. Entonces $K \subset (C_\infty)^0$ y $K^0 = C_\infty$. Se deduce que $\overline{K} = (C_\infty)^0$.*

Demostración : Sea $a \in C$. i) Consideremos $x^* \in K$. Entonces $\delta^*(x^*|C) < +\infty$. Si $d \in C_\infty$, entonces $\langle a+\lambda d, x^*\rangle \leq \delta^*(x^*|C)$ para todo $\lambda > 0$. Se deduce que $\langle d, x^*\rangle \leq 0$ para todo $d \in C_\infty$, de donde $x^* \in (C_\infty)^0$.

ii) Por dualidad y por el hecho que C_∞ es un cono convexo cerrado no vacío, $K^0 \supset (C_\infty)^{00} = C_\infty$. Para la inclusión contraria, sea $d \notin C_\infty$. Entonces existe $\lambda > 0$ tal que $a + \lambda d \notin C$, de donde, por los teoremas de separación, existe $x^* \neq 0$ tal que

$$\langle a + \lambda d, x^*\rangle > \langle x, x^*\rangle \ \text{ para todo } \ x \in C.$$

Por lo tanto, $\langle d, x^*\rangle > 0$ (haciendo $x = a$) y $\sup[\langle x, x^*\rangle : x \in C] \leq \langle a + \lambda d, x^*\rangle < \infty$. Se deduce que $d \notin K^0$. ■

Corolario 3.3 *Sea $C \subset \mathbb{R}^n$ no vacío. Entonces C es acotado si y solo si $K = \mathbb{R}^n$.*

Demostración : C es acotado si y solo si $\overline{\mathrm{co}}\,(C)$ es acotado si y solo si $(\overline{\mathrm{co}}\,(C))_\infty = \{0\}$ si y solo si $\overline{K} = \mathbb{R}^n$ si y solo si $K = \mathbb{R}^n$ (porque K es convexo). ■

3.5 Relaciones entre la función asintótica y la función conjugada

Proposición 3.4 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces* $f_\infty(d) = \delta^*\left(d \,|\, \overline{\mathrm{dom}\,(f^*)}\right) = \delta^*(d|\mathrm{dom}\,(f^*))$ *para todo* $d \in \mathbb{R}^n$.

Demostración : Sea $a \in \mathrm{dom}\,(f)$. De $f_\infty(d) = \sup_{t>0} \frac{f(a+td)-f(a)}{t}$ y

$$f(a+td) = f^{**}(a+td) = \sup_{x^* \in \mathrm{dom}\,(f^*)} [\langle x^*, a+td\rangle - f^*(x^*)],$$

tenemos

$$\begin{aligned} f_\infty(d) &= \sup_{t>0,\, x^* \in \mathrm{dom}\,(f^*)} \frac{1}{t}[\langle x^*, a+td\rangle - f^*(x^*) - f(a)] \\ &= \sup_{x^* \in \mathrm{dom}\,(f^*)} \left[\langle x^*, d\rangle + \sup_{t>0} \frac{1}{t}[\langle x^*, a\rangle - f^*(x^*) - f(a)]\right], \end{aligned}$$

de donde, $f_\infty(d) = \delta^*(d|\mathrm{dom}\,(f^*))$, debido a que $\langle x^*, a\rangle \leq f^*(x^*) + f(a)$. ■

Corolario 3.4 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces*

i) *f es inf-compacto si y solo si* $0 \in \mathrm{int}\,(\mathrm{dom}\,(f^*))$ *;*

ii) *f^* es inf-compacto si y solo si* $0 \in \mathrm{int}\,(\mathrm{dom}\,(f))$.

Demostración : i) f es inf-compacto si y solo si

$$\delta^*(d|\mathrm{dom}\,(f^*)) = f_\infty(d) = \begin{cases} 0 & \text{si } d = 0, \\ > 0 & \text{en otro caso.} \end{cases}$$

Este es equivalente a

$$0 \in \mathrm{int}\,(\mathrm{dom}\,(f^*)).$$

ii) Se obtiene de i) por dualidad. ■

Ejemplo 3.1 *Sea* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *convexa sci propia. Consideremos* $\psi : \mathbb{R}^n \times \mathbb{R} \to \overline{\mathbb{R}}$ *definida por*

$$\psi(x^*, \lambda^*) = \begin{cases} 0 & \text{si } f^*(x^*) + \lambda^* \leq 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Esta es convexa sci propia (es la función indicatriz de un conjunto convexo cerrado no vacío). Su función conjugada es

$$\psi^*(x, \lambda) = \sup_{x^*, \lambda^*} [\langle x, x^* \rangle + \lambda\lambda^* : f^*(x^*) + \lambda^* \leq 0].$$

Tenemos :

- *Si* $\lambda < 0$, *entonces* $\psi^*(x, \lambda) = \infty$ *(tomar* $x^* \in \operatorname{dom}(f^*)$ *y hacer* $\lambda^* \to -\infty$*),*
- *Si* $\lambda = 0$, *entonces* $\psi^*(x, \lambda) = \sup[\langle x, x^* \rangle : x^* \in \operatorname{dom}(f^*)] = f_\infty(x)$,
- *Si* $\lambda > 0$, *entonces*

$$\begin{aligned} \psi^*(x, \lambda) &= \sup_{x^*} [\langle x, x^* \rangle - \lambda f^*(x^*)] \\ &= \lambda \sup_{x^*} \left[\left\langle \frac{x}{\lambda}, x^* \right\rangle - f^*(x^*) \right] \\ &= \lambda f(\frac{x}{\lambda}). \end{aligned}$$

Denotando $\varphi = \psi^*$, *lo anterior se resume en*

$$\varphi(x, \lambda) = \begin{cases} \lambda f(\frac{x}{\lambda}) & \text{si } \lambda > 0, \\ \\ f_\infty(x) & \text{si } \lambda = 0, \\ \\ +\infty & \text{si } \lambda < 0. \end{cases}$$

Así, la conjugada de φ *es*

$$\varphi^*(x^*, \lambda^*) = \psi(x^*, \lambda) = \begin{cases} 0 & \text{si } f^*(x^*) + \lambda^* \leq 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

3.6 Subdiferencial de una función convexa

Sean $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa propia y $a \in \mathrm{dom}\,(f)$. El subdiferencial de f en a, denotado por $\partial f(a)$, es el conjunto definido por

$$\partial f(a) = \{x^* \in \mathbb{R}^n : f(a) + f^*(x^*) = \langle x^*, a\rangle\}. \tag{3.3}$$

Debido a que $f^*(x^*) = \sup_x[\langle x^*, x\rangle - f(x)]$, tenemos

$$\begin{aligned} \partial f(a) &= \{x^* \in \mathbb{R}^n : f(a) + f^*(x^*) \leq \langle x^*, a\rangle\} \\ &= \{x^* \in \mathbb{R}^n : f(a) + \langle x^*, x - a\rangle \leq f(x) \ \text{ para todo } x \in \mathbb{R}^n\}. \end{aligned}$$

Por extensión, definimos $\partial f(a) = \emptyset$ si $a \notin \mathrm{dom}\,(f)$. Aún cuando $a \in \mathrm{dom}\,(f)$, el subdiferencial $\partial f(a)$ puede ser vacío. Consideremos por ejemplo la función $f : \mathbb{R} \to \overline{\mathbb{R}}$ definida por $f(x) = -\sqrt{x}$, si $x \geq 0$; $+\infty$, en caso contrario. Se comprueba fácilmente que $\partial f(0) = \emptyset$.

Se deduce inmediatamente de la definición que

$$0 \in \partial f(a) \ \text{ si y solo si } \ f(a) = \inf_{x\in\mathbb{R}^n} f(x).$$

Para dar una interpretación geométrica del subdiferencial $\partial f(a)$, consideremos el hiperplano en $\mathbb{R}^n \times \mathbb{R}$,

$$H = \{(x, \lambda) : f(a) + \langle x^*, x - a\rangle = \lambda\}.$$

Este está por debajo del epígrafo de f, y coincide con este al menos en el punto $(a, f(a))$.

Proposición 3.5 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa propia y $a \in \mathrm{dom}\,(f)$. Entonces $\partial f(a)$ es convexo y cerrado.*

Demostración : Se deduce del hecho que la aplicación $x^* \to f^*(x^*) - \langle a, x^*\rangle$ es convexa y sci. ■

Proposición 3.6 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa propia. Si $\partial f(a) \neq \emptyset$, entonces f es sci en a.*

Demostración : Sean $\lambda < f(a)$ y $x^* \in \partial f(a)$. Tenemos $h_a(x) := f(a) + \langle x^*, x - a\rangle \leq f(x)$ para todo $x \in \mathbb{R}^n$, de donde, por la continuidad de h_a y el hecho que $\lambda < f(a) = h_a(a)$, existe una vecindad V de a tal que $\lambda < h_a(x) \leq f(x)$ para todo $x \in V$. Se sigue que f es sci en a. ■

Comentario: La recíproca de la proposición anterior es falsa. Consideremos por ejemplo $f : \mathbb{R} \to]-\infty, +\infty]$ definida por

$$f(x) = \begin{cases} +\infty & \text{si } x < 0, \\ -\sqrt{x} & \text{si } x \geq 0. \end{cases}$$

Esta es convexa sci y

$$f^*(x^*) = \sup_{x \geq 0}[xx^* + \sqrt{x}] = \begin{cases} +\infty & \text{si } x^* \geq 0, \\ -\frac{1}{4x^*} & \text{si } x^* < 0. \end{cases}$$

Se deduce que

$$x^* \in \partial f(x) \iff \begin{cases} -\sqrt{x} - \frac{1}{4x^*} = x^* x & \text{si } x \geq 0 \text{ y } x^* < 0, \\ \\ \emptyset & \text{en caso contrario,} \end{cases}$$

de donde

$$\partial f(x) = \begin{cases} \{-\frac{1}{2\sqrt{x}}\} & \text{si } x > 0, \\ \\ \emptyset & \text{en caso contrario.} \end{cases}$$

Proposición 3.7 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces $x^* \in \partial f(x)$ si y solo si $x \in \partial f^*(x^*)$.*

Demostración : Se deduce del hecho que $f = f^{**}$. ■

Teorema 3.7 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces $\partial f(a) \neq \emptyset$ para todo $a \in \mathrm{ri}\,(\mathrm{dom}\,(f))$.*

Demostración : Sea $C = \{(a, f(a))\}$. Entonces $C = \mathrm{ri}\,(C)$ y $C \cap \mathrm{ri}\,(\mathrm{epi}\,(f)) = \emptyset$. Por los teoremas de separación, existe (x^*, α^*) tal que

$$\langle a, x^* \rangle + \alpha^* f(a) > \langle x, x^* \rangle + \alpha^* \lambda \quad \text{para todo} \quad (x, \lambda) \in \mathrm{ri}\,(\mathrm{epi}\,(f)). \tag{3.4}$$

Se deduce que $\alpha^* < 0$, el cual, sin pérdida de generalidad, podemos asumir $\alpha^* = -1$. Haciendo $\lambda \to f(x)$, obtenemos

$$\langle a, x^* \rangle - f(a) \geq \langle x, x^* \rangle - f(x) \quad \text{para todo } x \in \mathrm{dom}\,(f),$$

lo que implica que $x^* \in \partial f(a)$. ∎

3.7 Derivadas direccionales de una función convexa

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, finita en $a \in \mathbb{R}^n$. La **derivada direccional** de f en a en la dirección $d \in \mathbb{R}^n$, denotada por $f'(a, d)$, es el límite, si este existe,

$$f'(a, d) = \lim_{t \to 0^+} \frac{f(a + td) - f(a)}{t}.$$

Se sabe que si f es convexa, la función $t \to \frac{f(a+td)-f(a)}{t}$ es creciente en $]0, +\infty[$ y por lo tanto $\lim_{t \to 0^+} \frac{f(a+td)-f(a)}{t}$ existe (con la eventualidad de ser $\pm\infty$) y coincide con $\inf_{t>0} \frac{f(a+td)-f(a)}{t}$.

Teorema 3.8 *Sean $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa propia y $a \in \mathrm{dom}\,(f)$. Entonces $f'(a, \cdot)$ es convexa positivamente homogénea.*

Demostración : Tenemos $f'(a, kd) = kf'(a, d)$ para todo $k > 0$. De otro lado, debido a que la función $d \to g(d) = f(a + d) - f(a)$ es convexa, la función $(d, \theta) \to \varphi(d, \theta) = \theta g(\frac{d}{\theta})$ es convexa en $\mathbb{R}^n \times]0, +\infty[$. Se deduce que la función $d \to f'(a, d)$ es convexa en $\mathbb{R}^n$ debido a que $f'(a, d) = \inf_{\theta > 0} \varphi(d, \theta)$ (Proposición 2.6). ∎

Proposición 3.8 *Sean* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *convexa propia y* $a \in \operatorname{dom}(f)$. *Entonces*

$$x^* \in \partial f(a) \quad \textit{si y solo si} \quad f'(a,d) \geq \langle x^*, d\rangle \quad \textit{para todo} \;\; d \in \mathbb{R}^n.$$

Demostración : Sea $x^* \in \partial f(a)$. Entonces $f(a+td) - f(a) \geq t\langle x^*, d\rangle$ para todo $d \in \mathbb{R}^n$ y todo $t > 0$. Dividiendo por t y luego haciendo $t \to 0$, obtenemos $f'(a,d) \geq \langle x^*, d\rangle$. Recíprocamente, asumamos que $f'(a,d) \geq \langle x^*, d\rangle$ para todo $d \in \mathbb{R}^n$. Para $d = x - a$, obtenemos

$$\langle x^*, x-a\rangle \leq f'(a,d) \leq \frac{f(a+td) - f(a)}{t} \quad \text{para todo } t > 0,$$

y por lo tanto, para $t = 1$, $x^* \in \partial f(a)$. ∎

Proposición 3.9 *Sean* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *convexa propia y* $a \in \operatorname{dom}(f)$.

a) *Si existe* d *tal que* $f'(a,d) = -\infty$, *entonces* $\partial f(a) = \emptyset$.

b) *Si* $f'(a,d) > -\infty$ *para todo* $d \in \mathbb{R}^n$, *entonces* $\overline{\theta}(d) = \delta^*(d|\partial f(a))$, *donde* $\theta(\cdot) = f'(a,\cdot)$. *Se deduce que* $\partial f(a) \neq \emptyset$.

Demostración : a) Por el teorema precedente, $\langle x^*, d\rangle \leq f'(a,d)$ desde que $x^* \in \partial f(a)$. Se deduce que $\partial f(a) = \emptyset$, si $f'(a,d) = -\infty$.

b) Por definición $\overline{\theta}$ es sci. Este también es convexa propia y positivamente homogénea debido a que θ lo es. Por el Teorema 3.5, $\overline{\theta}$ es la función soporte de un conjunto convexo cerrado no vacío, D.

Para $x^* \in \mathbb{R}^n$, consideremos la función μ_{x^*} definida por $\mu_{x^*}(d) = \theta(d) - \langle x^*, d\rangle$. Tenemos $\overline{\mu}_{x^*}(d) = \overline{\theta}(d) - \langle x^*, d\rangle$ y por lo tanto

$$\begin{aligned}
x^* \in \partial f(a) \quad &\Longleftrightarrow \quad \mu_{x^*}(d) \geq 0 \;\; \text{para todo} \;\; d \in \mathbb{R}^n \\
&\Longleftrightarrow \quad \overline{\mu}_{x^*}(d) \geq 0 \quad \text{para todo} \;\; d \in \mathbb{R}^n \\
&\Longleftrightarrow \quad \langle x^*, d\rangle \leq \delta^*(d|D) \quad \text{para todo} \;\; d \in \mathbb{R}^n \\
&\Longleftrightarrow \quad \sup_d [\langle x^*, d\rangle - \delta^*(d|D)] \leq 0 \\
&\Longleftrightarrow \quad \delta(x^*|D) \leq 0 \\
&\Longleftrightarrow \quad x^* \in D.
\end{aligned}$$

Se deduce que $\partial f(a) \neq \emptyset$ y $\overline{\theta}(d) = \delta^*(d|\partial f(a))$. ■

Comentario: En el siguiente ejemplo, f es una función convexa propia, y sci en $a = (0,0)$, con $\partial f(a) \neq \emptyset$; pero la función θ correspondiente no es sci :

$$f(x,y) = \begin{cases} y & \text{si } y < 0, \\ 0 & \text{si } x = y = 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Tenemos,

$$\begin{aligned} f^*(x^*,y^*) &= \begin{cases} 0 & \text{si } (x^*,y^*) \in \{0\} \times [1,+\infty[\,, \\ +\infty & \text{en caso contrario.} \end{cases} \\ \partial f(0,0) &= \{0\} \times [1,\infty[\,; \\ \theta(1,0) &= f'((0,0),(1,0)) = +\infty. \end{aligned}$$

De otro lado,

$$\overline{\theta}(1,0) = \sup_{x^*,y^*}[1x^* + 0y^* : (x^*,y^*) \in \partial f(0,0)] = 0.$$

Proposición 3.10 *Sean f convexa propia y $a \in \operatorname{dom}(f)$. Entonces $\partial f(a)$ es compacto no vacío si y solo si $a \in \operatorname{int}(\operatorname{dom}(f))$. En tal caso,*

$$f'(a,\cdot) = \delta^*(\cdot|\partial f(a)).$$

Demostración : a) Sea $a \in \operatorname{int}(\operatorname{dom}(f))$. Para $0 \neq d \in \mathbb{R}^n$, existe $\bar{t} > 0$ tal que $a + \bar{t}d \in \operatorname{dom}(f)$, de donde

$$\theta(d) = f'(a,d) = \inf_{t>0} \frac{f(a+td) - f(a)}{t} \leq \frac{f(a+\bar{t}d) - f(a)}{\bar{t}} < \infty.$$

De otro lado, la convexidad de θ implica $0 = \theta(0) \leq \frac{1}{2}\theta(d) + \frac{1}{2}\theta(-d) < +\infty$ y por consiguiente, $\theta(d) > -\infty$. Así, $\operatorname{dom}(\theta) = \mathbb{R}^n$ y por lo tanto $\operatorname{int}(\operatorname{dom}(\theta)) = \mathbb{R}^n$. Por las Proposiciones 2.10 y 3.9, tenemos $f'(a,d) = \theta(d) = \overline{\theta}(d) = \delta^*(d|\partial f(a)) < +\infty$ para todo $d \in \mathbb{R}^n$, lo que implica que $\partial f(a)$ es acotado (y por lo tanto compacto).

b) Asumamos $\partial f(a)$ compacto. Sean $e_1, \cdots, e_n$ los vectores de la base canónica. Debido a que $\mathrm{dom}\,(\overline{\theta}) = \mathbb{R}^n$, se tiene $\mathrm{dom}\,(\theta) = \mathbb{R}^n$ y por consiguiente $\theta = \overline{\theta}$. Debido a que $f'(a, \pm e_i)$ es finito, existe $t_i > 0$ tal que $f(a \pm t_i e_i) < +\infty$, y por consiguiente, $a \pm t_i e_i \in \mathrm{dom}\,(f)$. La envoltura convexa de los $2n$ puntos $a \pm t_i e_i$ es una vecindad de a contenida en $\mathrm{dom}\,(f)$. ■

Corolario 3.5 *Sean f convexa propia y $a \in \mathrm{ri}\,(\mathrm{dom}\,(f))$. Entonces $\partial f(a) \neq \emptyset$ y $f'(a, d) = \delta^*(d|\partial f(a))$.*

Demostración : Sean $\mathbb{R}^n = [\mathrm{aff}\,(\mathrm{dom}\,(f))] \times [\mathrm{aff}\,(\mathrm{dom}\,(f))]^{\perp}$ y $\tilde{f}$ la restricción de f sobre $\mathrm{aff}\,(\mathrm{dom}\,(f))$. Tenemos

$$f(x, y) = \begin{cases} \widetilde{f}(x) & \text{si } y = 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Por consiguiente, $\partial f(a_1, a_2) = \partial \widetilde{f}(a_1) \times [\mathrm{aff}\,(\mathrm{dom}\,(f))]^{\perp}$. ■

3.8 La derivada de una función convexa

Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, finita en $a \in \mathbb{R}^n$, es **Fréchet-diferenciable** en a, si existe $x^* \in \mathbb{R}^n$ tal que

$$\frac{f(a+h) - f(a) - \langle h, x^* \rangle}{\|h\|} \to 0, \qquad \text{cuando } h \to 0\,.$$

Observe que la definición no depende de la norma elegida en $\mathbb{R}^n$ y que el elemento x^*, si existe, es único. A este elemento se le suele denotar por $\nabla f(a)$, llamada la **derivada de Fréchet** de f en a. Se cumple,

$$\langle \nabla f(a), h \rangle = f'(a, h) \ \text{ para todo } h \in \mathbb{R}^n, \ \text{ si } \nabla f(a) \text{ existe}$$

Teorema 3.9 *Sean f convexa propia y $a \in dom(f)$. Entonces f es Fréchet diferenciable en a si y solo si $\partial f(a)$ es un conjunto unitario. En este caso, $a \in \mathrm{int}\,(\mathrm{dom}\,(f))$ y $\partial f(a) = \{\nabla f(a)\}$.*

Demostración : i) Asumamos f Fréchet-diferenciable en a. Tenemos, $f'(a,h) = \langle \nabla f(a), h\rangle$ para todo $h \in \mathbb{R}^n$ y por lo tanto, para $x^* \in \partial f(a)$, se cumple $\langle x^*, h\rangle \leq f'(a,h) = \langle \nabla f(a), h\rangle$ para todo h, lo que implica que $x^* = \nabla f(a)$.

ii) Asumamos ahora que $\partial f(a) = \{x^*\}$. Tenemos $a \in \operatorname{int}(\operatorname{dom}(f))$ por ser $\partial f(a)$ compacto no vacío. Sea $g : \mathbb{R}^n \to \overline{\mathbb{R}}$, definida por $g(x) = f(x) - f(a) - \langle x - a, x^*\rangle$. Esta es convexa, $g(a) = 0 \leq g(x)$ para todo $x \in \mathbb{R}^n$, y

$$g'(a,h) = 0 \text{ para todo } h \in \mathbb{R}^n.$$

Sean $\{e_1, \cdots, e_n\}$, la base canónica de $\mathbb{R}^n$, y

$$d_i := e_i, \quad d_{n+i} := -e_i \text{ para todo } i = 1, \cdots, n.$$

Para $h \in \mathbb{R}^n$, consideremos $\xi_i = 0$ y $\xi_{n+i} = -h_i$, si $h_i < 0$; y $\xi_i = h_i$ y $\xi_{n+i} = 0$, si $h_i \geq 0$. Se cumple $h = \sum_{i=1}^{2n} \xi_i d_i$. Denotemos $\|h\|_1 = \sum_{i=1}^{n} |h_i|$ y $t_i = \frac{\xi_i}{\|h\|_1}$ para $i = 1, \cdots, 2n$. Se cumple, $t_i \geq 0$ para todo i y $\sum t_i = 1$. Por la convexidad de g, se tiene

$$g(a+h) = g(a + \sum_{i=1}^{2n} t_i \|h\|_1 d_i) \leq \sum_{i=1}^{2n} t_i g(a + \|h\|_1 d_i),$$

de donde,

$$0 \leq \frac{g(a+h) - g(a)}{\|h\|_1} \leq \sum_{i=1}^{2n} t_i \frac{g(a + \|h\|_1 d_i) - g(a)}{\|h\|_1}.$$

Por lo tanto, debido a que $0 = g'(a, d_i) = \lim_{h\to 0} \frac{g(a+\|h\|_1 d_i) - g(a)}{\|h\|_1}$,

$$\lim_{h\to 0} \frac{g(a+h) - g(a)}{\|h\|_1} = 0.$$

Se deduce que $\nabla g(a) = 0$, y de este, $x^* = \nabla f(a)$. ■

Proposición 3.11 *Sea f convexa sci propia. Si f es estrictamente convexa, entonces* $\operatorname{int}(\operatorname{dom}(f^*)) \neq \emptyset$, *$f^*$ es diferenciable en todo* $\operatorname{int}(\operatorname{dom}(f^*))$ *y* $\partial f^*(x^*) = \emptyset$ *si* $x^* \notin \operatorname{int}(\operatorname{dom}(f^*))$.

Demostración : Sea $x \in \partial f^*(x^*)$. Entonces x es una solución óptima del problema $\inf_x[f(x) - \langle x, x^*\rangle]$. Siendo f estrictamente convexa, el problema admite a lo más una solución óptima. Por lo tanto, $\partial f^*(x^*)$ es un conjunto unitario desde que este es no vacío. De otro lado, como f es convexa propia sci, entonces f^* también lo es y por lo tanto, $\mathrm{ri}\,(\mathrm{dom}\,(f^*)) \neq \emptyset$. Debido a $\partial f^*(x^*) \neq \emptyset$ si $x^* \in \mathrm{ri}\,(\mathrm{dom}\,(f^*))$, se cumple, por el argumento anterior y por el Teorema 3.9, que f^* es diferenciable en $\mathrm{ri}\,(\mathrm{dom}\,(f^*))$, deduciendose que $\mathrm{int}\,(\mathrm{dom}\,(f^*)) = \mathrm{ri}\,(\mathrm{dom}\,(f)) \neq \emptyset$.

El argumento anterior muestra que $\partial f^*(x^*) = \emptyset$ para todo x^* fuera de $\mathrm{int}\,(\mathrm{dom}\,(f^*))$. ■

Proposición 3.12 *Sea f convexa sci propia. Si su conjugada f^* es diferenciable en todo x^* donde $\partial f^*(x^*) \neq \emptyset$, entonces f es estrictamente convexa en* $\mathrm{ri}\,(\mathrm{dom}\,(f))$.

Demostración : Asumamos, por contradicción, que f no es estrictamente convexa en $\mathrm{ri}\,(\mathrm{dom}\,(f))$. Entonces, existen $x, y \in \mathrm{ri}\,(\mathrm{dom}\,(f))$ $(x \neq y)$ y $t \in]0, 1[$, tales que

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y) \tag{3.5}$$

Debido a que $tx + (1-t)y \in \mathrm{ri}\,(\mathrm{dom}\,(f))$, existe $x^* \in \partial f(tx + (1-t)y)$ y por lo tanto

$$f(tx + (1-t)y) + \langle x - tx - (1-t)y, x^*\rangle \leq f(x) \tag{3.6}$$

$$f(tx + (1-t)y) + \langle y - tx - (1-t)y, x^*\rangle \leq f(y) \tag{3.7}$$

Multiplicando la desigualdad (3.6) por t, y la desigualdad (3.7) por $(1-t)$, la suma de estas dos expresiones resulta

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

Por lo tanto, de la relación (3.5), se deduce que las desigualdades en (3.6) y (3.7) son en realidad igualdades. Por otro lado

$$f(tx + (1-t)y) + f^*(x^*) = \langle tx + (1-t)y, x^*\rangle \tag{3.8}$$

Reemplazando en (3.6) y (3.7), resultan

$$f^*(x^*) + f(x) = \langle x, x^* \rangle \quad \text{y} \quad f^*(x^*) + f(y) = \langle y, x^* \rangle$$

y por lo tanto $x, y \in \partial f^*(x^*)$. Este es una contradicción porque por hipótesis $\partial f^*(x^*)$ es vacío o se reduce a un sólo elemento. ∎

3.9 Subdiferencial de la suma de dos funciones convexas

Para finalizar este capítulo, veamos las relaciones que hay entre el subdiferencial de la suma de dos funciones convexas con la suma de los subdiferenciales de cada una de estas funciones. En esta parte solamente enunciaremos estas relaciones, las demostraciones serán dadas en el siguiente capítulo.

Teorema 3.10 *Sean f y g convexas sci propias, $s = f + g$ y $a \in \mathrm{dom}\,(s) = [\mathrm{dom}\,(f)] \cap [\mathrm{dom}\,(g)]$ tales que $\partial f(a) \neq \emptyset$ y $\partial g(a) \neq \emptyset$. Entonces*

a) $\partial f(a) + \partial g(a) \subset \partial s(a)$.

b) *Si* $\mathrm{int}\,(\mathrm{dom}\,(f)) \cap \mathrm{int}\,(\mathrm{dom}\,(g)) \neq \emptyset$, *entonces* $\partial f(a) + \partial g(a) = \partial s(a)$.

Capítulo 4

Introducción a la Dualidad Convexa

4.1 Esquema General

Los problemas de optimización matemática se presentan en general en el contexto siguiente :

$$\alpha = \inf[\tilde{f}(x) : x \in C],$$

donde $C \subset \mathbb{R}^n$ y $\tilde{f} : C \to \mathbb{R}$, son dados. Considerando $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ definida por

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

el problema anterior es equivalente al problema, aparentemente sin restricción,

$$\alpha = \inf[f(x) : x \in \mathbb{R}^n]. \qquad (P)$$

A este problema le llamaremos **problema primal**.

Asociada a la función f, consideremos la función $\varphi : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$, llamada **función de perturbación**, tal que

$$\varphi(x, 0) = f(x) \text{ para todo } x \in \mathbb{R}^n,$$

y con esta, la función $h : \mathbb{R}^p \to \overline{\mathbb{R}}$ definida por

$$h(u) := \inf_{x\in\mathbb{R}^n} \varphi(x,u). \tag{P_u}$$

El problema (P_u) es el **problema primal perturbado** con perturbación u.

Se cumple

$$\sup_{u^*}[\langle 0,u^*\rangle - h^*(u^*) : u^* \in \mathbb{R}^n] = h^{**}(0) \leq h(0) = \inf_{x\in\mathbb{R}^n} \varphi(x,0) = \inf_{x\in\mathbb{R}^n} f(x) = \alpha$$

Denotando $-\beta = h^{**}(0)$, se obtiene el siguiente problema

$$\beta = \inf_{u^*\in\mathbb{R}^n} h^*(u^*), \tag{Q}$$

llamado el **problema dual** del problema (P).

En general, se cumple $-\beta \leq \alpha$. A la suma $\alpha+\beta$ se le llama **salto de dualidad**. Obviamente el salto de dualidad es cero (o no existe salto de dualidad) si $h^{**}(0) = h(0)$.

Calculemos la conjugada de h :

$$\begin{aligned} h^*(u^*) &= \sup_{u\in\mathbb{R}^n} [\langle u,u^*\rangle - h(u)] \\ &= \sup_{u\in\mathbb{R}^n} [\langle u,u^*\rangle - \inf_{x\in\mathbb{R}^n} \varphi(x,u)] \\ &= \sup_{(x,u)\in\mathbb{R}^n\times\mathbb{R}^n} [\langle 0,x\rangle + \langle u,u^*\rangle - \varphi(x,u)] \\ &= \varphi^*(0,u^*). \end{aligned}$$

El **problema dual perturbado** con perturbación x^*, es

$$k(x^*) := \inf_{u^*\in\mathbb{R}^p} \varphi^*(x^*,u^*). \tag{Q_{x^*}}$$

Se cumple $\beta = k(0)$.

Resumiendo, se cumple la siguiente simetría entre los problemas primal y dual :

$$\begin{array}{rclcrcl} \alpha &=& \inf_x f(x) \quad (P) & & \beta &=& \inf_{u^*} h^*(u^*) \quad (Q) \\ \varphi(x,0) &=& f(x) \ \ \forall\, x\in\mathbb{R}^n & & \varphi^*(0,u^*) &=& h^*(u^*) \ \ \forall\, u^*\in\mathbb{R}^p \\ h(u) &=& \inf_x \varphi(x,u) \quad (P_u) & & k(x^*) &=& \inf_{u^*} \varphi^*(x^*,u^*) \quad (Q_{x^*}) \\ \alpha &=& h(0) & & \beta &=& k(0) \end{array}$$

$$-\beta \leq \alpha.$$

El caso particular, $\varphi = \varphi^{**}$, el problema (P) también es el dual de (Q) y $f^* = k$.

Por construcción, h^* es convexa sci y por consiguiente, (Q) es un problema de optimización convexo (aún si (P) no lo fuera).

Si φ es convexa, entonces h es convexa. Si h es convexa y propia, entonces una condición necesaria y suficiente para que no exista salto de dualidad es que h sea sci en 0 (en general, φ sci no implica que h lo sea).

Referente a las soluciones óptimas de los problemas (P) y (Q), tenemos :

$$\begin{array}{ccc} \bar{u}^* \text{ es solución óptima de (Q)} & \iff & h^*(\bar{u}^*) \leq h^*(u^*) \;\; \text{para todo } u^* \in \mathbb{R}^p \\ & & \Updownarrow \\ \bar{u}^* \in \partial h^{**}(0) & \iff & 0 \in \partial h^*(\bar{u}^*). \end{array}$$

Por consiguiente, el conjunto de las soluciones óptimas del problema (Q) es el subdiferencial de h^{**} en 0.

Tenemos el siguiente resultado

Proposición 4.1 *Asumamos que h es convexa y propia.*

a) Si $0 \in \mathrm{ri}\,(\mathrm{dom}\,(h))$, entonces $\alpha = -\beta$ y el conjunto de soluciones óptimas del problema (Q) es no vacío.

b) Si $0 \in \mathrm{int}\,(\mathrm{dom}\,(h))$, el conjunto de soluciones óptimas de (Q) también es compacto.

Ahora, si φ es convexa sci propia, entonces el esquema de dualidad es simétrico : $f = k^*$ y $\varphi^{**} = \varphi$. Por lo tanto, el conjunto de las soluciones óptimas de (P) coincide con el subdiferencial de k^{**} en 0. Por lo tanto, se cumple el análogo de la Proposición 4.1.

Relacionemos ahora las soluciones óptimas de de los problemas (P) y (Q) por medio del subdiferencial de la función de perturbación φ. Sea $\bar{u}^*$

una solución óptima del problema (Q) y asumamos $\varphi^{**} = \varphi$ y $\alpha = -\beta$. Tenemos :

$$\bar{x} \text{ es solución óptima de } (P) \iff f(\bar{x}) = \alpha = -\beta = -h^*(\bar{u}^*),$$

$$\Updownarrow$$

$$\varphi(\bar{x},0) + \varphi^*(0,\bar{u}^*) = \langle \bar{x}, 0\rangle + \langle 0, \bar{u}^*\rangle$$

$$\Updownarrow$$

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

Por lo tanto, las soluciones óptimas de (P) se deducen del subdiferencial de φ^* en $(0,\bar{u}^*)$, donde $\bar{u}^*$ es una solución óptima de (Q). Análogamente para las soluciones óptimas de (Q).

4.1.1 Noción Lagrangiana

Consideremos la función $L : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$ definida por

$$L(x,u^*) := -\sup_u [\langle u, u^*\rangle - \varphi(x,u)] = \inf_u [\varphi(x,u) - \langle u, u^*\rangle]. \tag{4.1}$$

Puesto que

$$h^*(u^*) = \varphi^*(0,u^*) = \sup_{x,u} [\langle u, u^*\rangle - \varphi(x,u)] \quad \text{y} \quad \beta = \inf_{u^*} h^*(u^*), \tag{4.2}$$

tenemos,

$$-\beta = \sup_{u^*} \inf_x L(x,u^*).$$

De otro lado,

$$\sup_{u^*} L(x,u^*) = \sup_{u^*} \inf_u [\varphi(x,u) + \langle 0 - u, u^*\rangle],$$

y por lo tanto, denotando para cada x fijo, $\varphi_x(u) := \varphi(x,u)$, se cumple

$$\sup_{u^*} L(x,u^*) = \varphi_x^{**}(0) \le \varphi_x(0) = \varphi(x,0) = f(x), \tag{4.3}$$

de donde,

$$\inf_x \sup_{u^*} L(x, u^*) \leq \inf_x f(x) = \alpha.$$

Si φ es convexa sci propia, entonces φ_x es convexa sci y satisface $\varphi_x(u) > -\infty$ para todo $u \in \mathbb{R}^p$. Se deduce que

$$\varphi_x^{**}(0) = \varphi_x(0) = \varphi(x, 0) = f(x), \tag{4.4}$$

y por consiguiente,

$$\inf_x \sup_{u^*} L(x, u^*) = \inf_x f(x) = \alpha.$$

Por lo tanto,

$$-\beta = \sup_{u^*} \inf_x L(x, u^*) \leq \inf_x \sup_{u^*} L(x, u^*) = \alpha.$$

Note que si φ es convexa, entonces para cada $u^* \in \mathbb{R}^p$, la aplicación

$$x \to L(x, u^*) = \inf_u [\varphi(x, u) - \langle u, u^* \rangle]$$

es convexa debido a que la aplicación $(x, u) \to \varphi(x, u) - \langle u, u^* \rangle$ es convexa. De otro lado, para cada x fijo, la aplicación $u^* \to L(x, u^*)$ es cóncava y scs como ínfimo de funciones afínes.

Recordemos que un punto $(\bar{x}, \bar{u}^*) \in \mathbb{R}^n \times \mathbb{R}^p$ es **punto silla** de una función $l : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$, si

$$l(\bar{x}, u^*) \leq l(\bar{x}, \bar{u}^*) \leq l(x, \bar{u}^*) \quad \text{para todo } (x, u^*) \in \mathbb{R}^n \times \mathbb{R}^p.$$

Se deduce que

$$\inf_x \sup_{u^*} l(x, u^*) \leq \sup_{u^*} l(\bar{x}, u^*) = l(\bar{x}, \bar{u}^*)$$

y

$$l(\bar{x}, \bar{u}^*) = \inf_x l(x, \bar{u}^*) \leq \sup_{u^*} \inf_x l(x, u^*).$$

Por lo tanto, debido a que siempre se cumple la relación

$$\inf_x \sup_{u^*} l(x, u^*) \geq \sup_{u^*} \inf_x l(x, \bar{u}^*),$$

se tiene

$$\inf_x \sup_{u^*} l(x, u^*) = L(\bar{x}, \bar{u}^*) = \sup_{u^*} \inf_x l(x, \bar{u}^*).$$

Teorema 4.1 *Asumamos* $\varphi^{**} = \varphi$. *Entonces* $(\bar{x}, \bar{u}^*)$ *es un punto silla del lagrangiano* L*, definido en (4.1), si y solo si* $\bar{x}$ *es una solución óptima de* (P)*,* $\bar{u}^*$ *es una solución óptima de* (Q) *y* $\alpha = -\beta$.

Demostración : Por la regularidad de φ sabemos de (4.3) y (4.4) que $\sup_{u^*} L(\bar{x}, u^*) = f(\bar{x})$. De otro lado, de (4.2), tenemos $\inf_x [L(x, u^*)] = -h^*(u^*)$. Se deduce que $(\bar{x}, \bar{u}^*)$ es un punto silla de L si y solo si $f(\bar{x}) + h^*(\bar{u}^*) = 0$ ó $f(\bar{x}) = \alpha$, $h^*(\bar{u}^*) = \beta$ y $\alpha + \beta = 0$.

Por lo tanto, $(\bar{x}, \bar{u}^*)$ es un punto silla de L si y solo si

$$\left\{ \begin{array}{cl} \alpha + \beta = 0 & \text{(no existe salto de dualidad)}, \\ f(\bar{x}) = \alpha & \bar{x} \text{ es una solución óptima de } (P), \\ h^*(\bar{u}^*) = \beta & \bar{u}^* \text{ es una solución óptima de } (Q). \end{array} \right.$$

4.2 Dualidad en Programación Lineal

Consideremos el siguiente problema de programación lineal

$$\alpha := \inf_{x \geq 0} [\langle a, x \rangle : Bx \geq b], \tag{P}$$

con $a \in \mathbb{R}^n$, $b \in \mathbb{R}^p$ y $B \in \mathbb{R}^{p \times n}$, bajo la hipótesis siguiente :

$$\exists\, \bar{x} \geq 0 \text{ tal que } B\bar{x} \geq b, \tag{H_0}$$

que es equivalente a la condición $\alpha < +\infty$.

Consideremos las funciones

$$f(x) := \left\{ \begin{array}{ll} \langle a, x \rangle & \text{si } Bx \geq b,\ x \geq 0, \\ +\infty & \text{en caso contrario.} \end{array} \right.$$

y

$$\varphi(x, u) := \left\{ \begin{array}{ll} \langle a, x \rangle & \text{si } Bx \geq b + u,\ x \geq 0, \\ +\infty & \text{en caso contrario.} \end{array} \right.$$

Se deduce que f y φ son convexas sci y propias, y satisfacen,

$$f(x) = \varphi(x, 0) \text{ para todo } x \in \mathbb{R}^n.$$

Por lo tanto, la función h definida por

$$h(u) := \inf \varphi(x, u)$$

es convexa y satisfece $h(0) = \alpha$.

De otro lado, debido a que

$$S_\lambda(\varphi) = \{(x, u) : \langle a, x\rangle \leq \lambda, Bx - u \geq b, x \geq 0\}$$

es un poliedro convexo, el conjunto

$$S_\lambda(h) = \bigcap_{\mu > \lambda} \operatorname{proj}_u S_\mu(\varphi),$$

también es un poliedro convexo y por lo tanto cerrado. Se deduce que h es sci.

En resumen, la función h es convexa y sci, lo que significa que

$$-\beta = h^{**}(0) = h(0) = \alpha.$$

Para obtener una expresión explícita del problema dual, calculemos la conjugada de h :

$$h^*(u^*) = \varphi^*(0, u^*) = \sup_{x,u}[\langle u, u^*\rangle - \langle a, x\rangle : Bx - u \geq b,\ x \geq 0].$$

Distinguimos dos casos :

— Si alguna componente (digamos i_0) de u^* es negativa, entonces considerando $x = \bar{x}$, $u_i = (B\bar{x} - b)_i$ para $i \neq i_0$ y $u_{i_0} \to -\infty$, se obtiene $h^*(u^*) = +\infty$.

— Si $u^* \geq 0$, obtenemos

$$h^*(u^*) = \sup_{x \geq 0}[\langle -b + Bx, u^*\rangle - \langle a, x\rangle] = \langle -b, u^*\rangle + \sup_{x \geq 0}[\langle x, B^t u^* - a\rangle],$$

de donde

$$h^*(u^*) = \begin{cases} \langle -b, u^*\rangle & \text{si } B^t u^* - a \leq 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Por lo tanto,

$$-\alpha = \beta = \inf_{u^*}[\langle -b, u^*\rangle : u^* \geq 0,\ \ B^t u^* \leq a].$$

De esta manera

$$\sup_{u^*}[\langle b, u^*\rangle : u^* \geq 0,\ B^t u^* \leq a] = \inf_x[\langle a, x\rangle : x \geq 0,\ Bx \geq b].$$

4.3 Pertubación vertical para problemas con restricciones de desigualdades

Consideremos el problema

$$\alpha := \inf[f(x) : g_i(x) \leq 0,\ i = 1, \cdots, p\,],$$

con las siguientes hipótesis :

(H_0) Las funciones f y g_i (para $i = 1, \cdots, p$) son convexas, sci y propias.

(H_1) Existe $\hat{x}$ tal que $f(\hat{x}) < +\infty$ y $g_i(\hat{x}) < 0$ para todo $i = 1, \cdots, p$.

(H_2) $\alpha > -\infty$.

Note que (H_1) y (H_2) juntas, implican $|\alpha| < +\infty$.

Consideremos las funciones

$$\varphi(x, u) = \begin{cases} f(x) & \text{si } g_i(x) + u_i \leq 0 \text{ para todo } i = 1, \cdots, p, \\ +\infty & \text{en caso contrario.} \end{cases}$$

y

$$h(u) = \inf_x \varphi(x, u).$$

De las hipótesis, φ es convexa sci propia (lo que implica en particular que $\varphi = \varphi^{**}$), de donde, la convexidad de h. Además, $h(u) \leq f(\hat{x})$, desde que $u_i \leq -g_i(\hat{x})$, para todo $i = 1, \cdots, p$.

La última parte implica que

$$V := \prod_{i=1}^{n}]-\infty, -g_i(\hat{x})[\subset \operatorname{dom}(h).$$

Debido a que V es una vecindad de 0, se cumple que $0 \in \operatorname{int}(\operatorname{dom}(h))$.

Por lo tanto,

- $h(0) = h^{**}(0)$ (no hay salto de dualidad).
- $\partial h(0)$ es compacto no vacío, y por consiguiente el conjunto de soluciones óptimas del problema dual es convexo, compacto y no vacío.
- El problema dual correspondiente es

$$-\alpha = \beta = \inf_{u^*} h^*(u^*)$$

con

$$\begin{aligned} h^*(u^*) &= \sup_{x,u} [\langle u, u^* \rangle - \varphi(x,u)] \\ &= \sup_{x,u} [\langle u, u^* \rangle - f(x) : g_i(x) + u_i \leq 0]. \end{aligned}$$

Distinguimos dos casos :

— Si alguna componente (digamos i_0) de u^* es negativa, entonces considerando $x = \hat{x}$, $u_i = -g_i(\hat{x})$ para todo $i \neq i_0$ y $u_{i_0} \to -\infty$, se obtiene $h^*(u^*) = +\infty$.

— Si $u^* \geq 0$, entonces

$$h^*(u^*) = \sup_{x \in \mathrm{dom}\,(f) \cap (\cap_{i=1}^p \mathrm{dom}\,(g_i))} \left[-f(x) - \sum u_i^* g_i(x) \right].$$

Se deduce que

$$-\beta = \sup_{u^* \geq 0} \inf_x \left[L(x, u^*) = f(x) + \sum u_i^* g_i(x) \right],$$

que es la dualidad lagrangiana.

De esta manera, si $\bar{u}^*$ es una solución óptima del problema dual, entonces $\bar{x}$ es una solución óptima del problema primal si, y solo si,

$$f(\bar{x}) + \sum_{i=1}^p \bar{u}_i^* g_i(\bar{x}) \leq f(x) + \sum \bar{u}_i^* g_i(x) \text{ para todo } x \in \mathbb{R}^n \tag{4.5}$$

y

$$f(\bar{x}) + \sum_{i=1}^p u_i^* g_i(\bar{x}) \leq f(\bar{x}) + \sum_{i=1}^p \bar{u}_i^* g_i(\bar{x}) \text{ para todo } u^* \geq 0. \tag{4.6}$$

La última desigualdad es equivalente a las dos condiciones siguientes :

$$g_i(\bar{x}) \leq 0 \;\text{ y }\; \bar{u}_i^* g_i(\bar{x}) = 0 \text{ para todo } i = 1, \cdots, p.$$

En general, la desigualdad en (4.5) no es una condición suficiente de optimalidad del problema primal, como se ve en el siguiente ejemplo.

Ejemplo 4.1 *Consideremos el siguiente problema*

$$\alpha := \inf_x [x : -x \leq 0].$$

Aquí, $\tilde{f}(x) = x$ *y* $g(x) = -x$. *La solución óptima (única) de este problema es* $\bar{x} = 0$. *El problema dual correspondiente es*

$$-\alpha = \inf_{u^*} h^*(u^*),$$

con

$$h^*(u^*) = \sup_x [-x + u^* x] = \begin{cases} 0 & \textit{si} \;\; u^* = 1, \\ +\infty & \textit{en caso contrario.} \end{cases}$$

Luego, la solución óptima (única) del problema dual es $\bar{u}^* = 1$. *De otro lado, el conjunto de los* x'*s satisfaciendo (4.5) es todo* $\mathbb{R}$.

4.4 Perturbación Vertical - Caso General

Consideremos el problema

$$\alpha := \inf[f(x) : g_i(x) \leq 0, \; i = 1, \cdots p, \; Ax + a = 0],$$

con las siguientes hipótesis :

(H_1) Las funciones f y g_i ($i = 1, \cdots, p$) son convexas, sci y propias, $a \in \mathbb{R}^q$ y A es una matriz de orden $q \times n$ de rango q.

(H_2) Existe $\hat{x} \in \operatorname{int}(\operatorname{dom}(f)) \cap [\cap_{i=1}^p \operatorname{int}(\operatorname{dom}(g_i))]$ tal que $A\hat{x} + a = 0$ y $g_i(\hat{x}) < 0$ para todo $i = 1, \cdots, p$.

La hipótesis (H_2) implica que $\alpha < +\infty$.

Consideremos las funciones

$$\varphi(x, u, v) = \begin{cases} f(x) & \text{si } g_i(x) + u_i \leq 0, \; Ax + a + v = 0, \\ +\infty & \text{en caso contrario,} \end{cases}$$

y

$$h(u, v) = \inf_x \varphi(x, u, v).$$

Se deduce que $\alpha = h(0,0)$ y, en virtud de las hipótesis (H_1) y (H_2), la función φ es convexa sci propia y, por lo tanto, h es convexa.

Mostraremos que h es continua en $(0,0)$. Para ello mostraremos que $(0,0) \in \operatorname{int}(\operatorname{dom}(h))$. Por la hipótesis (H_2), existe una vecindad X de $\hat{x}$ tal que para todo $x \in X$ y para todo $i = 1, \cdots, p$,

$$f(x) \leq f(\hat{x}) + 1 < +\infty \quad \text{y} \quad g_i(x) \leq \frac{1}{2} g_i(\hat{x}) < 0.$$

Además, debido a que A es de rango máximo y $A\hat{x} + a = 0$, existe una vecindad V de 0 en $\mathbb{R}^q$ tal que para todo $v \in V$, existe $x \in X$ con $Ax+a+v = 0$. También, debido a que $g_i(\hat{x}) < 0$, existe una vecindad U de 0 en $\mathbb{R}^p$ tal que para todo $u \in U$,

$$\frac{1}{2} g_i(\hat{x}) + u_i \leq 0 \text{ para todo } i = 1, \cdots, p.$$

Por lo tanto, si $(u,v) \in U \times V$, existe $x \in X$ tal que

$$Ax + a + v = 0, \quad g_i(x) + u_i \leq 0 \quad \text{y} \quad f(x) \leq f(\hat{x}) + 1.$$

Luego

$$h(u,v) \leq f(\hat{x}) + 1 < +\infty \;\text{ para todo } (u,v) \in U \times V,$$

y por consiguiente, $(0,0) \in \operatorname{int}(\operatorname{dom}(h))$. Se deduce por lo tanto que h es continua en $(0,0)$ y $\alpha = -\beta$.

Si $\alpha > -\infty$, entonces el conjunto de soluciones óptimas del problema (Q) es compacto no vacío.

El problema dual correspondiente es

$$-\alpha = \beta = \inf_{u^*, v^*} h^*(u^*, v^*)$$

con

$$\begin{aligned} h^*(u^*, v^*) &= \sup_{x,u,v} [\langle u, u^* \rangle + \langle v, v^* \rangle - \varphi(x,u,v)] \\ &= \sup_{u,v,x} [\langle u, u^* \rangle + \langle v, v^* \rangle - f(x) : g(x) + u \leq 0,\; Ax + a + v = 0]. \end{aligned}$$

Distinguimos dos casos :

— Si alguna componente (digamos i_0) de u^* es negativa, entonces considerando $x = \hat{x}$, $u_i = 0$ para todo $i \neq i_0$, $v_j = 0$ para todo $j = 1, \cdots, q$ y $u_{i_0} \to -\infty$, se obtiene $h^*(u^*, v^*) = +\infty$.

— Si $u^* \geq 0$, denotando $D = \operatorname{dom}(f) \cap (\cap_{i=1}^p \operatorname{dom}(g_i))$, obtenemos

$$h^*(u^*, v^*) = \sup_{x \in D} [-\langle a + Ax, v^* \rangle - \sum u_i^* g_i(x) - f(x)].$$

Se deduce por lo tanto que

$$\alpha = -\beta = \sup_{u^* \geq 0, v^*} \inf_{x \in D} [\langle a + Ax, v^* \rangle + \sum u_i^* g_i(x) + f(x)].$$

Condición débil de (H_2) : referente a la condición de la hipótesis (H_2) podemos debilitarla por la siguiente condición :

(H_2') Existe $\hat{x} \in \operatorname{ri}(\operatorname{dom}(f)) \cap [\cap_{i=1}^p \operatorname{ri}(\operatorname{dom}(g_i))]$ tal que $A\hat{x} + a = 0$ y $g_i(\hat{x}) < 0$ para todo $i = 1, \cdots, p$.

Esta condición implica que $(0, 0) \in \operatorname{ri}(\operatorname{dom}(h))$. En efecto, es suficiente demostrar que si $(\bar{u}, \bar{v}) \in \operatorname{dom}(h)$, entonces $(-\bar{t}\bar{u}, -\bar{t}\bar{v}) \in \operatorname{dom}(h)$ para algún $\bar{t} > 0$. Sea $(\bar{u}, \bar{v}) \in \operatorname{dom}(h)$ fijo arbitrario. Entonces existe $\bar{x} \in \operatorname{dom}(f)$ tal que

$$g(\bar{x}) + \bar{u} \leq 0 \quad \text{y} \quad A\bar{x} + a + \bar{v} = 0, \tag{4.7}$$

de donde, en particular, $\bar{x} \in \operatorname{dom}(g_i)$. De otro lado, debido a que $\hat{x} \in \operatorname{ri}(\operatorname{dom}(f)) \cap [\cap_{i=1}^p \operatorname{ri}(\operatorname{dom}(g_i))]$, existe $\hat{t} > 0$ tal que

$$\hat{x} + t(\hat{x} - \bar{x}) \in \operatorname{ri}(\operatorname{dom}(f)) \cap [\cap_{i=1}^p \operatorname{ri}(\operatorname{dom}(g_i))] \quad \text{para todo } t \in [-1, \hat{t}]. \tag{4.8}$$

Para cada $i = 1, \cdots, p$, consideremos la función $r_i : (-\infty, \hat{t}\,] \to \overline{\mathbb{R}}$ definida por

$$r_i(t) := g_i(\hat{x} + t(\hat{x} - \bar{x})).$$

Se sigue que $0 \in \operatorname{int}(\operatorname{dom}(r_i))$ y por lo tanto r_i es continua en 0. Debido a que $r_i(0) = g_i(\hat{x}) < 0$, existe $\bar{t} \in (0, \hat{t}\,]$ tal que

$$g_i(\hat{x} + \bar{t}(\hat{x} - \bar{x})) - \bar{t}\bar{u}_i = r_i(\bar{t}) - \bar{t}\bar{u}_i \leq 0 \quad \text{para todo } i = 1, \cdots, p. \tag{4.9}$$

De otro lado, por la igualdad $A\hat{x} + a = 0$ (hipótesis (H_2')) y la igualdad en (4.7), tenemos

$$A(\hat{x} + \bar{t}(\hat{x} - \bar{x})) + a - \bar{t}\bar{v} = 0. \tag{4.10}$$

Así, de (4.8), (4.9) y (4.10), tenemos

$$h(-\bar{t}\bar{u}, -\bar{t}\bar{v}) \leq f(\hat{x} + \bar{t}(\hat{x} - \bar{x})) < +\infty,$$

y por lo tanto, $(-t\bar{u}, -t\bar{v}) \in \operatorname{dom}(h)$. ■

4.5 Ejemplo de perturbaciones no Verticales

Consideremos el siguiente problema

$$\alpha := \inf_{x \in \mathbb{R}^n} [f(Ax) + g(Bx)],$$

con las siguientes hipótesis :

(H_1) Las funciones f y g son convexas, sci y propias; A y B son matrices de orden $p \times n$ y $q \times n$, respectivamente.

(H_2) Existe $\hat{x} \in \mathbb{R}^n$ tal que $A\hat{x} \in \operatorname{int} \operatorname{dom}(f)$ y $B\hat{x} \in \operatorname{int}(\operatorname{dom}(g))$.

La hipótesis (H_2) implica $\alpha < +\infty$.

Consideremos las funciones

$$\varphi(x, u, v) := f(Ax + u) + g(Bx + v)$$

y

$$h(u, v) := \inf_x \varphi(x, u, v).$$

Tenemos, $h(0, 0) = \alpha$ y, por las hipótesis precedentes, φ es convexa sci propia, y por lo tanto, h es convexa.

Mostraremos que h es continua en $(0, 0)$. En efecto, debido a que $A\hat{x} \in \operatorname{int} \operatorname{dom}(f)$, existe una vecindad U de 0 en $\mathbb{R}^p$ tal que $f(A\hat{x}+u) \leq f(A\hat{x})+1$ para todo $u \in U$. Análogamente, existe una vecindad V de 0 en $\mathbb{R}^q$ tal que $g(B\hat{x} + v) \leq g(B\hat{x}) + 1$ para todo $v \in V$. Por lo tanto,

$$h(u, v) \leq f(A\hat{x}) + g(B\hat{x}) + 2 \quad \text{para todo } (u, v) \in U \times V.$$

Se deduce que $(0, 0) \in \operatorname{int}(\operatorname{dom}(h))$, lo que implica que h es continua en $(0, 0)$ y $\alpha = -\beta$.

Si $\alpha > -\infty$, entonces el conjunto de soluciones óptimas del problema dual es convexo, compacto y no vacío.

El problema dual correspondiente es

$$-\alpha = \beta = \inf_{u^*,v^*} h^*(u^*, v^*),$$

donde

$$\begin{aligned}
h^*(u^*, v^*) &= \sup_{u,v,x}[\langle u, u^*\rangle + \langle v, v^*\rangle - f(Ax+u) - g(Bx+v)] \\
&= \sup_{u,v,x}[\langle Ax+u, u^*\rangle - f(Ax+u) + \langle Bx+v, v^*\rangle \\
&\qquad \cdots - g(Bx+v) - \langle x, A^t u^* + B^t v^*\rangle] \\
&= \sup_{x,y,z}[\langle y, u^*\rangle - f(y) + \langle z, v^*\rangle - g(z) - \langle x, A^t u^* + B^t v^*\rangle] \\
&= \begin{cases} f^*(u^*) + g^*(v^*) & \text{si } A^t u^* + B^t v^* = 0, \\ +\infty & \text{en caso contrario .} \end{cases}
\end{aligned}$$

Por consiguiente,

$$-\alpha = \beta = \inf_{u^*,v^*}[f^*(u^*) + g^*(v^*) : A^t u^* + B^t v^* = 0].$$

Aquí, el lagrangiano es

$$l(x, u^*, v^*) = \langle x, A^t u^* + B^t v^*\rangle - f^*(u^*) - g^*(v^*).$$

Así, si $(\bar{u}^*, \bar{v}^*)$ es una solución óptima del problema dual, entonces $\bar{x}$ es una solución óptima del problema primal si y solo si

$$l(\bar{x}, u^*, v^*) \leq l(\bar{x}, \bar{u}^*, \bar{v}^*) \leq l(x, \bar{u}^*, \bar{v}^*) \text{ para todo } x, u^*, v^*,$$

lo que significa que

$$A\bar{x} \in \partial f^*(\bar{u}^*), \quad B\bar{x} \in \partial g^*(\bar{v}^*) \quad \text{y} \quad A^t\bar{u}^* + B^t\bar{v}^* = 0.$$

El caso particular $A = B = I$ (matriz identidad), los problemas primal y dual son

$$\alpha = \inf_{x\in\mathbb{R}^n}[f(x) + g(x)] \quad \text{(Problema primal)}$$

y

$$\beta = \inf_{x^*}[f^*(x^*) + g^*(-x^*)] \qquad \text{(Problema dual)}.$$

Finalmente, similar a los casos anteriores, podemos debilitar la condición de la hipótesis (H_2) por la condición :

(H_2') Existe $\hat{x} \in \mathbb{R}^n$ tal que $A\hat{x} \in \mathrm{ri}\,(\mathrm{dom}\,(f))$ y $B\hat{x} \in \mathrm{ri}\,(\mathrm{dom}\,(g))$.

Esta condición implica que $0 \in \mathrm{ri}\,(\mathrm{dom}\,(h))$ y, por lo tanto, se cumple $-\alpha = \beta$.

4.6 Inf–convolución y el subdiferencial de la suma de dos funciones convexas

Para $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$, convexas y propias, sean $A = \mathrm{epi}\,(f)$, $B = \mathrm{epi}\,(g)$ y $C = \mathrm{epi}\,(f) + \mathrm{epi}\,(g)$. Por ser A y B conjuntos convexos, C también es convexo y satisface la siguiente propiedad :

$$(x, \lambda) \in C \ \text{ y } \ \mu > \lambda \ \Rightarrow \ (x, \mu) \in C.$$

Sea $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ definida por $h(x) := \inf_\lambda[\lambda : (x, \lambda) \in C]$ para todo $x \in \mathbb{R}^n$. Esta es convexa debido a que C es convexo. Además, $C \subset \mathrm{epi}\,(h) \subset \overline{C}$.

Note que h no es necesariamente sci porque C no es necesariamente cerrado, aún cuando A y B lo fuesen.

De la definición resulta

$$\begin{aligned} h(x) &= \inf[\lambda_1 + \lambda_2 : (x_1, \lambda_1) \in A,\ (x_2, \lambda_2) \in B] \\ &= \inf[f(x_1) + g(x_2) : x = x_1 + x_2]. \end{aligned}$$

Denotaremos a la función h por

$$h = f \nabla g.$$

Se deduce de la defición que $f \nabla g = g \nabla f$ y

$$h(x) = \inf_y[f(x - y) + g(y)] \ \text{ para todo } x \in \mathbb{R}^n.$$

Proposición 4.2 *Sean f y g convexas, sci y propias. Entonces,*

a) $(f\nabla g)^* = f^* + g^*$;

b) $(f+g)^* = \overline{f^*\nabla g^*}$.

Demostración : a) Sea $h = f\nabla g$. Tenemos

$$\begin{aligned} h^*(x^*) &= \sup_{x,y}[\langle x, x^*\rangle - f(x-y) - g(y)] \\ &= \sup_{x,y}[\langle x-y, x^*\rangle + \langle y, x^*\rangle - f(x-y) - g(y)] \\ &= f^*(x^*) + g^*(x^*). \end{aligned}$$

b) De la primera parte,

$$(f^* + g^*)^* = (f\nabla g)^{**} = \overline{(f\nabla g)},$$

y por lo tanto se cumple la identidad de la parte **b)**. ∎

Haremos uso de estas identidades para el cálculo del subgradiente de la suma de dos funciones convexas.

Antes veamos el siguiente resultado.

Proposición 4.3 *Sean f y g convexas y propias, $s = f + g$ y $a \in \operatorname{dom}(s)$. Entonces*

$$\partial s(a) \supset \partial f(a) + \partial g(a).$$

Demostración : $x^* \in \partial f(a)$ si y solo si $f(a) + \langle x^*, x - a\rangle \leq f(x)$ para todo x. Análogamente, $y^* \in \partial g(a)$ si y solo si $g(a) + \langle y^*, x - a\rangle \leq g(x)$ para todo x. Se deduce que

$$s(a) + \langle x^* + y^*, x - a\rangle \leq s(x) \quad \text{para todo } x,$$

lo que significa que $x^* + y^* \in \partial s(a)$. ∎

Teorema 4.2 *Sean f y g convexas, sci y propias y $s = f+g$. Si* $\operatorname{ri}(\operatorname{dom}(f)) \cap \operatorname{ri}(\operatorname{dom}(g)) \neq \emptyset$, *entonces*

a) $(f+g)^* = f^*\nabla g^*$.

b) *Para todo* $a^* \in \mathrm{dom}\,(f^*\nabla g^*)$ *existen* x^* *y* y^* *tales que* $x^* + y^* = a^*$ *y* $(f^*\nabla g^*)(a^*) = f^*(x^*) + g^*(y^*)$.

c) *Si* $\partial f(a) \neq \emptyset$ *y* $\partial g(a) \neq \emptyset$, *entonces* $\partial f(a) + \partial g(a) = \partial s(a)$.

Demostración : La función s es convexa, sci y propia, y

$$-s^*(a^*) = \inf_x [f(x) + g(x) - \langle a^*, x\rangle]. \qquad (D)$$

Por definición, s^* es convexa y sci. Esta también es propia debido a que s lo es. Por lo tanto $-s^*(a^*) < +\infty$ para todo a^*. Sean

$$\varphi(x,u) := f(x+u) + g(x) - \langle a^*, x\rangle \quad \text{ y } \quad h(u) := \inf_x \varphi(x,u).$$

La función φ es convexa (y por lo tanto, h también), sci y propia, y por consiguiente, $\varphi^{**} = \varphi$.

Mostraremos que $0 \in \mathrm{ri}\,(\mathrm{dom}\,(h))$. Sea $u \in \mathrm{dom}\,(h)$. Entonces existe $\bar{x}$ tal que $\bar{x} + u \in \mathrm{dom}\,(f)$ y $\bar{x} \in \mathrm{dom}\,(g)$.

Sea $\hat{x} \in \mathrm{ri}\,(\mathrm{dom}\,(f)) \cap \mathrm{ri}\,(\mathrm{dom}\,(g))$. Entonces existe $t > 0$ tal que

$$\hat{x} + t(\hat{x} - \bar{x} - u) \in \mathrm{dom}\,(f) \quad \text{ y } \quad \hat{x} + t(\hat{x} - \bar{x}) \in \mathrm{dom}\,(g),$$

de donde,

$$h(-tu) \leq f(\hat{x} + t(\hat{x} - \bar{x} - u)) + g(\hat{x} + t(\hat{x} - \bar{x})) - \langle a^*, \hat{x} + t(\hat{x} - \bar{x})\rangle < +\infty,$$

lo que significa que $-tu \in \mathrm{dom}\,(h)$. Por lo tanto, $0 \in \mathrm{ri}\,(\mathrm{dom}\,(h))$, lo que implica que no hay salto de dualidad y que el problema dual admite solución si $s^*(a^*)$ es finito.

Tenemos

$$\begin{aligned} s^*(a^*) &= \inf_{u^*} \varphi^*(0, u^*) \\ &= \inf_{u^*} [f^*(u^*) + g^*(a^* - u^*)] \\ &= (f^*\nabla g^*)(a^*). \end{aligned}$$

La segunda igualdad se debe a que

$$\begin{aligned} \varphi^*(x^*, u^*) &= \sup_{x,u} [\langle x, x^*\rangle + \langle u, u^*\rangle - f(x+u) - g(x) + \langle a^*, x\rangle] \\ &= \sup_{x,u} [\langle x+u, u^*\rangle - f(x+u) + \langle x, x^* + a^* - u^*\rangle - g(x)] \\ &= f^*(u^*) + g^*(x^* + a^* - u^*). \end{aligned}$$

Por lo tanto, si $a^* \in \text{dom}\,(f^*\nabla g^*)$, entonces $s^*(a^*)$ es finito y por consiguiente existe u^* tal que

$$s^*(a^*) = \varphi^*(0, u^*) = f^*(u^*) + g^*(a^* - u^*) = (f^*\nabla g^*)(a^*).$$

Mostremos ahora la igualdad del item c). Sea $a^* \in \partial s(a)$. Entonces $s^*(a^*) + s(a) = \langle a, a^* \rangle$ y por consiguiente, existen x^* y y^*, con $x^* + y^* = a^*$, tales que

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

Debido a que $f^*(x^*) + f(a) \geq \langle a, x^* \rangle$ y $g^*(y^*) + g(a) \geq \langle a, y^* \rangle$, se cumplen

$$f^*(x^*) + f(a) = \langle a, x^* \rangle \quad \text{y} \quad g^*(y^*) + g(a) = \langle a, y^* \rangle,$$

que es equivalente a

$$x^* \in \partial f(a) \quad \text{e} \quad y^* \in \partial g(a).$$

Por lo tanto $\partial s(a) \subset \partial f(a) + \partial g(a)$. ■

Comentario: La condición $\text{ri}\,(\text{dom}\,(f)) \cap \text{ri}\,(\text{dom}\,(g)) \neq \emptyset$ en el Teorema 4.2 no es superflua como se nota en el siguiente ejemplo. Sean $f, g : \mathbb{R}^2 \to \overline{\mathbb{R}}$ definidas por $f = \delta(\cdot|C_1)$ y $g = \delta(\cdot|C_2)$, con $C_1 = \{(x,y) : y \geq x^2\}$ y $C_2 = \{(x,y) : y = 0\}$. Tenemos $f\nabla g = \delta(\cdot|C_3)$ y $f^* + g^* = \delta(\cdot|C_4)$, con $C_3 = \{(x,y) : y \geq 0\}$ y $C_4 = \{(x,y) : x = 0\}$.

4.7 El subdiferencial de la función $\max_{i\in I} f_i$

Sean $f_1, \cdots, f_p$ convexas, sci y propias de $\mathbb{R}^n$ en $\overline{\mathbb{R}}$. Consideremos la función h definida por

$$h(x) := \max_{i=1,\cdots,p} f_i(x) \text{ para todo } x \in \mathbb{R}^n,$$

y asumamos que su dominio, $\text{dom}\,(h) = \cap_{i=1}^p \text{dom}\,(f_i)$, es no vacío. Entonces, h es convexa, sci y propia.

Calculemos el subdiferencial $\partial h(a)$ en $a \in \text{dom}\,(h)$. Sabemos que

$$a^* \in \partial h(a) \text{ si y solo si } \langle a^*, x - a \rangle + h(a) \leq h(x) \text{ para todo } x \in \mathbb{R}^n,$$

que es equivalente a

$$h(a) - \langle a^*, a\rangle \leq h(x) - \langle x^*, x\rangle \text{ para todo } x \in \mathbb{R}^n,$$

lo que significa que a es una solución óptima del problema

$$\alpha := \inf_x [h(x) - \langle a^*, x\rangle]$$

o equivalentemente del problema

$$\alpha = \inf_{x,t} [t - \langle a^*, x\rangle : f_i(x) - t \leq 0,\ i = 1, \cdots, p]. \tag{P}$$

Note que la condición de Slater se cumple en esta última formulación (haciendo por ejemplo $t = 1 + h(a)$). Por consiguiente, $a^* \in \partial h(a)$ si y solo si existe $\bar{u}^* \geq 0$ tal que la terna $(a, \bar{t} = h(a), \bar{u}^*)$ es un punto silla del lagrangiano

$$l(x, t, u^*) := t - \langle a^*, x\rangle + \sum_{i=1}^{p} u_i^*(f_i(x) - t).$$

Por lo tanto,

$$\partial h(a) = \left\{ a^* \in \partial \left(\sum_{i=1}^{p} u_i^* f_i \right)(a) : \begin{array}{l} u^* \geq 0,\ \sum_{i=1}^{p} u_i^* = 1, \\ u_i^*(f_i(a) - h(a)) = 0 \text{ para todo } i \end{array} \right\}.$$

Así, si $\cap_{i=1}^{p} \mathrm{ri}\,(\mathrm{dom}\,(f_i)) \neq \emptyset$, entonces

$$\partial h(a) = \left\{ a^* = \sum_{i=1}^{p} u_i^* a_i^* : \begin{array}{l} u^* \geq 0,\ \sum_{i=1}^{p} u_i^* = 1,\ a_i^* \in \partial f_i(a), \\ u_i^*(f_i(a) - h(a)) = 0 \text{ para todo } i \end{array} \right\}.$$

Capítulo 5

Monotonía de una multiaplicación

5.1 Nociones generales de las multiaplicaciones

Sean X y Y dos conjuntos arbitrarios. Se dice que Γ es una **multiaplicación** (aplicación multivaluada, correspondencia,$\cdots$) de X en Y, y lo denotaremos por $\Gamma : X \rightrightarrows Y$, si para cada $x \in X$, $\Gamma(x)$ es un subconjunto de Y. La gráfica de Γ es el conjunto

$$\mathrm{gr}\,(\Gamma) = \{(x, y) \in X \times Y : y \in \Gamma(x)\},$$

que es un subconjunto de $X \times Y$.

Recíprocamente, dado un subconjunto G de $X \times Y$, la multiaplicación Γ definida por $\Gamma(x) = \{y : (x, y) \in G\}$ tiene como gráfica al conjunto G.

En general se cumple

$$\Gamma(x) = \{y : (x, y) \in \mathrm{gr}\,(\Gamma)\},$$

lo que significa que una multiaplicación queda completamente definida a partir de su gráfica.

De forma natural definimos la inversa de Γ, denotada por $\Gamma^{-1} : Y \rightrightarrows X$, como

$$\Gamma^{-1}(y) = \{x : (x, y) \in \mathrm{gr}\,(\Gamma)\}.$$

Se cumple por lo tanto que $(\Gamma^{-1})^{-1} = \Gamma$.

A diferencia de las aplicaciones univaluadas (funciones), cada multiaplicación admite una inversa, que también es una multiaplicación . Por ejemplo, si $\Gamma : \mathbb{R} \rightrightarrows \mathbb{R}$ es definida por $\Gamma(x) = \{x^2\}$ para todo $x \in \mathbb{R}$, entonces

$$\Gamma^{-1}(y) = \begin{cases} \{+\sqrt{y}, -\sqrt{y}\} & \text{si } y \geq 0, \\ \emptyset & \text{en caso contrario.} \end{cases}$$

El **dominio** y la **imagen** de $\Gamma : X \rightrightarrows Y$, son respectivamente

$$\mathrm{dom}\,(\Gamma) = \{x \in X : \Gamma(x) \neq \emptyset\} \;\text{ y }\; \mathrm{img}\,(\Gamma) = \{y \in Y : \exists\, x \in X \text{ con } y \in \Gamma(x)\}.$$

Se cumplen,

$$\mathrm{dom}\,(\Gamma) = \mathrm{proj}_X[\mathrm{gr}\,(\Gamma)] = \mathrm{img}\,(\Gamma^{-1})$$

y

$$\mathrm{dom}\,(\Gamma^{-1}) = \mathrm{proj}_Y[\mathrm{gr}\,(\Gamma)] = \mathrm{img}\,(\Gamma).$$

Si $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es una función convexa sci propia, entonces ∂f es una multiaplicación de $\mathbb{R}^n$ en sí mismo. De otro lado, debido a que $x^* \in \partial f(x)$ si y solo si $x \in \partial f^*(x^*)$, se deduce que ∂f^* es la multiaplicación inversa de ∂f.

5.2 Monotonía del subdiferencial

Si $I \subset \mathbb{R}$ un intervalo y $f : I \to \mathbb{R}$ convexa diferenciable, entonces f' es creciente, es decir,

$$f'(x_1) \leq f'(x_2) \;\text{ si } x_1 < x_2,$$

que es equivalente a

$$(f'(x_2) - f'(x_1))(x_2 - x_1) \geq 0 \text{ para todo } x_1, x_2 \in I. \tag{5.1}$$

Una propiedad análoga se cumple para las funciones convexas definidas en $\mathbb{R}^n$. Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función convexa y propia. Para $(x_i, x_i^*) \in \mathrm{gr}\,(\partial f)$,

$i = 0, \cdots, p$, tenemos

$$\begin{array}{rcl} \langle x_0^*, x_1 - x_0 \rangle & \leq & f(x_1) - f(x_0), \\ \langle x_1^*, x_2 - x_1 \rangle & \leq & f(x_2) - f(x_1), \\ & \vdots & \\ \langle x_{p-1}^*, x_p - x_{p-1} \rangle & \leq & f(x_p) - f(x_{p-1}), \\ \langle x_p^*, x_0 - x_p \rangle & \leq & f(x_0) - f(x_p). \end{array}$$

Denotando $x_{p+1} = x_0$ y sumando todas las expresiones anteriores, obtenemos

$$\sum_{i=0}^{p} \langle x_i^*, x_{i+1} - x_i \rangle \leq 0.$$

Se dice entonces que ∂f es una multiaplicación **cíclicamente monótona**. En particular

$$\langle x_0^*, x_1 - x_0 \rangle + \langle x_1^*, x_0 - x_1 \rangle \leq 0,$$

esto es,

$$\langle x_1^* - x_0^*, x_1 - x_0 \rangle \geq 0, \text{ para todo } x_1^* \in \partial f(x_1),\ x_0^* \in \partial f(x_0),$$

que generaliza la propiedad (5.1).

El siguiente resultado establece, en particular, una condición necesaria y suficiente para que una función creciente definida sobre $\mathbb{R}$ (o sobre un intervalo de $\mathbb{R}$) sea la derivada de una función convexa.

Teorema 5.1 *Si Γ es una multiaplicación maximal cíclicamente monótona (esto es, si Σ es cíclicamente monótona y $\mathrm{gr}\,(\Sigma) \supset \mathrm{gr}\,(\Gamma)$, entonces $\Gamma = \Sigma$), entonces existe una función f convexa, sci y propia tal que $\Gamma = \partial f$.*

La demostración de este maravilloso resultado se puede encontrar en el texto clásico [7].

5.3 Continuidad de una multiaplicación

Sean X y Y dos espacios topológicos y $\Gamma : X \rightrightarrows Y$. Γ es **semicontinua inferior** (sci) en $\bar{x} \in X$, si para todo conjunto abierto Ω de Y satisfaciendo $\Omega \cap \Gamma(\bar{x}) \neq \emptyset$, existe una vecindad V de $\bar{x}$ (conjunto abierto de X conteniendo a $\bar{x}$) tal que

$$\Omega \cap \Gamma(x) \neq \emptyset \text{ para todo } x \in V.$$

Γ es **semicontinua superior** (scs) en $\bar{x}$, si para todo conjunto abierto Ω de Y satisfaciendo $\Omega \supset \Gamma(\bar{x})$, existe una vecindad V de $\bar{x}$ tal que

$$\Omega \supset \Gamma(x) \text{ para todo } x \in V.$$

Γ es sci (scs) en $Z \subset X$ (sin hacer referencia al conjunto cuando $Z = X$), si este es sci (scs) en todo punto de Z.

Γ es cerrada si su gráfica es un conjunto cerrado en $X \times Y$.

Finalmente, Γ es de valor cerrado (compacto, $\cdots$) en $Z \subset X$ (sin hacer referencia al conjunto cuando $Z = X$) si $\Gamma(x)$ es cerrado (compacto, $\cdots$), para todo $x \in Z$.

Teorema 5.2 *Sean $\Gamma : X \rightrightarrows Y$ scs y $K \subset X$ compacto. Si Γ es de valor compacto en K, entonces*

$$\Gamma(K) := \bigcup_{x \in K} \Gamma(x) = \{y \in \mathbb{R}^p : \exists\, x \in K \text{ con } y \in \Gamma(x)\}$$

es compacto de Y.

Demostración : Sea $\{W_i\}_{i \in I}$ un cubrimiento abierto de $\Gamma(K)$. Vamos a mostrar que existe un conjunto finito $J \subset I$ tal que $\cup_{j \in J} W_j \supset \Gamma(K)$. En efecto, sea $x \in K$. Debido a la compacidad de $\Gamma(x)$, existe $I(x) \subset I$ finito tal que $\Gamma(x) \subset \left(\cup_{i \in I(x)} W_i\right) =: W(x)$, que es un conjunto abierto. Por la semicontinuidad superior de Γ, existe una vecindad $V(x)$ de x tal que $\Gamma(x') \subset W(x)$ para todo $x' \in V(x)$. Debido a que $\cup_{x \in K} V(x) \supset K$, existen $x_1, \cdots, x_p$ tales que $\cup_{i=1}^{p} V(x_i) \supset K$. Por lo tanto

$$\Gamma(K) \subset \bigcup_{i=1}^{p} W(x_i) = \bigcup_{i=1}^{p} \left(\bigcup_{j \in I(x_i)} W_j \right),$$

lo que demuestra que $\Gamma(K)$ es recubierto por una cantidad finita de $W_j's$. ■

Teorema 5.3 *Sea $\Gamma : X \rightrightarrows Y$ con $X \subset \mathbb{R}^n$ y $Y \subset \mathbb{R}^p$. Se cumplen las siguientes propiedades :*

a) *Si Γ es scs y de valor cerrado, entonces Γ es cerrada.*

b) *Si Y es compacto y Γ cerrada, entonces Γ es de valor compacto y scs.*

Demostración : a) Sea $(\bar{x}, \bar{y}) \notin \operatorname{gr}(\Gamma)$, es decir $\bar{y} \notin \Gamma(\bar{x})$. Siendo $\Gamma(\bar{x})$ cerrado, existe un abierto W de Y, con $\bar{y} \in W$, tal que $\Gamma(\bar{x}) \subset [\overline{W}]^c$ que es abierto en Y. Por la semicontinuidad superior de Γ, existe un abierto V de X, con $\bar{x} \in V$, tal que $\Gamma(x) \subset [\overline{W}]^c$ para todo $x \in V$, lo que significa que

$$\operatorname{gr}(\Gamma) \cap (V \times W) = \emptyset.$$

b) Note que $\{x\} \times \Gamma(x) = \operatorname{gr}(\Gamma) \cap (\{x\} \times Y)$. Entonces $\Gamma(x)$ es compacto debido a que Y lo es. Si Γ no es scs en algún $\bar{x} \in X$, entonces existe un conjunto abierto Ω de Y conteniendo a $\Gamma(\bar{x})$ en el que cualquier abierto V de X, con $\bar{x} \in V$, contiene un elemento donde Γ no está contenido en Ω. Así entonces, para todo $n \in \mathbb{N}$, existe $x_n \in B(\bar{x}, \frac{1}{n}) \cap X$ y $y_n \in \Gamma(x_n)$ tal que $y_n \notin \Omega$. Por la compacidad de Y, existe una subsucesión convergente de $\{y_n\}$, con límite $\bar{y}$, lo que implica que $(\bar{x}, \bar{y}) \in \operatorname{gr}(\Gamma)$, debido a que Γ es cerrada.

De otro lado, debido a que $Y \cap \Omega^c$ es compacto (entonces cerrado) y $\{y_n\} \in Y \cap \Omega^c$, tenemos $\bar{y} \in Y \cap \Omega^c$, lo que significa que $(\bar{x}, \bar{y}) \notin \operatorname{gr}(\Gamma)$, en contradicción con lo anterior. Por consiguiente, Γ es scs. ■

Sean X un espacio topológico y $f : X \to \overline{\mathbb{R}}$. La definición de semicontinuidad de f es análoga a la definición dada en el Capítulo 2 : f es sci en $\bar{x} \in X$ si para cada $\lambda < f(\bar{x})$, existe una vecindad V de $\bar{x}$ (abierto de X conteniendo a $\bar{x}$) tal que $\lambda < f(x)$ para todo $x \in V$. Esta función es sci si este es semicontinua inferior en todo punto de X.

Se cumple que f es sci si y solo si $\operatorname{epi}(f)$ es cerrado en $X \times \mathbb{R}$ si y solo si para cada $\lambda \in \mathbb{R}$, el conjunto $S_\lambda(f)$ es cerrado en X.

Análogamente para la semicontinuidad superior.

Teorema 5.4 *Sean X y Y espacios topológicos y $\varphi : X \times Y \to \overline{\mathbb{R}}$ sci. Entonces $\Gamma : X \rightrightarrows Y$ definida por*

$$\Gamma(x) = \{y : \varphi(x, y) \leq 0\},$$

es cerrada.

Demostración : Se deduce del hecho que $\mathrm{gr}\,(\Gamma) = \{(x, y) : \varphi(x, y) \leq 0\}$ es cerrado debido a que φ es sci. ■

5.4 Continuidad del subgradiente

Teorema 5.5 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ es de valor convexo compacto no vacío y scs en* $\mathrm{int}\,(\mathrm{dom}\,(f))$.

Demostración : Por lo que ya vimos en el Capítulo 3, sólo nos queda mostrar que ∂f es scs en $\mathrm{int}\,(\mathrm{dom}\,(f))$. Sea $\bar{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Entonces, existe $r > 0$ tal que $B(\bar{x}, 3r) \subset \mathrm{int}\,(\mathrm{dom}\,(f))$, de donde, por la continuidad de f en esta bola, existen m_1 y m_2 en $\mathbb{R}$ tales que $m_1 \leq f(x) \leq m_2$ para todo $x \in B(\bar{x}, 2r)$. Sean B la clausura de $B(\bar{x}, r)$, $x \in B$ y $x^* \in \partial f(x)$. Tenemos,

$$\langle x^*, y - x\rangle \leq f(y) - f(x) \quad \text{para todo } y \in B$$

y por consiguiente

$$\langle x^*, y - x\rangle \leq m_2 - m_1 \quad \text{para todo } y \in B.$$

Por lo tanto, para todo $x \in B$,

$$\partial f(x) \subset C := \{x^* \in \mathbb{R}^n : \|x^*\| \leq \frac{m_2 - m_1}{r}\} \qquad \text{(compacto).}$$

De otro lado, debido a que

$$\partial f(x) = \{x^* \in \mathbb{R}^n : f(x) + f^*(x^*) - \langle x^*, x\rangle \leq 0\},$$

del Teorema 5.4 se deduce que la multiaplicación ∂f es cerrada y por lo tanto, del Teorema 5.3, $\Gamma : B \rightrightarrows C$, definida por $\Gamma(x) = \partial f(x)$ para todo $x \in B$, es scs. Se deduce, en particular, que ∂f es scs en $\bar{x}$. ■

Corolario 5.1 *Sean $U \subset \mathbb{R}^n$ convexo abierto no vacío y $f : U \to \mathbb{R}$ convexa diferenciable. Entonces ∇f es continua en U.*

Demostración : Sea $\hat{f} : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci tal que $\hat{f}(x) = f(x)$ si $x \in U$; y $+\infty$ si $x \notin \overline{U}$. Entonces $U = \operatorname{int}(\operatorname{dom}(f))$ y $\partial f(x) = \{\nabla f(x)\}$ para todo $x \in U$. Por el Teorema 5.5, ∇f es continua en U. ■

5.4.1 Aplicación al análisis de sensibilidad

Consideremos el siguiente problema de optimización

$$\alpha := \inf_x [f(x) : g_i(x) \leq 0, \ i = 1, \cdots, p; \ Ax + a = 0 \,], \qquad (P)$$

con las siguientes hipótesis :

$(H1)$ Las funciones f y g_i $(i = 1, \cdots, p)$ son convexas sci propias definidas en $\mathbb{R}^n$, $a \in \mathbb{R}^q$ y A una matriz de orden $q \times n$ de rango q.

(H_2) Existe $\hat{x} \in \mathbb{R}^n$ tal que $A\hat{x} + a = 0$, f y g_i finitas y continuas en $\hat{x}$, y $g_i(\hat{x}) < 0$ para $i = 1, \cdots, p$.

(H_3) $\alpha > -\infty$.

Por lo que vimos en el Capítulo 4, el problema dual de (P) es

$$\alpha = \sup_{u^* \geq 0, v^*} \inf_x [f(x) + \sum u_i^* g_i(x) + \langle Ax + a, v^* \rangle \,] \qquad (Q)$$

Denotemos por $\overline{S}$ y $\overline{T}$ a los conjuntos de soluciones óptimas de (P) y (Q), respectivamente. Entonces, $\overline{T}$ es convexo compacto no vacío.

Asociados a los parámetros (de perturbación) $(u, v) \in \mathbb{R}^p \times \mathbb{R}^q$, consideremos el siguiente problema perturbado

$$h(u, v) := \inf_x [f(x) : g_i(x) + u_i \leq 0, \ i = 1, \cdots, p; \ Ax + a + v = 0]. \qquad (P_{u,v})$$

La función h es convexa y $(0, 0) \in \operatorname{int}(\operatorname{dom}(h))$. Se deduce que h es continua en una vecindad $U_0 \times V_0 \subset \mathbb{R}^p \times \mathbb{R}^q$ de $(0, 0)$, y por consiguiente

$$h(u, v) = h^{**}(u, v) = \sup_{u^*, v^*} [\langle u, u^* \rangle + \langle v, v^* \rangle - h^*(u^*, v^*) \,] \ \text{ para todo } (u, v) \in U_0 \times V_0$$

donde,

$$-h^*(u^*,v^*) = \begin{cases} \inf_x[f(x) + \sum u_i^* g_i(x) + \langle Ax+a, v^*\rangle\ ] & \text{si } u^* \geq 0, \\ -\infty & \text{en caso contrario.} \end{cases}$$

Por lo tanto,

$$h(u,v) = \sup_{u^* \geq 0, v^*} \inf_x [f(x) + \sum u_i^*(g_i(x)+u_i) + \langle Ax+a+v, v^*\rangle]. \qquad (Q_{u,v})$$

Denotando por $T(u,v)$ al conjunto de soluciones óptimas de $(Q_{u,v})$, tenemos $T(u,v) = \partial h(u,v)$, lo que implica que la multiaplicación T es scs en la vecindad $U_0 \times V_0$.

Por último, sea $S(u,v) := \{x : \varphi(x,u,v) - h(u,v) \leq 0\}$, donde

$$\varphi(x,u,v) = \begin{cases} f(x) & \text{si } Ax+a+v=0,\ g_i(x)+u_i \leq 0, i=1,\cdots,p, \\ +\infty & \text{en otro caso.} \end{cases}$$

Siendo φ sci (por que f y g_i lo son), la multiaplicación S es cerrada en la vecindad $U_0 \times V_0$ debido a que h es continua en esa vecindad. Eso significa en particular que si $(u_n,v_n) \to (0,0)$ y $x_n \to \bar{x}$, con $x_n \in S(u_n,v_n)$ para todo n, entonces $\bar{x} \in S(0,0) = \overline{S}$.

Asumamos ahora que $\overline{S}$ es compacto no vacío. Mostraremos que la multiaplicación S es scs en una vecindad de $(0,0)$. Por la continuidad de h en $U_0 \times V_0$, existe $r > 0$ tal que $R_1 := [-r,r]^p \subset U_0$, $R_2 := [-r,r]^q \subset V_0$ y $h(u,v) \leq \alpha + 1$ para todo $(u,v) \in R_1 \times R_2$.

Sea

$$\widehat{S} = \left\{ x \in \mathbb{R}^n : \begin{array}{c} f(x) \leq \alpha+1, \\ g_i(x) - r \leq 0,\ i=1,\cdots,p, \\ -r \leq (Ax+a)_j \leq r,\ j=1,\cdots,q. \end{array} \right\}.$$

Este es compacto y $S(u,v) \subset \widehat{S}$ para todo $(u,v) \in R_1 \times R_2$. La compacidad se debe al hecho que

$$\begin{aligned} (\widehat{S})_\infty &= \{x : f(x) \leq \alpha+1\}_\infty \cap \{x : g(x) \leq r\}_\infty \cap \{x : |(Ax+a)_j| \leq r\}_\infty \\ &= \{x : f(x) \leq \alpha\}_\infty \cap \{x : g_i(x) \leq 0\}_\infty \cap \{x : Ax = a\}_\infty \\ &= (\overline{S})_\infty \\ &= \{0\}. \end{aligned}$$

Se deduce por lo tanto, del Teorema 5.3, que la multiaplicación S es scs en $R_1 \times R_2$.

Se tiene $S(u,v) \subset \widehat{S}$ para todo $(u,v) \in R_1 \times R_2$. También $\widehat{S}$ es compacto debido a que $(\widehat{S})_\infty = \{0\}$.

5.4.2 Propiedades genéricas de las funciones convexas

Teorema 5.6 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces f es localmente Lipschitz en* $\operatorname{int}(\operatorname{dom}(f))$.

Demostración : Sea $\bar{x} \in \operatorname{int}(\operatorname{dom}(f))$. En la demostración del Teorema 5.5 vimos que para algún $r > 0$, $B(\bar{x}, r) \subset \operatorname{int}(\operatorname{dom}(f))$ y

$$\|x^*\| \leq k := \frac{m_2 - m_1}{r} \quad \text{para todo } x^* \in \partial f(B(\bar{x}, r)).$$

Sean $x, y \in B(\bar{x}, r)$, $x^* \in \partial f(x)$ y $y^* \in \partial f(y)$. Tenemos

$$f(x)-f(y) \leq \langle x^*, x-y\rangle \leq k\|y-x\| \quad \text{y} \quad f(y)-f(x) \leq \langle y^*, y-x\rangle \leq k\|y-x\|,$$

de donde,

$$|f(y) - f(x)| \leq k\|y - x\|.$$

Por lo tanto f es Lipschitz en $B(\overline{x}, r)$. ■

Un resultado bastante conocido debido a Rademacher [Hans Adolph Rademacher (3 de Abril de 1892, Wandsbeck, ahora Hamburg-Wandsbek – 7 de Febrero de 1969, Haverford, Pennsylvania, USA)$\cdots$], dice que si $f : U \to \mathbb{R}^m$ (U abierto de $\mathbb{R}^n$) es Lipschitz, entonces este es diferenciable en casi todo punto de U; esto es, el conjunto de todos los puntos de U en el que f no es diferenciable, tiene medida de Lebesgue cero. Su demostración es un poco extensa y puede encontrarse en los textos clásicos de análisis.

Cuando la función es considerada convexa, existe una demostración más directa y mucho más simple de este resultado. Para dar la demostración en este último caso, necesitamos algunos resultados previos.

Proposición 5.1 *Sean X y Y dos espacios topológicos, $\varphi : X \times Y \to \mathbb{R}$ scs y $\Gamma : X \rightrightarrows Y$ scs de valor compacto no vacío. Entonces la función $\upsilon(x) := sup_y[\varphi(x,y) : y \in \Gamma(x)]$, es scs.*

Demostración : Sean $x_0 \in X$ y $\lambda > \upsilon(x_0)$. Para cada $y \in \Gamma(x_0)$ se cumple $\varphi(x_0, y) < \lambda$ lo que implica que existen V_y y W_y, vecindades de x_0 e y, respectivamente, tales que $\varphi(x', y') < \lambda$ para todo $(x', y') \in V_y \times W_y$. Siendo $\Gamma(x_0)$ compacto y $\cup_{y\in\Gamma(x_0)} W_y \supset \Gamma(x_0)$, existen $y_1, \cdots, y_p \in \Gamma(x_0)$ tales que $\Gamma(x_0) \subset \cup_{i=1,\cdots,p} W_{y_i}$.

Sean $V = \cap_{i=1}^{p} V_{y_i}$ y $W = \cup_{i=1}^{p} W_{y_i}$. Entonces, V es una vecindad de x_0, y W es un abierto conteniendo a $\Gamma(x_0)$. Por la semicontinuidad superior de Γ en x_0, existe una vecindad U de x_0 tal que $\Gamma(x) \subset W$ para todo $x \in U$.

Por lo tanto, para todo $x \in U \cap V$,

$$\varphi(x, y) < \lambda \text{ para todo } y \in \Gamma(x),$$

de donde, $\upsilon(x) \leq \lambda$. ∎

Corolario 5.2 *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Entonces para $h \in \mathbb{R}^n$ fijo, la función $x \to f'(x, h)$ es scs en* $\operatorname{int}(\operatorname{dom}(f))$.

Demostración : Se deduce inmediatamente de la Proposición 5.1 debido a que $f'(x, h) = \sup_{x^*}[\langle x^*, h\rangle : x^* \in \partial f(x)]$. ∎

Comentario: La función $x \to f'(x, h)$ puede no ser sci. Consideremos por ejemplo $f : \mathbb{R} \to \mathbb{R}$ con $f(x) = |x|$. Entonces

$$f'(x, h) = \begin{cases} h & \text{si } x > 0, \\ |h| & \text{si } x = 0, \\ -h & \text{si } x < 0. \end{cases}$$

Proposición 5.2 *Sean $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia y $x \in$* $\operatorname{dom}(f)$. *Entonces f es diferenciable en x si y solo si f admite todas las derivadas parciales en x.*

Demostración : Asumamos que f admite todas las derivadas parciales en x. Entonces,

$$\frac{\partial f}{\partial x_i}(x) = f'(x, e_i) = -f'(x, -e_i) \text{ para todo } i = 1, \cdots, n.$$

Se deduce que para cada $i = 1, \cdots, n$, la función $\mathbb{R} \ni t \to f(x + te_i)$ es derivable en 0, lo que implica que existe $t_i > 0$ tal que $x \pm t_i e_i \in \operatorname{dom}(f)$. Así, $\operatorname{co}(\{x \pm t_i e_i\}_{i=1}^n) \subset \operatorname{dom}(f)$ y por lo tanto, $x \in \operatorname{int}(\operatorname{dom}(f))$. Eso implica en particular que $\partial f(x) \neq \emptyset$. Mostraremos que $\partial f(x)$ tiene exactamente un elemento. En efecto, sean $x^*, y^* \in \partial f(x)$. Tenemos,

$$f'(x, e_i) = \sup_{z^* \in \partial f(x)} \langle z^*, e_i \rangle \geq y_i^* \quad \text{y} \quad f'(x, -e_i) = \sup_{z^* \in \partial f(x)} \langle z^*, -e_i \rangle \geq -x_i^*,$$

de donde,

$$0 = f'(x, e_i) + f'(x, -e_i) \geq y_i^* - x_i^*.$$

Análogamente,

$$0 = f'(x, e_i) + f'(x, -e_i) \geq x_i^* - y_i^*.$$

Por lo tanto, $|x_i^* - y_i^*| = 0$ para todo $i = 1, \cdots, n$. ■

Teorema 5.7 *Sea* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *convexa sci propia. Entonces este es diferenciable en casi todo punto de* $\operatorname{int}(\operatorname{dom}(f))$.

Demostración : El caso particular $n = 1$ es simple. En efecto, si f es convexa en un interval $I \subset \mathbb{R}$, entonces podemos dividir I en dos subintervalos en el que f es creciente en uno y decreciente en el otro. De otro lado sabemos (ver cualquier texto de análisis) que las funciones monótonas de una variable real son diferenciables en casi todo punto.

Asumamos ahora $n > 1$. Denotemos $C = \operatorname{int}(\operatorname{dom}(f))$,

$$D_i := \left\{ x \in C : \frac{\partial f}{\partial x_i}(x) \text{ existe} \right\}, \quad \text{para } i = 1, \cdots, n,$$

y

$$D := \bigcap_{i=1}^n D_i = \{x \in C : \nabla f(x) \text{ existe}\}.$$

Puesto que la aplicación $x \to f'(x, e_i) + f'(x, -e_i)$ es scs, esta es medible y por lo tanto, D_i es medible. Se sigue que D también es medible.

Para mostrar que $\operatorname{med}(D^c) = 0$, mostraremos que $\operatorname{med}(D_i^c) = 0$ para todo $i = 1, \cdots, n$.

Mostraremos para $i = n$ (para los otros se usan los mismos argumentos). Sean $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_n) \in C$ y $\epsilon > 0$ tal que

$$P = \widehat{P} \times [\bar{x}_n, \bar{x}_n + \epsilon] \subset C \quad \text{ con } \widehat{P} = \prod_{i=1}^{n-1} [\bar{x}_i, \bar{x}_i + \epsilon].$$

Tenemos

$$\operatorname{med}(P \cap D_n^c) = \int_P 1_{D_n^c}(x)dx = \int_{\widehat{P}} \left[\int_{\bar{x}_n}^{\bar{x}_n + \epsilon} 1_{D_n^c}(x) dx_n \right] dx_1 \cdots dx_n,$$

donde la última igualdad se debe al Teorema de Fubini. Denotemos $x = (y, x_n)$ con $y = (x_1, \cdots, x_{n-1})$. Fijado y, la integral $\int_{\bar{x}_n}^{\bar{x}_n+\epsilon} 1_{D_n^c}(x)dx_n$ es la medida del conjunto $N(y) := [\bar{x}_n, \bar{x}_n + \epsilon] \setminus D(y)$ con

$$D(y) = \{x_n \in [\bar{x}_n, \bar{x}_n + \epsilon] : (y, x_n) \in D_n\}.$$

Por definición, $D(y)$ es el conjunto de puntos donde la aplicación convexa $[\bar{x}_n, \bar{x}_n + \epsilon] \ni t \to f(y, t)$ es derivable. Luego, $N(y)$ tiene medida cero, y por lo tanto, $\int_{\bar{x}_n}^{\bar{x}_n+\epsilon} 1_{D_n^c}(x)dx_n = 0$. De este, $\operatorname{med}(P \cap D_n^c) = 0$. ■

Teorema 5.8 *Sean* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *convexa sci propia,* $\bar{x} \in \operatorname{int}(\operatorname{dom}(f))$ *y* $\Im = \{x \in \operatorname{dom}(f) : \nabla f(x)$ *existe* $\}$*. Entonces* $\partial f(\overline{x}) = \overline{co}(S)$ *con*

$$S = \{e : \exists \{x_n\} \text{ en } \Im \text{ tal que } x_n \to \overline{x} \text{ y } \nabla f(x_n) \to e\}.$$

Demostración : Por el Teorema 5.7, el conjunto S es no vacío; y por el Teorema 5.6, es acotado. Se deduce que $\overline{\text{co}}(S)$ es convexo compacto no vacío.

De otro lado, por la semicontinuidad superior de ∂f en $\bar{x}$, el conjunto S está contenido en $\partial f(\bar{x})$ y por lo tanto, $\overline{\text{co}}(S) \subset \partial f(\bar{x})$.

Para mostrar la inclusión contraria, asumamos por contradicción que $x^* \notin \overline{\text{co}}(S)$ para algún $x^* \in \partial f(\bar{x})$. Por el teorema de separación fuerte, existen $d \in \mathbb{R}^n$, con $\|d\| = 1$, y $\alpha > 0$ tales que

$$\langle l, d \rangle + \alpha \leq \langle x^*, d \rangle \quad \text{para todo } l \in \overline{\text{co}}(S),$$

de donde,

$$\langle l, d \rangle + \alpha \leq \langle x^*, d \rangle \leq f'(\overline{x}, d) \quad \text{para todo } l \in S. \tag{5.2}$$

Consideremos la función convexa θ definida por $\theta(t) := f(\bar{x} + td)$, para todo $t \in \mathbb{R}$. Debido a que θ es derivable en casi todo punto, entonces para cada $k \in \mathbb{N}$, existe $t_k \in]0, \frac{1}{k}[$ en el que θ es derivable y $\bar{x} + t_k d \in \operatorname{int}(\operatorname{dom}(f))$. Se cumple $\theta'(t_k) \geq \theta'_+(0) \geq \langle x^*, d\rangle$.

Sea $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}$ definida por $\Gamma(x) := \{\langle x^*, d\rangle : x^* \in \partial f(x)\}$. Este es scs en $\operatorname{int}(\operatorname{dom}(f))$ debido a que ∂f lo es. Puesto que $\Gamma(\bar{x} + t_k d) = \{\theta'(t_k)\}$, la semicontinuidad superior de Γ implica que existe una vecindad U de $\bar{x} + t_k d$ tal que

$$\Gamma(x) \subset \left[\theta'(t_k) - \frac{\alpha}{10}, \theta'(t_k) + \frac{\alpha}{10}\right] \quad \text{para todo } x \in U.$$

Sea $x_k \in U \cap \Im$ con $\|x_k - (\bar{x} + t_k d)\| \leq \frac{1}{k}$. Se cumple,

$$\langle \nabla f(x_k), d\rangle \geq \theta'(t_k) - \frac{\alpha}{10} \geq \langle x^*, d\rangle - \frac{\alpha}{10}.$$

Haciendo $k \to \infty$, se obtiene $\langle \hat{l}, d\rangle \geq \langle x^*, d\rangle - \frac{\alpha}{10}$, para algún valor de adherencia $\hat{l}$ de $\{\nabla f(x_k)\}$. Este es una contradicción con (5.2) debido a que $\hat{l} \in S$. ■

5.5 $\epsilon-$subdiferencial

Sean $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa propia, $x \in \operatorname{dom}(f)$ y $\epsilon > 0$. Se dice que x^* es un $\epsilon-$subdiferencial de f en x, si

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

Denotaremos por $\partial_\epsilon f(x)$ al conjunto de todos los $\epsilon-$subdiferenciales de f en x.

Este es convexo y cerrado porque la aplicación $x^* \to f^*(x^*) - \langle x, x^*\rangle$ es convexa sci y

$$\partial_\epsilon f(x) = \{x^* : f^*(x^*) - \langle x, x^*\rangle \leq \epsilon - f(x)\}.$$

Otra propiedad de los ϵ- subdiferenciales es que $\partial_{\epsilon_1} f(x) \subset \partial_{\epsilon_2} f(x)$ desde que $0 < \epsilon_1 < \epsilon_2$. Este implica en particular que

$$\bigcap_{\epsilon > 0} \partial_\epsilon f(x) = \partial f(x).$$

De la definición de $f^*(x^*)$ obtenemos

$$x^* \in \partial_\epsilon f(x) \text{ si y solo si } f(y) \geq f(x) + \langle x^*, y - x\rangle - \epsilon \text{ para todo } y \in \mathbb{R}^n.$$

Teorema 5.9 *Si $\partial f(x)$ es compacto no vacío, entonces $\partial_\epsilon f(x)$ también es compacto no vacío para todo $\epsilon > 0$.*

Demostración : Sea $g : X^* \to \mathbb{R}$ definida por $g(x^*) := f^*(x^*) + f(x) - \langle x, x^*\rangle$. Esta es convexa sci y $S_\epsilon(g) = \partial_\epsilon f(x)$. Se deduce que

$$[S_\epsilon(g)]_\infty = [S_0(g)]_\infty = \{0\}$$

debido a que $S_0(g) = \partial f(x)$ es compacto no vacío. ■

Teorema 5.10 *Si f es convexa sci propia, entonces $\partial_\epsilon f(x) \neq \emptyset$ para todo $x \in \operatorname{dom}(f)$ y para todo $\epsilon > 0$.*

Demostración : Es consecuencia inmediata de la definición del supremo y del hecho que $f(x) = \sup_{x^* \in X^*}[\langle x, x^*\rangle - f^*(x^*)]$. ■

Capítulo 6

Análisis convexo matricial

6.1 El espacio vectorial de matrices rectángulares

En todo este capítulo usaremos las siguientes notaciones :

— $\mathcal{M}_{np}$ es el conjunto de todas las matrices reales de orden $n \times p$,

— $\mathcal{S}_n$ es el conjunto de todas las matrices de $\mathcal{M}_{nn}$ simétricas,

— $\mathcal{S}_n^+$ es el conjunto de todas las matrices de $\mathcal{S}_n$ semidefinida positiva,

— $\mathcal{S}_n^{++}$ es el conjunto de todas las matrices de $\mathcal{S}_n$ definida positiva,

— Ω_n es el conjunto de todas las matrices ortonormales de orden $n \times n$, es decir, las matrices $P \in \mathcal{M}_{nn}$ tal que $PP^t = I_n$.

— I_n denotará la matriz identidad de orden $n \times n$.

La traza de una matriz $A \in \mathcal{M}_{nn}$ se define como

$$\operatorname{tr}(A) := \sum\nolimits_{i=1}^{n} a_{ii}.$$

Es claro que la función traza $\operatorname{tr}(\cdot)$ es una función lineal en $\mathcal{M}_{nn}$. Note que $\operatorname{tr}(A) = \operatorname{tr}(A^t)$.

Para $A, B \in \mathcal{M}_{np}$, definimos

$$\langle A, B \rangle := \operatorname{tr}(AB^t) = \sum_{\substack{i=1,\cdots,n \\ j=1,\cdots,p}} a_{ij} b_{ij}. \tag{6.1}$$

Se deduce fácilment de la definición :

— $\langle A, B\rangle = \langle B, A\rangle$ para todo $A, B \in \mathcal{M}_{np}$,
— $\langle A, A\rangle \geq 0$ para todo $A \in \mathcal{M}_{np}$ y $\langle A, A\rangle = 0$ si y solo si $A = 0$,
— $\langle\cdot,\cdot\rangle$ es bilineal en $\mathcal{M}_{np} \times \mathcal{M}_{np}$.

Por lo tanto, $\langle\cdot,\cdot\rangle$ define un producto interno en el espacio vectorial $\mathcal{M}_{np}$. La norma asociada se define como

$$|||A||| = \sqrt{\langle A, A\rangle} = \sqrt{\operatorname{tr}(AA^t)} = \left(\sum_{i=1,\cdots,n}\sum_{j=1,\cdots,p} a_{ij}^2\right)^{1/2} \quad \forall\, A \in \mathcal{M}_{np}. \tag{6.2}$$

Note que $\operatorname{tr}(AB^t) = \operatorname{tr}(A^tB)$. Siendo AB^t una matriz de orden $n \times n$ y A^tB una matriz de orden $p\times p$, el producto interno $\langle A, B\rangle$ puede interpretarse como el producto escalar de A y B en el espacio $\mathcal{M}_{np}$ o como el producto escalar de A^t y B^t en el espacio $\mathcal{M}_{pn}$. Por lo tanto, escribiremos $|||A||| = |||A^t|||$ teniendo en cuenta que $|||\cdot|||$ representa, según el caso, la norma en $\mathcal{M}_{np}$ o en $\mathcal{M}_{pn}$.

La traza satisface también las siguientes dos propiedades muy importantes :

— Para todo $A \in \mathcal{M}_{np}$, $B \in \mathcal{M}_{pq}$ y $A \in \mathcal{M}_{qn}$,

$$\operatorname{tr}(\mathrm{ABC}) = \operatorname{tr}(\mathrm{CAB}) = \operatorname{tr}(\mathrm{BCA}). \tag{6.3}$$

— Para todo $A, B \in \mathcal{M}_{np}$,

$$\operatorname{tr}(\mathrm{AB^t}) = \operatorname{tr}(\mathrm{BA^t}) = \operatorname{tr}(\mathrm{B^tA}) = \operatorname{tr}(\mathrm{A^tB}). \tag{6.4}$$

Recordemos dos propiedades más referente a las matrices simétricas. Para $A \in \mathcal{S}_n$, denotemos por $\det(A)$ al determinante de la matriz A.

Proposición 6.1 *Sea $A \in \mathcal{S}_n$. Entonces*
— $\det(A) = \prod_{i=1}^n \lambda_i(A)$,
— $\operatorname{tr}(A) = \sum_{i=1}^n \lambda_i(A)$,
donde los $\lambda_i(A)$ son los valores propios de A.

Demostración : Existen $D \in \mathcal{S}_n$ diagonal y $P \in \Omega_n$ tal que $A = P^tDP$. Se deducen

$$\det(A) = \det(P^tDP) = \det(PP^tD) = \det(D),$$

$$\mathrm{tr}\,(A) = \mathrm{tr}\,(\mathrm{P^tDP}) = \mathrm{tr}\,(\mathrm{PP^tD}) = \mathrm{tr}\,(\mathrm{D}).$$

Las dos expresiones de la proposición se obtienen a partir de estas últimas debido a que D es la matriz diagonal donde los elementos de la diagonal son los valores propios de A. ■

Volviendo a la definición de la norma en (6.2), $|||A|||$ no es otra cosa que la norma euclidiana de A considerada como un vector de $\mathbb{R}^{np}$.

Si la matriz A es vista como una aplicación lineal de $\mathbb{R}^p$ en $\mathbb{R}^n$, definimos su norma como

$$n^{\mathrm{o}}\,(A) := \max\,[\,\|Ax\|\;:\;\|x\| = 1\,].$$

En la siguiente proposición mostramos algunas relaciones entre las normas $|||\cdot|||$ and $n^{\mathrm{o}}\,(\cdot)$, donde las normas consideradas en $\mathbb{R}^p$ y $\mathbb{R}^n$ es la euclidiana, denotadas para ambas por $\|\cdot\|$.

Proposición 6.2 *Sea $A \in \mathcal{M}_{np}$. Se cumplen*

i) $n^o\,(A) = n^o\,(A^t) = \sqrt{\lambda_{max}(AA^t)} = \sqrt{\lambda_{max}(AA^t)}$, *donde* $\lambda_{max}(\cdot)$ *asigna el más grande valor propio de la matriz.*

ii) $[n^o\,(A)]^2 \leq |||A|||^2 \leq \min(n,p)\,n^{o\,2}(A)$.

Demostración : Por definición,

$$n^{\mathrm{o}}\,(A) = \sup_x\,[\,\|Ax\|\;:\;\|x\| = 1\,] = \sup_{x,y}\,[\,y^tAx\;:\;\|x\| = \|y\| = 1\,] = n^{\mathrm{o}}\,(A^t).$$

Sabemos que existen $D \in S_n^+$ diagonal y $P \in \Omega_n$ tales que $A^tA = P^tD^2P$, de donde, denotando $y = Px$, tenemos

$$[n^{\mathrm{o}}\,(A)]^2 = \sup_x\,[\,\|Ax\|^2\;:\;\|x\| = 1\,] = \sup_y\,[\,\|Dy\|^2\;:\;\|y\| = 1\,] = \lambda^2_{max}(D).$$

El resto de la demostración se deducen de las siguientes relaciones :

— $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^tAx, x\rangle = \mathrm{tr}\,(x^tA^tAx)$,

— $\mathrm{tr}\,(x^tA^tAx) = \mathrm{tr}\,(xx^tA^tA) = \langle xx^t, A^tA\rangle \leq |||xx^t|||\,|||A^tA||| = |||A^tA|||$,

— $\mathrm{tr}\,(A^tA) = \mathrm{tr}\,(P^tD^2P) = \mathrm{tr}\,(PP^tD^2) = \sum_i \lambda_i(A^tA) \leq p\,\lambda_{max}(A^tA)$,

— $\mathrm{tr}\,(AA^t) \leq n\,\lambda_{max}(AA^t)$.

Las relaciones de la proposición quedan establecidas. ■

Veamos otras propiedades referente a la función determinante. Asociada a esta función, consideremos las siguientes dos funciones $f, g : S_n^{++} \to \mathbb{R}$ definidas por

$$g(A) = \ln(\det(A)) \;\text{ y }\; f(A) = [\det(A)]^{\frac{1}{n}} \;\text{ para todo } A \in S_{++}^n.$$

Teorema 6.1 *La función determinante* det *y las funciones f y g son infinitamente diferenciables en el cono S_n^{++}. En todo $A \in S_n^{++}$ se cumplen*

$$\nabla \det(A) = \det(A)A^{-1}, \quad \nabla g(A) = A^{-1} \quad y \quad \nabla f(A) = \frac{1}{n}[\det(A)]^{\frac{1}{n}} A^{-1}.$$

Además, f y g son cóncavas en S_n^{++}, siendo además g fuertemente cóncava en todo convexo compacto contenido en S_n^{++}.

Demostración : Sea $A \in S_n^{++}$. Existe una matriz invertible L tal que $A = LL^t$. Tal matriz L no es única, considere por ejemplo $L = PDP^t$, siendo D una matriz diagonal definida positiva y $P \in \Omega_n$ tal que $A = PD^2P^t$ (por lo tanto $L = L^t$ y $A = L^2$). Para esta descomposición se necesita determinar los valores y vectores propios de A. Otra matriz L más fácil de obtener es considerando la matriz triangular inferior obtenida a partir de la descomposición de Cholewsky.

Sean $H \in \mathcal{S}_n \setminus \{0\}$. Se cumple

$$\det(A + H) = \det(L(I + R)L^t) = \det(A) \det(I + R),$$

con $R = L^{-1}H[L^{-1}]^t$. Debido a que R es simétrica, existe una matriz diagonal Δ y una matriz $Q \in \Omega_n$ tales que $R = Q\Delta Q^t$. Por lo tanto

$$\det(A + H) = \det(A) \det(Q(I + \Delta)Q^t) = \det(A) \prod_{i=1}^{n}(1 + \delta_i).$$

Note que

$$\prod_{i=1}^{n}(1 + \delta_i) = 1 + \sum_{i=1}^{n} \delta_i + |||\Delta||| \, \varepsilon(|||\Delta|||),$$

donde $\varepsilon(t) \to 0$ desde que $t \to 0$. Ahora, debido a que

$$\mathrm{tr}\,(\Delta) = \mathrm{tr}\,(\mathrm{R}) = \mathrm{tr}\,([\mathrm{L}^{-1}]^{\mathrm{t}}\mathrm{L}^{-1}\mathrm{H}) = \mathrm{tr}\,(\mathrm{A}^{-1}\mathrm{H}) = \langle \mathrm{A}^{-1}, \mathrm{H}\rangle,$$

la expresión anterior de $\det(A+H)$ se puede reescribir como

$$\det(A+H) = \det(A) + \det(A)\langle A^{-1}, H\rangle + |||H|||\,\varepsilon(|||H|||).$$

Se deduce que las funciones det y f son diferenciables en todo $A \in \mathcal{S}_n^{++}$, con

$$\nabla \det(A) = \det(A)A^{-1} \quad \text{y} \quad \nabla f(A) = \frac{1}{n}[\det(A)]^{\frac{1}{n}} A^{-1}.$$

Concerniente a la función g, sea H tal que $A+H \in \mathcal{S}_n^{++}$. Tenemos

$$g(A+H) - g(A) = \sum_{i=1}^{n} \ln(1+\delta_i) = \langle A^{-1}, H\rangle + |||H|||\,\varepsilon(|||H|||)$$

y por lo tanto $\nabla g(A) = A^{-1}$. Se deduce que g es infinitamente diferenciable en A. Esta regularidad también es válida para la función $A \to \det(A)$ debido a que

$$\det(A) = \exp(g(A)).$$

Mostremos la concavidad de la función g. Sea $H \in \mathcal{S}_n$, con $H \neq 0$, tal que $A+H \in \mathcal{S}_n^{++}$. Sea

$$I = \{t \in \mathrm{I\!R} : A + tH \in \mathcal{S}_n{++}\} \tag{6.5}$$

y consideremos la función $\mu : I \to \mathrm{I\!R}$ definida por

$$\mu(t) := g(A+tH) \quad \text{para todo } t \in I.$$

Note que I es un intervalo abierto conteniendo a $[0,1]$. Tenemos

$$\mu(t) = \mu(0) + \sum_{i=1}^{n} \ln(1+t\delta_i),$$

de donde,

$$\mu'(t) = \sum_{i=1}^{n} \frac{\delta_i}{(1+t\delta_i)} \quad \text{y} \quad \mu''(t) = -\sum_{i=1}^{n} \frac{\delta_i^2}{(1+t\delta_i)^2}, \tag{6.6}$$

siendo los δ_i valores propios de $R = L^{-1}H[L^{-1}]^t$. Note que $\mu''(t) < 0$ debido a que $R \neq 0$.

De otro lado, por Taylor, existe $t \in (0,1)$ tal que

$$\mu(1) = \mu(0) + \mu'(0) + \frac{1}{2}\mu''(t),$$

y por lo tanto,

$$\mu(1) < \mu(0) + \mu'(0). \tag{6.7}$$

Por la expresión $\mu'(t)$ en (6.6), tenemos

$$\mu'(0) = \sum_{i=1}^{n} \delta_i = \mathrm{tr}\,(R) = \langle A^{-1}, H\rangle,$$

y por lo tanto de (6.7), denotando $B = A + H$, tenemos

$$g(B) < g(A) + \langle A^{-1}, B - A\rangle \ \text{ para todo } A,\, B \in \mathcal{S}_n^{++} \ \text{ con } \ A \neq B,$$

lo que significa que la función g es estrictamente cóncava en $\mathcal{S}_n^{++}$.

Para mostrar la fuerte concavidad local de g, consideremos $V \subset \mathcal{S}_n^{++}$, convexo compacto conteniendo a A. Sean $B \in V$ y $H = B - A$. Debido a que los $1 + t\delta_i$ son los valores propios de la matriz $I + tL^{-1}H[L^{-1}]^t$, existen $\alpha > 0$ y $\beta > 0$ tales que

$$\alpha\, |||B - A|||^2 \leq \sum_{i=1}^{n} \frac{\delta_i^2}{(1 + t\delta_i)^2} \leq \beta\, |||B - A|||^2 \ \text{ para todo } t \in [0,1].$$

Por lo tanto,

$$g(B) \leq g(A) + \langle A^{-1}, B - A\rangle - \frac{\alpha}{2}\, |||B - A|||^2 \ \text{ para todo } A, B \in V,$$

lo que significa que g es fuertemente cóncava en V.

Mostremos ahora la concavidad de f. Sea I el intervalo definido en (6.5) y sea $\nu : I \to \mathbb{R}$ la función definida por

$$\nu(t) := f(A + tH) \ \text{ para todo } t \in I.$$

Tenemos

$$\nu(t) = \nu(0) \prod_{i=1}^{n} (1 + t\delta_i)^{\frac{1}{n}},$$

de donde, denotando

$$v_i = \frac{\delta_i}{(1 + t\delta_i)}, \quad \bar{v} = \frac{1}{n} \sum_{i=1}^{n} v_i \;\; \text{y} \;\; \sigma_v^2 = \frac{1}{n} \sum_{i=1}^{n} (v_i - \bar{v})^2 = \left(\frac{1}{n} \sum_{i=1}^{n} v_i^2 \right) - \bar{v}^2,$$

obtenemos

$$\nu'(t) = \nu(0) \frac{\nu(t)}{n} \sum_{i=1}^{n} \frac{\delta_i}{(1 + t\delta_i)} \quad \text{y} \quad \nu''(t) = -\nu(0) \sigma_v^2 \, \nu(t) \leq 0.$$

Por Taylor, existe $t \in (0, 1)$ tal que

$$\nu(1) = \nu(0) + \nu'(0) + \frac{1}{2} \nu''(t),$$

de donde se deduce la concavidad de f en $\mathcal{S}_n^{++}$. ■

6.2 Rango de una matriz

Sea $A \in \mathcal{M}_{np}$. Su rango, denotado por rang (A), es la dimensión del subespacio $A(\mathbb{R}^p)$. Por definición,

$$\text{rang}\,(A) = \text{rang}\,(A^t) = \text{rang}\,(AA^t) = \text{rang}\,(A^t A) \leq \min(n, p).$$

Mostremos ahora que la función rang es semicontinua inferior. Para ello, notemos primeramente que para $P \in \Omega_n$ y $Q \in \Omega_p$, se cumple

$$\text{rang}\,(A) = \text{rang}\,(PA) = \text{rang}\,(AQ) = \text{rang}\,(PAQ).$$

Siendo AA^t simétrica semidefinida positiva, existen $D \in \mathcal{S}_{++}^r$ diagonal y $P \in \Omega_n$ tales que

$$PAA^t P^t = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}.$$

Teniendo en cuenta que

$$\mathrm{rang}\,(A+H) = \mathrm{rang}\,(P(A+H)(A+H)^t P^t),$$

vemos que para H suficientemente pequeño, se cumple $\mathrm{rang}\,(A+H) \geq \mathrm{rang}\,(A)$. Por lo tanto se deduce que la function rang es sci. Esta no es scs debido a que no es continua.

Es claro que $\mathrm{rang}\,(X) = n$ si X es simétrica definida positiva.

En la siguiente proposición estudiamos a la función rang en el conjunto de las matrices simétricas semidefinidas positivas.

Proposición 6.3 *Consideremos la función* $f_\alpha : \mathcal{S}_n^+ \to \mathrm{I\!R}$, *con* $0 < \alpha < 1$, *definida por*

$$f_\alpha(X) = \frac{1}{\ln(\alpha)}\ \ln(\,\det(\alpha I_n + X)\,).$$

Esta función es convexa. Además,

$$\lim_{\alpha \downarrow 0} f_\alpha(X) = n - \mathrm{rang}\,(\mathrm{X}).$$

Se deduce que la función rang *es cóncava en* $\mathcal{S}_n^+$.

Demostración : Debido a que $\ln(\alpha) < 0$, por el Teorema 6.1 la función f_α es convexa. Sea $r = \mathrm{rang}\,(\mathrm{X})$, entonces existen $D \in \mathcal{S}_n^{++}$ diagonal y $P \in \Omega_n$ tales que

$$X = P \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} P^t,$$

de donde

$$\det(\alpha I_n + X) = \alpha^{n-r} \det(\alpha I_r + D),$$

$$f_\alpha(X) = n - r + \frac{1}{\ln(\alpha)} \sum_{i=1}^{r} \ln(\, d_i + \alpha\,),$$

siendo d_i $(i = 1, \cdots, r)$ elemento de la diagonal de D. Se deduce que la función $n - r$ es el límite puntual, cuando $\alpha \to 0^+$, de las funciones convexas f_α en el cono convexo $\mathcal{S}_n^+$ (esta función límite alcanza su valor mínimo que vale cero en todo $\mathcal{S}_n^{++}$). Por lo tanto, la función rang es cóncava en $\mathcal{S}_n^+$. ■

6.3 Dualidad

Recordemos que el cono positivo dual de un conjunto $K \subset S_n$ es

$$K^+ := \{X^* \in S_n \,:\, \langle X^*, X\rangle \geq 0 \ \text{ para todo } X \in K\}.$$

Proposición 6.4 *El cono $\mathcal{S}_n^+$ coincide con su cono positivo dual.*

Demostración : Sea $X \in \mathcal{S}_n^+$. Mostraremos que $\langle Y, X\rangle \geq 0$ para todo $Y \in \mathcal{S}_n^+$. Existen $v_1, \cdots, v_n \in \mathbb{R}^n$ tal que $X = \sum_{i=1}^n v_i v_i^t$, de donde

$$\langle X, Y\rangle = \operatorname{tr}(\sum_{i=1}^n v_i v_i^t Y) = \sum_{i=1}^n \operatorname{tr}(v_i^t Y v_i) = \sum_{i=1}^n v_i^t Y v_i \geq 0.$$

Por lo tanto $X \in (\mathcal{S}_n^+)^+$.

Recíprocamente, consideremos $X \in \mathcal{S}_n$ y asumamos $X \notin \mathcal{S}_n^+$. Entonces existe $v \in \mathbb{R}^n$ tal que

$$0 > v^t X v = \operatorname{tr}(\mathrm{v^tXv}) = \operatorname{tr}(\mathrm{vv^tX}) = \langle \mathrm{vv^t}, \mathrm{X}\rangle.$$

Por lo tanto $X \notin (\mathcal{S}_n^+)^+$ debido a que $vv^t \in \mathcal{S}_n^+$. ■

Es claro que si $0 \neq X \in \mathcal{S}_n^+$, entonces $\langle X, Y\rangle > 0$ para todo $Y \in \mathcal{S}_n^{++}$.

6.3.1 Dualidad en programación lineal semi-definida

La programación lineal semi-definida es bastante parecida a la programación lineal clásica desde su formulación y sus propiedades. Con el deseo de poner en evidencia estas similitudes, denotaremos por $X \succeq 0$ si $X \in \mathcal{S}_n^+$, y por $X \succ 0$ si $X \in \mathcal{S}_n^{++}$.

Dadas las matrices $C, A_1, \cdots, A_p \in \mathcal{S}_n^+$ y el vector $b \in \mathbb{R}^p$, consideremos el problema, llamado problema primal,

$$\alpha := \inf_{X \succeq 0} [\langle C, X\rangle : \langle A_i, X\rangle \leq b_i, \ i = 1, 2, \cdots, p]. \qquad (P)$$

Con el fin de definir el problema dual de (P), consideremos la función de perturbación $\varphi : \mathcal{S}_n \times \mathbb{R}^p \to \overline{\mathbb{R}}$ definida por

$$\varphi(X, u) := \begin{cases} \langle C, X\rangle & \text{si } \ X \succeq 0, \ \langle A_i, X\rangle + u_i \leq b_i \ \text{ para todo } i, \\ +\infty & \text{en caso contrario,} \end{cases}$$

y su función marginal asociada $h : \mathbb{R}^p \to \overline{\mathbb{R}}$ definida por

$$h(u) := \inf_X \varphi(X, u).$$

Aplicando la conjugada de Fenchel, tenemos

$$\varphi^*(X^*, u^*) = \begin{cases} \langle b, u^* \rangle & \text{si } u^* \geq 0,\ C + \sum_{i=1}^p u_i^* A_i \succeq X^*, \\ +\infty & \text{en caso contrario,} \end{cases}$$

y

$$h^*(u^*) = \varphi^*(0, u^*) = \begin{cases} \langle b, u^* \rangle & \text{si } u^* \geq 0,\ C + \sum_{i=1}^p u_i^* A_i \succeq 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Por razones de simetría, consideremos la función $k : \mathcal{S}_n \to \overline{\mathbb{R}}$ definida por

$$k(X^*) := \inf_{u^*} \varphi^*(X^*, u^*).$$

Por definición, el problema dual del problema (P) es

$$\beta := -h^{**}(0) = \inf_u h^*(u^*) = \inf_{u^* \geq 0} [\, \langle b, u^* \rangle : C + \sum_{i=1}^p u_i^* A_i \succeq 0 \,]. \qquad (D)$$

Se cumple la siguiente desigualdad, conocida como salto de dualidad :

$$-\beta = h^{**}(0) \leq h(0) = \alpha.$$

Se puede verificar fácilmente que la función φ es convexa sci propia en $\mathcal{S}_n \times \mathbb{R}^p$. Tenemos, $\varphi = \varphi^{**}$ y por lo tanto

$$\varphi(X, 0) = \varphi^{**}(X, 0) = k^*(X) \ \text{ para todo } X \in \mathcal{S}_n,$$

deduciendose nuevamente la desigualdad de dualidad

$$-\alpha = k^{**}(0) \leq k(0) = \beta.$$

Del argumento anterior podemos afirmar que la dualidad entre los dos problemas de optimización considerados son completamente simétricos en el sentido que *el problema dual del problema dual es el problema primal*. Tenemos

$$\inf_{u^*\geq 0} [\, \langle b, u^* \rangle : \sum_{i=1}^{p} u_i^* A_i \succeq -C \,] \geq \sup_{X \succeq 0} [\, \langle -C, X \rangle : \langle A_i, X \rangle \leq b_i \text{ para todo } i \,]$$

que es análoga a la desigualdad correspondiente al problema de programación lineal clásico.

Los conjuntos de soluciones óptimas (eventualmente vacíos) de (P) y (D) con convexos cerrados.

Condiciones de optimalidad

a) **Condición de complementaridad :** Si $\overline{X}$ y $\bar{u}^*$ son respectivamente admisibles de los problemas (P) y (D), entonces estos son soluciones óptimas de estos problemas si y solo si

$$\langle C, \overline{X} \rangle = \langle b, \bar{u}^* \rangle \quad \text{y} \quad \bar{u}_i^* (\langle A_i, \overline{X} \,\rangle - b_i) = 0 \;\; \text{para todo } i.$$

b) **Condición suficiente y salto de dualidad cero :** Consideremos las siguientes condiciones de factibilidad :

$$\exists\, \widehat{X} \succeq 0 \text{ tal que } \langle A_i, \widehat{X} \,\rangle - b_i \leq 0 \text{ para todo } i, \tag{R_p}$$

$$\exists\, \widehat{y} \geq 0 \text{ tal que } C + \sum_{i=1}^{p} \widehat{y}_i A_i \succeq 0, \tag{R_d}$$

y las condiciones de factibilidad estricta :

$$\exists\, \widetilde{X} \succeq 0 \text{ tal que } \langle A_i, \widetilde{X} \,\rangle - b_i < 0 \text{ para todo } i. \tag{S_p}$$

$$\exists\, \widetilde{y} \in \mathbb{R}_+^n \text{ tal que } C + \sum_{i=1}^{p} \widetilde{y}_i A_i \succ 0. \tag{S_d}$$

Es claro que $\alpha < +\infty$ desde que (R_p) se cumple; y $\beta < +\infty$ desde que (R_d) se cumple.

i) **Asumamos que la condición** (R_d) **se cumple** : Entonces la condición (S_p) implica que 0 pertenece al interior del dominio de la función convexa h. Por lo tanto,

$$-\infty < -\beta = h^{**}(0) = h(0) = \alpha < +\infty.$$

Además el conjunto $\partial h(0)$ es convexo compacto no vacío. Este es exactamente el conjunto de las soluciones óptimas del problema (D). Recíprocamente, si el conjunto de las soluciones óptimas del problema (D) es compacto no vacío, entonces la condición (S_p) se cumple.

ii) **Asumamos que la condición** (R_p) **se cumple** : Entonces la condición (S_d) implica que 0 pertenece al interior del dominio de la función convexa k. Por lo tanto

$$-\infty < -\alpha = k^{**}(0) = k(0) = \beta < +\infty.$$

Además el conjunto $\partial k(0)$ es convexo compacto no vacío. Este es exactamente el conjunto de las soluciones óptimas del problema (P). Recíprocamente, si el conjunto de las soluciones óptimas del problema (P) es compacto no vacío, entonces la condición (S_d) se cumple.

Contrariamente a la dualidad en programación lineal clásica, el hecho que los problemas (P) y (D) sean factibles, no implican la existencia de solución de ambos problemas, como se muestra en el siguiente ejemplo. Es más, en este ejemplo, el salto de dualidad es cero.

Ejemplo 6.1 *Consideremos el siguiente ejemplo con $n = 2$ y $p = 2$:*

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Tenemos

$$\alpha = \inf_X \left[x_{11} : \begin{array}{c} x_{11}x_{22} - 1 \geq 0 \\ x_{12} = x_{21} = 1 \\ x_{11} \geq 0, \quad x_{22} \geq 0 \end{array} \right] = 0 \qquad (P)$$

y

$$\begin{aligned}\beta &= \inf_{y\geq 0}\left[2y_1-2y_2 \,:\, \begin{pmatrix} 1 & y_1-y_2 \\ y_1-y_2 & 0\end{pmatrix} \succeq 0\right] \qquad (D)\\ &= \inf_{y\geq 0}\,[\,2y_1-2y_2 \,:\, y_1-y_2=0\,]=0.\end{aligned}$$

Las condiciones R_p y R_d se cumplen y no existe salto de dualidad. Además, el conjunto de soluciones óptimas del problema (D) coincide con el conjunto de todos sus elementos factibles, que no es acotado. De otro lado, el conjunto de soluciones óptimas del problema (P) es vacío.

En este ejemplo ninguna de las condiciones (S_p) y (S_d) se cumple.

Capítulo 7

Ejercicios

Ejercicio 1 (Caracterización de segundo orden de la convexidad) *Considere la función f definida en $\mathbb{R}^{n+1}$ $(n \geq 1)$ por*

$$f(t,x) = \begin{cases} t^{n+1}\Pi_{i=1}^{n} x_i^{-1} & \text{si } \; x > 0, \; t \geq 0, \\ 0 & \text{si } \; x \geq 0, \; t = 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Demuestre que f es convexa sci, halle su función conjugada.

Solución. a) La semicontinuidad inferior de f se puede mostrar fácilmente usando la definición.

b) Con respecto a la convexidad, mostremos primeramente que f es convexa en $\operatorname{int}(\operatorname{dom}(f)) =]0,+\infty[^{\,n+1}$. Para ello mostraremos que $\nabla^2 f(t,x)$ es semi-defnida positiva en todo $]0,+\infty[^{n+1}$. En este conjunto,

$$\nabla f(t,x) = f(t,x) \begin{pmatrix} \frac{n+1}{t} \\ -X^{-1}e \end{pmatrix},$$

donde $e = (1,1,\cdots,1)^t$ y X es la matriz diagonal cuya diagonal es x. Tenemos,

$$\begin{aligned} \frac{\nabla^2 f(t,x)}{f(t,x)} &= \begin{pmatrix} -\frac{n+1}{t^2} & 0 \\ 0 & X^{-2} \end{pmatrix} + \begin{pmatrix} \frac{n+1}{t} \\ -X^{-1}e, \end{pmatrix} \begin{pmatrix} \frac{n+1}{t} & -X^{-1}e, \end{pmatrix}, \\ &= \begin{pmatrix} \frac{n+1}{t} & 0 \\ 0 & X^{-1} \end{pmatrix} \begin{pmatrix} \frac{n}{n+1} & -e^t \\ -e & I + ee^t \end{pmatrix} \begin{pmatrix} \frac{n+1}{t} & 0 \\ 0 & X^{-1} \end{pmatrix}. \end{aligned}$$

Así, es suficiente mostrar que la matriz

$$M = \begin{pmatrix} \frac{n}{n+1} & -e^t \\ -e & I + ee^t \end{pmatrix}$$

es semi-definida positiva. Para ello, note que

$$M \begin{pmatrix} n+1 \\ e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{y} \quad \text{M} \begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ v \end{pmatrix}$$

se cumple para todo $v \in \mathbb{R}^n$ tal que $e^t v = 0$. Se deduce que 0 y 1 son valores propios de M de orden 1 y $n-1$, respectivamente. Sea λ el valor propio faltante de M, entonces

$$\lambda = \operatorname{tr}(M) - n + 1 = \frac{n}{n+1} + \operatorname{tr}(I + ee^t) - n + 1 = \frac{n(n+2)}{n+1} + 1.$$

Por lo tanto, M es semi-definida positiva.

Siendo f sci en todo $\mathbb{R}^{n+1}$ y, por lo que acabamos de demostrar, convexa en el interior de su dominio, este también es convexa en todo $\mathbb{R}^{n+1}$.

c) Ahora hallemos la conjugada de f. Por definición,

$$\begin{array}{rcl} f^*(t^*, x^*) & = & \sup\limits_{t,x} [\, tt^* + \langle x^*, x \rangle - f(t,x) \,] \\ \text{(por continuidad)} & = & \sup\limits_{t,x} [\, tt^* + \langle x^*, x \rangle - t^{n+1} \Pi_{i=1}^n x_i^{-1} \, : \, t > 0, \, x > 0 \,] \\ \text{(haciendo } x = ty\text{)} & = & \sup\limits_{t>0} [\, t(t^* + \sup\limits_{y>0} [\, \langle x^*, y \rangle - \Pi_{i=1}^n y_i^{-1} \,] \,) \,]. \end{array}$$

Denotemos

$$A(x^*) := \sup_{y>0} [\, \langle x^*, y \rangle - \Pi_{i=1}^n y_i^{-1} \,]. \tag{7.1}$$

i) Si existe i tal que $x_i^* > 0 : A(x^*) = +\infty$ y por lo tanto $f^*(t^*, x^*) = +\infty$.

ii) Si $x^* = 0 : A(x^*) = 0$ y por lo tanto $f^*(t^*, 0) = 0$ si $t^* \leq 0$, y $+\infty$ si $t^* > 0$.

iii) Si $x^* < 0$: el supremo en (7.1) se alcanza en y tal que

$$-x_i^* y_i = [\Pi_{j=1}^n (-x_j^*)]^{\frac{1}{n+1}},$$

de donde

$$A(x^*) = -(n+1)[\Pi_{j=1}^n (-x_j^*)]^{\frac{1}{n+1}}.$$

Por lo tanto, $f^*(t^*, x^*) = 0$ si $t^* - (n+1)[\Pi_{j=1}^n (-x_j^*)]^{\frac{1}{n+1}} \le 0$, y $+\infty$ en caso contrario.

iv) Si $0 \neq x^* \le 0$ con $x_i^* = 0$ para algún i : pasando al límite la expresión en iii), tenemos $A(x^*) = 0$ y por lo tanto $f^*(t^*, x^*) = 0$ si $t^* \le 0$, y $+\infty$ en caso contrario.

En resumen, tenemos

$$f^*(t^*, x^*) = \begin{cases} 0 & \text{si } \ x^* \le 0 \ \text{ y } \ t^* \le (n+1)[\Pi_{j=1}^n |x_j^*|]^{\frac{1}{n+1}}, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Ejercicio 2 (Conjugada de una convexa compuesta con una lineal) *Sean* $A \in \mathbb{R}^{p\times n}$ *de rango* p, $a \in \mathbb{R}^p$, $b \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, *y* f *una función convexa sci propia en* $\mathbb{R}^p$. *Calcule la conjugada de la función* g *definida por* $g(x) = f(Ax + a) + \langle b, x\rangle + \alpha$.

Solución. Por definición,

$$\begin{aligned} g^*(x^*) &= \sup_x [\, \langle x^*, x\rangle - f(Ax+a) - \langle b, x\rangle - \alpha \,] \\ &= -\alpha + \sup_{x,y} [\, \langle x^* - b, x\rangle - f(y) \ : \ y = Ax + a \,]. \end{aligned}$$

Aplicando la dualidad lagrangiana a la última expresión, tenemos

$$\begin{aligned} g^*(x^*) &= -\alpha + \inf_u \sup_{x,y} [\, \langle x^* - b, x\rangle - f(y) + \langle u, y - Ax - a\rangle \,] \\ &= -\alpha + \inf_u \sup_x [\, \langle x^* - b - A^t u, x\rangle + f^*(u) - \langle u, a\rangle \,] \\ &= -\alpha + \inf_u [\, f^*(u) - \langle u, a\rangle \ : \ x^* = b + A^t u] \\ &= -\alpha + f^*((AA^t)^{-1} A(x^* - b)) - \langle (AA^t)^{-1} A(x^* - b)), a\rangle. \end{aligned}$$

La última expresión se debe al hecho que AA^t es invertible. ∎

Ejercicio 3 (Funciones de recesión versus función soporte)

i) *Exprese la relación entre la función de recesión de una función convexa sci propia con la función soporte del dominio de su función conjugada.*

ii) *Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa sci propia. Considere la función $\varphi : \mathbb{R}^n \times \mathbb{R} \to \overline{\mathbb{R}}$ definida por*

$$\varphi(x,t) = \begin{cases} tf(\frac{x}{t}) & \text{si} \quad t > 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Halle φ^* y φ^{**}. Deduzca que φ es convexa. ¿Cuál es su clausura sci?.

Solución. 1) Sea $a \in \operatorname{int}(\operatorname{dom}(f)$. Entonces,

$$\begin{aligned} f_\infty(d) &= \limsup_{t\to+\infty} \frac{f(a+td)-f(a)}{t} \\ &= \sup_{t>0} \frac{f(a+td)-f(a)}{t} \\ &= \sup_{t>0} \sup_{x^*\in\operatorname{dom}(f^*)} \frac{\langle x^*, a+td\rangle - f^*(x^*) - f(a)}{t} \\ &= \sup_{x^*\in\operatorname{dom}(f^*)} \left[\langle x^*, d\rangle + \sup_{t>0} \frac{\langle x^*, a\rangle - f^*(x^*) - f(a)}{t} \right] \\ &= \delta^*(d, \operatorname{dom}(f^*)) \text{ (debido a que } \langle x^*, a\rangle - f^*(x^*) - f(a) \le 0). \end{aligned}$$

2) Por definición,

$$\varphi^*(x^*,t^*) = \sup_{t>0} \left[tt^* + t \sup_x \left[\langle x^*, \frac{x}{t}\rangle - f(\frac{x}{t}) \right] \right] = \sup_{t>0} [\, tt^* + tf^*(x^*)\,].$$

Por lo tanto,

$$\varphi^*(x^*,t^*) = \begin{cases} 0 & \text{si} \quad f^*(x^*) + t^* \le 0, \\ +\infty & \text{en caso contrario.} \end{cases}$$

Se deduce que

$$\varphi^{**}(x,t) = \sup_{x^*,t^*} [\, tt^* + \langle x^*, x\rangle \ : \ f^*(x^*) + t^* \le 0\,],$$

de donde

$$\varphi^{**}(x,t) = \begin{cases} +\infty & si \quad t < 0, \\ f_\infty(x) & si \quad t = 0, \\ tf(\frac{x}{t}) & si \quad t > 0. \end{cases}$$

Así, las funciones φ y φ^{**} coinciden en $\mathbb{R}^n \times]0, \infty\,]$. Por lo tanto φ es convexa y su clausura sci es la expresión de φ^{**}. ∎

Ejercicio 4 (La conjugada de la composición de funciones convexas) *Sea f convexa sci propia definida en $\mathbb{R}^n$ y k convexa sci propia estrictamente creciente definida en la imagen de f. Calcule la conjugada de la función g definida por $g(x) = k(f(x))$.*

Solución. Por definición,

$$\begin{aligned} g^*(x^*) &= \sup_x [\, \langle x^*, x\rangle - k(f(x))\,] \\ &= \sup_{x,t} [\, \langle x^*, x\rangle - k(t) \,:\, f(x) - t \leq 0]. \end{aligned}$$

Esta última expresión es un problema de programación convexa donde la condition de Slater se cumple, por lo tanto

$$g^*(x^*) = \inf_{\lambda \geq 0} \sup_{x,t} [\, \langle x^*, x\rangle - k(t) + \lambda(f(x) - t)].$$

Tenemos,

$$\sup_{x,t} [\, \langle x^*, x\rangle - k(t) + \lambda(f(x) - t)] = \begin{cases} -\inf_t k(t) + \delta^*(x^*, \mathrm{dom}\,(f)) & \text{si} \quad \lambda = 0, \\ \\ k^*(-\lambda) + \lambda f^*(\frac{x^*}{\lambda}) & \text{si} \quad \lambda > 0. \end{cases}$$

Por las relaciones obtenidas en el Ejercicio 3, el supremo para $\lambda = 0$, es el límite cuando λ tiende a 0. Por lo tanto

$$g^*(x^*) = \inf_{\lambda > 0} [\, k^*(-\lambda) + \lambda f^*(\frac{x^*}{\lambda})\,].$$

Ejercicio 5 (Clausura sci de una función convexa) *Sea* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *convexa sci propia y sea* $C \subset \operatorname{int}(\operatorname{dom}(f)$ *convexo cerrado acotado. El objetivo de este ejercicio es hallar la función convexa sci más pequeña que coincida con* f *en el conjunto* C*. Dar algunas propiedades de esta función.*

Solución. Sea $A = [C \times \mathbb{R}^n] \cap \operatorname{gr}(\partial f)$. Este es compacto porque ∂f es scs en el compacto C. Este también es no vacío porque $C \subset \operatorname{int}(\operatorname{dom}(f)$ (lo que implica que $\partial f(y) \neq \emptyset$ para cada $y \in C$).

Consideremos la función

$$g(x) = \sup_{y,y^*} [\, f(y) + \langle y^*, x - y\rangle \ : \ (y, y^*) \in A\,]. \qquad (prol)$$

Mostraremos que esta función cumple con las propiedades requeridas.

i) Esta es convexa sci como supremo de funciones afines. Además, para cada $x \in \mathbb{R}^n$, el supremo se alcanza en A debido a que f es continua en $\operatorname{int}(\operatorname{dom}(f))$. Se deduce que g es una función convexa de $\mathbb{R}^n$ en $\mathbb{R}$, y por lo tanto continua en todo $\mathbb{R}^n$.

ii) Sea $x \in \mathbb{R}^n$. Debido a que $f(x) \geq f(y) + \langle y^*, x - y\rangle$ para todo $(y, y^*) \in A$, tenemos $f(x) \geq g(x)$.

iii) Si $(x, x^*) \in A$, entonces $f(x) = f(x) + \langle x^*, x - x\rangle \leq g(x)$.

Por lo tanto, de ii) y iii), tenemos

$$g(x) = f(x) \text{ para todo } x \in C.$$

iv) Por la definición de g, esta es la función convexa sci más pequeña que coincide con f en C.

v) Sea $x \in \mathbb{R}^n$. Existe $(y(x), y^*(x)) \in A$ tal que

$$g(x) = f(y(x)) + \langle y^*(x), x - y(x)\rangle$$

y por lo tanto, para todo $x' \in \mathbb{R}^n$,

$$g(x') \geq f(y(x)) + \langle y^*(x), x' - y(x)\rangle = g(x) + \langle y^*(x), x' - x\rangle.$$

Se deduce que $y^*(x) \in \partial g(x)$, y de este

$$g^*(y^*(x)) = g(x) - \langle y^*(x), x\rangle = f(y(x)) - \langle y^*(x), x\rangle = f^*(y^*(x)).$$

Ejercicio 6 (Hessiana de la conjugada de una función convexa) *Sea f una función convexa $\mathcal{C}^2$ en una vecindad V de $\bar{x}$ con $\nabla^2 f(\bar{x})$ definida positiva. Sea $\bar{x}^* = \nabla f(\bar{x})$. Demuestre que $\nabla^2 f^*$ existe y es continua en una vecindad de $\bar{x}^*$. Demuestre también que $\nabla^2 f(\bar{x})\nabla^2 f^*(\bar{x}^*) = I$.*

Sea $\varphi : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$ convexa sci propia y $h : \mathbb{R}^p \to \overline{\mathbb{R}}$ definida por $h(u) = \inf_x \varphi(x, u)$. Sean $\bar{x}$ y $\bar{u}$ tales que $h(\bar{u}) = \varphi(\bar{x}, \bar{u})$. Asuma que la función φ es $\mathcal{C}^2$ en una vecindad de $(\bar{x}, \bar{u})$ y que la matriz $\nabla^2\varphi(\bar{x}, \bar{u})$ es definida positiva. Demuestre que h es dos veces diferenciable en $\bar{u}$. Calcule $\nabla h(\bar{u})$ y $\nabla^2 h(\bar{u})$.

Solución. 1) Primeramente demostremos que f^* es diferenciable en $\bar{x}^*$. Por la definición del subdiferencial, $x \in \partial f^*(\bar{x}^*)$ si y solo si x es una solución optimal del problema convexo

$$\min_y [f(y) - \langle \bar{x}^*, y\rangle]. \tag{p_x}$$

Debido a que $\bar{x}^* = \nabla f(\bar{x})$ y f fuertemente convexa en una vecindad de $\bar{x}$, entonces $\bar{x}$ es la única solución del problema $(p_{\bar{x}})$. Así, el conjunto $\partial f^*(\bar{x}^*)$ se reduce al singleton $\{\bar{x}\}$ y por lo tanto f^* es diferenciable en $\bar{x}^*$, con $\bar{x} = \nabla f^*(\bar{x}^*)$.

Debido a la semicontinuidad superior de ∂f^*, existe una vecindad W de $\bar{x}^*$ tal que $\partial f^*(x^*) \subset V$. Sea $(x^*, x) \in [W \times V] \cap \mathrm{gr}\,(\partial f^*)$, entonces $x^* = \nabla f(x)$ debido a que f es diferenciable en V. Aplicando la primera parte, tenemos $x = \nabla f^*(x^*)$.

Así, hemos demostrado que en la vecindad $V \times W$ de $(\bar{x}, \bar{x}^*)$, se cumple

$$x^* = \nabla f(x) \Longleftrightarrow x = \nabla f^*(x^*).$$

Debido a que $\nabla^2 f$ es inversible en V, se cumple $\nabla^2 f(x)\nabla^2 f^*(x^*) = I$ en $V \times W$.

2) Por definición, $h^*(u^*) = \varphi^*(0, u^*)$ para todo u^*. Debido a que $0 = \nabla_x\varphi(\bar{x}, \bar{u})$, existe $\bar{u}^*$ tal que $(0, \bar{u}^*) = \nabla\varphi(\bar{x}, \bar{u})$. Por la parte 1), $\nabla\varphi^*$ y $\nabla^2\varphi^*$ existen en una vecindad de $(0, \bar{u}^*)$, con $\nabla\varphi^*(0, \bar{u}^*) = (\bar{x}, \bar{u})$ y $\nabla^2\varphi^*(0, \bar{u}^*) = [\nabla^2\varphi(\bar{x}, \bar{u})]^{-1}$. Se deduce que

$$\nabla h^*(\bar{u}^*) = \bar{u} \quad \text{y} \quad \nabla^2 h^*(\bar{u}^*) = (0\ I_p)[\nabla^2\varphi(\bar{x}, \bar{u})]^{-1}(0\ I_p)^t.$$

Ejercicio 7 (Distancia a un conjunto convexo cerrado) *Considere un conjunto* $C \subset \operatorname{int}(\operatorname{dom}(f)$ *convexo cerrado. Para* $x \in \mathbb{R}^n$*, sea* $p(x)$ *la proyección ortogonal (con la norma euclidiana) de* x *sobre* C*. Demuestre que las funciones* d_C *y* f_C *definidas por*

$$d_C(x) = \|x - p(x)\| \quad y \quad f_C(x) = \frac{1}{2}\|x - p(x)\|^2 \quad \textit{para todo } x \in \mathbb{R}^n,$$

son convexas. ¿Son diferenciables ?.

Solución. Considere la función φ definida por

$$\varphi(x,y) = \begin{cases} \|x - y\| & \text{si } \ y \in C, \\ +\infty & \text{en otro caso.} \end{cases}$$

Esta es convexa en $\mathbb{R}^n \times \mathbb{R}^n$. Debido a que $d_C(x) = \inf_{y \in C} \varphi(x,y)$, esta también es convexa e igualmente f_C. Ambas funciones son finitas en todo $\mathbb{R}^n$ y por lo tanto continuas. Por definición, $f_C(x) = d_C(x) = 0$ si $x \in C$. Para todo $x, z \in \mathbb{R}^n$, se cumple

$$f_C(z) \le \frac{1}{2}\|z - x + x - p(x)\|^2 = f_C(x) + \langle z - x, x - p(x)\rangle + \frac{1}{2}\|x - z\|^2.$$

Si $x \notin C$, entonces

$$x \notin D = \{y \,:\, \langle p(x) - x, y - p(x)\rangle \ge 0\} \supset C$$

debido a que $p(x)$ es la proyección de x.

Sea V_x una vecindad de x que no interseca a D. Si $z \in V_x$, entonces

$$f_C(z) \ge \inf_{y \in D} \frac{1}{2}\|z - y\|^2 \ge f_C(x) + \langle z - x, x - p(x)\rangle.$$

Se deduce que para todo z en una vecindad de x, se cumple

$$0 \le \frac{f_C(z) - f_C(x) - \langle z - x, x - p(x)\rangle}{\|z - x\|} \le \frac{1}{2}\|z - x\|.$$

Note que esta desigualdad también es válida si $x \in C$. Por lo tanto, f_C es diferenciable en todo $x \in \mathbb{R}^n$, con $\nabla f_C(x) = x - p(x))$.

De este se deduce que d_C es diferenciable en todo x que no pertenece a la frontera de C, con

$$\nabla d_C(x) = \begin{cases} 0 & \text{si} \quad x \in \operatorname{int}(\mathrm{C}), \\ \|x-p(x)\|^{-1}(x-p(x)) & \text{si} \quad x \notin C. \end{cases}$$

Ejercicio 8 (Caracterización de un conjunto convexo) *Sea $C \neq \mathbb{R}^n$ convexo cerrado con interior no vacío. Para cada $x \in \mathbb{R}^n$, considere*

$$\Pi(x) := \arg\min[\, \|y-x\| \,:\, y \in C,\, y \notin \operatorname{int}(\mathrm{C})\,].$$

1) *Dado $x \in \operatorname{int}(\mathrm{C})$. Demuestre que para todo $\pi \in \Pi(x)$,*

$$\langle x-\pi, y-\pi\rangle \geq 0 \quad \textit{para todo} \;\; y \in C. \tag{7.2}$$

2) *Sea $\delta_C^*(x^*) = \sup\,[\,\langle x^*, y\rangle \,:\, y \in C]$, la función soporte de C, y sean*

$$B = \{x^* \,:\, \|x^*\| \leq 1\} \qquad y \qquad S = \{x^* \,:\, \|x^*\| = 1\},$$

la bola y la esfera unitaria (con la norma euclideana) de $\mathbb{R}^n$.

Estudie y compare las tres funciones siguientes :

$$\begin{aligned} \delta_C(x) &= \sup_{x^* \in \mathbb{R}^n} [\,\langle x^*, x\rangle - \delta_C^*(x^*)] = \sup_{x^* \in \mathbb{R}^n} \inf_{y \in C} \langle x^*, x-y\rangle, && (7.3) \\ d_C(x) &= \sup_{x^* \in B} [\,\langle x^*, x\rangle - \delta_C^*(x^*)\,] = \sup_{x^* \in B} \inf_{y \in C} \langle x^*, x-y\rangle, && (7.4) \\ \Delta_C(x) &= \sup_{x^* \in S} [\,\langle x^*, x\rangle - \delta_C^*(x^*) \,= \sup_{x^* \in S} \inf_{y \in C} \langle x^*, x-y\rangle. && (7.5) \end{aligned}$$

Solución. 1) Note primeramente que $\Pi(x)$ es cerrado y no vacío. Sea $\pi \in \Pi(x)$. Entonces $A := \{z \,:\, \|z-x\| < \|\pi - x\|\} \subset \operatorname{int}(C)$.

Asumamos, por contradicción, que existe $y \in C$ tal que $\langle x-\pi, y-\pi\rangle < 0$. Sea $z_t = \pi + t(\pi - y)$ con $t > 0$. Para $t > 0$ suficientemente pequeño, se tiene

$$\|x-z_t\|^2 = \|x-\pi\|^2 + 2t\langle x-\pi, y-\pi\rangle + t^2\|y-\pi\|^2 < \|\pi - x|^2,$$

y por lo tanto $z_t \in A \subset \operatorname{int}(\mathrm{C})$. Se deduce que $\pi \in \operatorname{int}(\mathrm{C})$, en contradicción con la hipótesis.

2) Las tres funciones δ_C, d_C y Δ_C son convexas sci como supremo de funciones afines. Note que

$$\delta_C \geq d_c \geq \Delta_C. \tag{7.6}$$

La función δ_C es la función característica de C y d_C es la función distancia al conjunto C.

a) Sea $x \in C$. Haciendo $y = x$ en (7.3) y $x^* = 0$ en (7.4), se obtiene

$$\delta_C(x) = d_C(x) = 0 \quad (\text{de donde } \Delta_C(x) \leq 0).$$

b) Sea $x \notin C$. Denotando por p a la projección de x sobre C (entonces $\Pi(x) = \{p\}$), para todo x^* se cumple

$$\inf_{y \in C} \langle x^*, x - y\rangle \leq \langle x^*, x - p\rangle \leq \|x^*\| \, \|x - p\|,$$

y por lo tanto,

$$\Delta_C(x) \leq d_C(x) \leq \|x - p\|. \tag{7.7}$$

Sea $\bar{x}^* = \|p - x\|^{-1}(p - x)$. Debido a que p es la proyección de x sobre C, para todo $y \in C$ se cumple

$$0 < \|x - p\| = \langle \bar{x}^*, p - x\rangle \leq \langle \bar{x}^*, y - x\rangle, \tag{7.8}$$

de donde

$$\Delta_C(x) \geq \|p - x\|. \tag{7.9}$$

De (7.7) y (7.9) se obtiene

$$\Delta_C(x) = d_C(x) = \|x - p\|.$$

Multiplicando por $\lambda > 0$ a la expresión en (7.8), se obtiene

$$\lambda\|x - p\| = \langle \lambda x^*, p - x\rangle \leq \langle \lambda x^*, y - x\rangle \leq \delta_C(x).$$

Haciendo $\lambda \to +\infty$, se obtiene $\delta_C(x) = +\infty$.

c) Sea $x \in \operatorname{int}(\mathrm{C})$ y $\pi \in \Pi(x)$. Sea $\bar{x}^* = \|x-\pi\|^{-1}(\pi - x)$. Debido a (7.2), se obtiene

$$-\|x-\pi\| = \langle \bar{x}^*, x-\pi\rangle = \langle \bar{x}^*, x-y\rangle + \langle \bar{x}^*, y-\pi\rangle \leq \langle \bar{x}^*, x-y\rangle \quad \forall\, y \in C$$

y por lo tanto

$$\Delta_C(x) \geq -\|x-\pi\|. \tag{7.10}$$

Sea $A = \{y \,:\, \|x-y\| \leq \|x-\pi\|\}$. Se cumple $A \subset C$ y por lo tanto para todo x^* tal que $\|x^*\| = 1$, se obtiene

$$\inf_{y\in C} \leq x^*, x-y\rangle \leq \inf_{y\in A} \leq x^*, x-y\rangle = -\|x-\pi\|,$$

de donde

$$\Delta_C(x) \leq -\|x-\pi\|. \tag{7.11}$$

De (7.10) y (7.11), se obtiene

$$\Delta_C(x) = -\|x-\pi\|.$$

d) En resumen, para todo $\pi \in \Pi(x)$:
— Si $x \in \operatorname{int}(C)$, entonces $0 = \delta_C(x) = d_C(x) > \Delta_C(x) = -\|x-\pi\|$,
— Si $x \notin \operatorname{int}(C)$, $x \in C$, entonces $0 = \delta_C(x) = d_C(x) = \Delta_C(x)$,
— Si $x \notin C$, entonces $+\infty = \delta_C(x) > d_C(x) = \Delta_C(x) = \|x-\pi\|$.

Ejercicio 9 (Gauges convexos, normas y duales) *Considere $C \subset \mathbb{R}^n$ convexo cerrado acotado tal que $0 \in \operatorname{int}(C)$. La función*

$$j(x) = \inf_t \,[\, t > 0 \,:\, x \in tC\,]$$

es llamada gauge (o funcional de Minkowski) de C.

1) Estudie las propiedades de j. Demuestre que su función conjugada j^* es la función característica de un conjunto convexo cerrado, que lo denotaremos por C°. Demuestre que C° es acotado y que $0 \in \operatorname{int}(C^\circ)$.

2) Denote por j° a la función gauge de C°, que será llamada gauge dual de j. Demuestre que $\langle x^*, x\rangle \leq j(x)j^\circ(x^*)$ para todo $x, x^* \in \mathbb{R}^n$ y que $C = C^{\circ\circ}$. Deduzca que $j = j^{\circ\circ}$ y de este, j es la función soporte de C°.

3) Demuestre que j es una norma desde que $C = -C$.

Solución. Es fácil verificar que

$$C = \{x \, : \, j(x) \leq 1\} \quad \text{y} \quad j(x) = \begin{cases} 0 & \text{si} \quad x = 0, \\ \in]0, \infty[& \text{si} \quad x \neq 0. \end{cases}$$

También, para todo $x \in \mathbb{R}^n$ y $\lambda \geq 0$, se cumple

$$j(\lambda x) = \lambda j(x).$$

Mostraremos que j es convexa. Sean $x, y \in \mathbb{R}^n$ y $\lambda \in]0, 1[$, mostraremos que

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

Sin pérdida de generalidad asumiremos que $x \neq 0$ e $y \neq 0$. Denotemos

$$x' = \frac{x}{j(x)}, \; y' = \frac{y}{j(y)}, \; \mu = \lambda j(x) + (1-\lambda)j(y), \; z = \frac{\lambda j(x)}{\mu} x' + \frac{(1-\lambda)j(y)}{\mu} y'.$$

Se cumple,

$$j(\lambda x + (1-\lambda)y) = j(\mu z) = \mu j(z).$$

Debido a que x' e y' pertenecen a C, también z pertenece a C, y por lo tanto $j(z) \leq 1$, lo que demuestra la convexidad de j.

2) Por la definición de conjugada,

$$j^*(x^*) = \sup_{x \in \mathbb{R}^n, \, \lambda \geq 0} [\, \langle x^*, \lambda x \rangle - j(\lambda x) \,] = \sup_{x \in \mathbb{R}^n, \, \lambda \geq 0} [\, \lambda \, [\, \langle x^*, x \rangle - j(x)] \,].$$

Se deduce que $j^*(x^*) = +\infty$ desde que si existe x tal que $\langle x^*, x \rangle - j(x) > 0$, y $j^*(x^*) = 0$ en caso contrario. Por lo tanto, j^* es la función característica del conjunto

$$C^\circ = \bigcap_{x \in \mathbb{R}^n} \{x^* \, : \, \langle x^*, x \rangle \leq j(x)\} = \bigcap_{x \in C} \{x^* \, : \, \langle x^*, x \rangle \leq 1\}.$$

Se sigue que j es la función soporte de C°.

El conjunto C° es convexo cerrado como intersección de convexos cerrados. Debido a que $0 \in \operatorname{int}(\mathrm{C})$, existe $r > 0$ tal que $x \in C$ cada vez que $\|x\| \leq r$. Se deduce que $\|x^*\| \leq r^{-1}$ para todo $x^* \in C^\circ$.

De otro lado, debido a que C es acotado, existe ρ tal que $\|x\| < \rho$ para todo $x \in C$. Por lo tanto todo x^* satisfaciendo $\|x^*\| \leq \rho^{-1}$, está en C°. Así, hemos demostrado que C° es convexo cerrado acotado con $0 \in \operatorname{int}(C^\circ)$.

Mostremos ahora que $C = C^{\circ\circ}$. Sea $x \in C$, entonces $\langle x^*, x\rangle \leq 1$ para todo $x^* \in C^\circ$ y por lo tanto, $x \in C^{\circ\circ}$. Si $x \notin C$, entonces debido a que C es convexo cerrado, existe $\alpha \in \mathbb{R}$ y $x^* \in \mathbb{R}^n$ tal que $\langle x^*, x\rangle > \alpha > \langle x^*, y\rangle$ para todo $y \in C$. La condición $0 \in \operatorname{int}(\mathrm{C})$, implica que $\alpha > 0$, el cual podemos asumir $\alpha = 1$. Por lo tanto $x^* \in C^\circ$, lo que implica que $x \notin C^{\circ\circ}$. Acabamos de demostrar que

$$C = C^{\circ\circ} = \bigcap_{x^* \in C^\circ} \{x \,:\, \langle x^*, x\rangle \leq 1\}.$$

Consideremos ahora la función gauge j° de C°. Por construcción se tiene $\langle x^*, x\rangle \leq 1$ para todo $x \in C$ y todo $x^* \in C^\circ$. Se deduce que

$$\langle x^*, x\rangle \leq j(x)\, j^\circ(x^*) \quad \text{para todo} \quad x, x^* \in \mathbb{R}^n.$$

La dualidad simétrica entre los conjuntos C y $C^{\circ\circ}$ implica una dualidad simétrica entre sus gauges asociadas j y j°. Así, j° es la función soporte de $C = C^{\circ\circ}$.

Asumamos ahora que $C = -C$. Entonces $j(\lambda x) = |\lambda| j(x)$ para todo $x \in \mathbb{R}^n$ y todo $\lambda \in \mathbb{R}$. Por lo tanto, la función gauge j es una norma. ∎

Ejercicio 10 (La norma p y su dual) *Para $p > 1$ finito, considere la función $\|\cdot\|_p$ definida por*

$$\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{\frac{1}{p}} \quad \textit{para todo } x \in \mathbb{R}^n.$$

Usando el ejercicio precedente verifique que este es una norma. Halle su norma dual.

Solución. Sean $f(x) = \sum_{i=1}^n |x_i|^p$ y $C = \{x \,:\, f(x) \leq 1\}$. La función f es convexa y toma valor finito en todo $\mathbb{R}^n$. Por lo tanto, C es convexo cerrado.

Este también es acotado. Note que $0 \in \operatorname{int}(\mathrm{C})$ y $C = -C$. Por el ejercicio precedente, $\|\cdot\|_p$ es la norma de bola unitaria C.

La norme dual es

$$n^\circ(x^*) := \max_x [\, \langle x^*, x\rangle \;:\; x \in C\,] = \max_x [\, \langle x^*, x\rangle \;:\; f(x) \le 1\,].$$

Sea $x^* \neq 0$. Usando la condición necesaria y suficiente de optimalidad, x es una solución óptima si y solo si $f(x) = 1$ y existe $\lambda > 0$ tal que

$$x_i^* = \lambda \varepsilon(x_i)|x_i|^{p-1} \quad \text{para todo } \; i = 1, \cdots n,$$

donde $\varepsilon(x_i)$ es 1 si $x_i \geq 0$, y -1 en otro caso. Se deduce que x_i y x_i^* tienen el mismo signo y por lo tanto $\varepsilon(x_i) = \varepsilon(x_i^*)$. Así, para todo i,

$$|x_i| = [\frac{|x_i^*|}{\lambda}]^{\frac{1}{p-1}}.$$

De la igualdad $f(x) = 1$ se obtiene λ, de donde $n^\circ(x^*) = \|x^*\|_q$, siendo q (> 1) tal que $\frac{1}{p} + \frac{1}{q} = 1$. Las normas p y q son llamadas normas duales entre si. La norma cuadrática corresponde a $p = 2$. La desigualdad (que se deduce del ejercicio precedente)

$$|\langle x^*, x\rangle| \le \|x^*\|_q \, \|x\|_p \;\; \text{para todo } \; x, x^* \in \mathbb{R}^n,$$

generaliza la desigualdad de Schwartz.

Ejercicio 11 (Función de recesión y penalidades cuadráticas) *Considere el siguiente problema*

$$\alpha = \inf_{x \in C} f(x) \quad \textit{donde} \quad C = \{\, x \in \mathbb{R}^n \;:\; g_i(x) \le 0, \; i = 1, \cdots p, \; Ax = a \,\}, \quad (P)$$

y los problemas de penalidad asociados

$$\alpha(r) = \inf_x \left[f_r(x) = f(x) + r \left(\sum_{i=1}^{p} [g_i^+(x)]^2 + \|Ax - a\|^2 \right) \right], \qquad (P_r)$$

donde las funciones f, g_i son convexas sci propias, $a \in \mathbb{R}^q$, $A \in \mathbb{R}^{q \times n}$, $g_i^+(x) = \max[0, g_i(x)]$ y $r > 0$. Asuma que existe $\bar{x}$ tal que $A\bar{x} = a$, $f(\bar{x}) < +\infty$ y $g_i(\bar{x}) < 0$ para todo i.

Denote por S y S_r a los conjuntos de soluciones óptimas de (P) y (P_r), respectivamente. Demuestre que

a) *S y S_r son convexos cerrados.*

b) *S es acotado no vacío si y solo si S_r lo es.*

c) *Asuma que S es acotado no vacío. Considere una sucesión $\{r_k\}$ estrictamente creciente tal que $r_k \to +\infty$ cuando $k \to +\infty$. Sea x_k una solución óptima del problema (P_{r_k}). Demuestre que la sucesión $\{x_k\}$ es acotada y que todo sus valores de adherencia son soluciones óptimas de (P).*

Solución. a) Por las hipótesis, S y S_r son convexos cerrados.

b) El conjunto S es acotado no vacío si y solo si la función convexa sci propia $\hat{f}$ definida por $\hat{f}(x) = f(x)$ si $x \in C$, y $+\infty$ en caso contrario, es inf-compacto. Este es equivalente a que el conjunto convexo cerrado no vacío

$$D = \{\, x \,:\, f(x) \leq f(\bar{x}),\, g_i(x) \leq 0,\, i = 1, \cdots, p,\, Ax = a \,\}$$

es acotado, lo que que a su vez es equivalente a que el cono

$$D_\infty = \{\, d \,:\, f_\infty(d) \leq 0,\, (g_i)_\infty(d) \leq 0,\, i = 1, \cdots, p,\, Ad = 0 \,\}$$

se reduce a $\{0\}$.

De otro lado, S_r es acotado no vacío si y solo si la función convexa sci propia f_r es inf-compacto.

i) Asuma que S es acotado no vacío. Si existe $d \neq 0$ tal que

$$(f_r)_\infty(d) = \lim_{t \to +\infty} \frac{f_r(\bar{x} + td) - f_r(\bar{x})}{t} \leq 0,$$

entonces $d \in D_\infty$. Se deduce que S_r es acotado no vacío.

ii) Asuma que S_r es acotado no vacío. Si existe $d \neq 0$ tal que $Ad = 0$ y $(g_i)_\infty(d) \leq 0$ para todo $i = 1, \cdots, p$, entonces $f_r(\bar{x} + td) = f(\bar{x} + td)$ para todo $t \geq 0$, y por lo tanto $f_\infty(d) > 0$. Se deduce que S es acotado no vacío.

c) See $D_{r_k} := \{\, x \,:\, f_r(x) \leq f_r(\bar{x}) = f(\bar{x}) = \hat{f}(\bar{x}) \,\}$. Debido a que $f_{r_1}(x) \leq f_{r_k}(x) \leq \hat{f}(x)$ para todo x, se cumple $D_{r_1} \supset D_{r_k} \supset D$. Por lo tanto la sucesión $\{x_k\}$ está contenida en el compacto D_{r_1}. Sea $\hat{x}$ uno de sus valores de adherencia. Es fácil verificar que la sucesión $\alpha(r_k)$ es creciente y mayorada por α. Así, $f(x_k) \leq f_{r_k} = \alpha(r_k) \leq \alpha$ y por lo tanto $f(\hat{x}) \leq \alpha$. Si $\hat{x} \notin C$, entonces $f_r(x_k) \to +\infty$ cuando $k \to +\infty$, lo que es imposible. Se deduce por lo tanto que $\hat{x}$ es una solución óptima del problema (P). ■

Ejercicio 12 (Optimización DC) *Considere el siguiente problema*

$$\alpha = \inf_x [\, f(x) - g(x)\,] \tag{DC}$$

donde f y g son convexas sci propias. La función $h = f - g$ es llamada función DC (diferencia de funciones convexas).

a) *Sea $\bar{x} \in \operatorname{int}(\operatorname{dom}(f)) \cap \operatorname{int}(\operatorname{dom}(g))$. Demuestre que una condición necesaria para que $\bar{x}$ sea una solución optimal del problema es que $\partial f(\bar{x})) \supset \partial g(\bar{x}))$. Dar un ejemplo donde esta condición no es suficiente.*

b) *Demuestre que*

$$\alpha = \inf_{x^* \in \mathbb{R}^n} [\, g^*(x^*) - f^*(x^*)\,].$$

Solución. a) Las funciones f y g siendo convexas, admiten derivadas direccionales en $\bar{x}$. Por lo tanto, una condición necesaria de optimalidad en $\bar{x}$ es

$$0 \leq h'(\bar{x}, d) = f'(\bar{x}, d) - g'(\bar{x}, d) \quad \text{para todo } d \in \mathbb{R}^n.$$

Recordemos que

$$f'(\bar{x}, d) = \sup [\, \langle x^*, d\rangle \;:\; x^* \in \partial f(\bar{x})\,] = \delta^*(d, \partial f(\bar{x})),$$

$$g'(\bar{x}, d) = \sup [\, \langle x^*, d\rangle \;:\; x^* \in \partial g(\bar{x})\,] = \delta^*(d, \partial g(\bar{x})).$$

La condición de optimalidad en término de las funciones soportes es

$$\delta^*(\cdot, \partial g(\bar{x})) \leq \delta^*(\cdot, \partial f(\bar{x})).$$

En término de la función indicatriz, la desigualdad anterior es equivalente a

$$\delta(\cdot, \partial f(\bar{x})) \leq \delta(\cdot, \partial g(\bar{x}))$$

que a su vez es equivalente a

$$\partial g(\bar{x}) \subset \partial f(\bar{x}).$$

Para mostrar que la condición necesaria no es suficiente, considere como ejemplo las funciones convexas $f(x,y) = x^2$ y $g(x,y) = y^2$. En el punto $(0,0)$ se verifica la inclusión precedente pero este punto no es ni siquiera solución local.

b) Siendo g convexa sci propia, se verifica

$$g(x) = \sup\left[\, \langle x^*, x\rangle - g^*(x^*) \;:\; x^* \in \mathbb{R}^n \,\right].$$

Por lo tanto,

$$\begin{aligned} \alpha &= \inf_{x,x^*}\left[\, f(x) + g^*(x^*) - \langle x^*, x\rangle \,\right] \\ &= \inf_{x^*}\left[\, g^*(x^*) - \sup_{x}\left[\, \langle x^*, x\rangle - f(x) \,\right]\right] \\ &= \inf_{x^*}\left[\, g^*(x^*) - f^*(x^*) \,\right]. \end{aligned}$$

Ejercicio 13 (Conjugada de funciones matriciales) *1) Denote por K al ortante estrictamente positivo de $\mathbb{R}^n$. Sean $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ definidas por*

$$f(x) = \begin{cases} -n\prod_{i=1}^{n} x_i^{1/n} & si \quad x \in K, \\ +\infty & si \quad x \notin K. \end{cases} \qquad g(x) = \begin{cases} -\sum_{i=1}^{n} \ln(x_i) & si \quad x \in K, \\ +\infty & si \quad x \notin K. \end{cases}$$

Halle sus funciones conjugadas.

2) Sean E el conjunto de las matrices simétricas de orden n y K el subconjunto de E formado por las matrices definidas positivas. Considere las funciones $f, g : E \to \overline{\mathbb{R}}$ definidas por

$$f(X) = \begin{cases} -n[\det(X)]^{1/n} & si \quad X \in K, \\ +\infty & si \quad X \notin K. \end{cases} \qquad g(X) = \begin{cases} -\ln(\det(X)) & si \quad X \in K, \\ +\infty & si \quad X \notin K. \end{cases}$$

Halle sus funciones conjugadas.

Solución.

1a) Por la definición de conjugada,

$$\begin{aligned} f^*(x^*) &= \sup_{x\in K}\left[\sum_{i=1}^n x_i^* x_i + n\prod_{i=1}^n x_i^{1/n}\right] \\ &= \sup_{x\in K,\lambda\geq 0}\left[\sum_{i=1}^n x_i^*(\lambda x_i) + n\prod_{i=1}^n [\lambda x_i]^{1/n}\right]. \end{aligned}$$

Sea

$$C=\left\{x^*\in \mathbb{R}^n \,:\, \sum_{i=1}^n x_i^* x_i + n\prod_{i=1}^n x_i^{1/n}\leq 0 \;\text{ para todo }\; x\in K\right\}.$$

Claramente $f^*(x^*) = +\infty$ si $x \notin C$, y $f^*(x^*) = 0$ si $x \in C$. Por lo tanto, f^* es la función indicatriz del conjunto C.

Busquemos una expresión más explícita del conjunto C. Sea $x^* \in \mathbb{R}^n$ tal que su $r-$ésima coordenada x_r^* es no negativa. Entonces para $x \in \mathbb{R}^n$ con $x_j = 1$ si $j \neq r$ y x_r tendiendo a $+\infty$, la expresión $\sum_{i=1}^n x_i^* x_i + n\prod_{i=1}^n x_i^{1/n}$ también tiende a $+\infty$. Se deduce por lo tanto que $C \subset -K$. Se puede verificar fácilmente que

$$\sup_{x\in K}\left[n\prod_{i=1}^n x_i \,:\, \sum_{i=1}^n x_i^* x_i = -1\right] = \prod_{i=1}^n |x_i^*|^{-1/n}. \qquad (7.12)$$

Este un problema de maximizatión de una función cóncava con una restricción lineal. Usando la condición de optimalidad, el supremo se alcanza en x tal que $x_i = [-nx_i^*]^{-1}$ para todo i. Se deduce que

$$C=\left\{x^*\in -K \,:\, \prod_{i=1}^n |x_i^*|^{1/n}\geq 1\right\}.$$

1b) Usando nuevamente la definición de conjugada, tenemos

$$\begin{aligned} g^*(x^*) &= \sup_{x\in K}\left[\sum_{i=1}^n [x_i^* x_i + \ln(x_i)]\right] \\ &= \begin{cases} -n-\sum_{i=1}^n \ln(-x_i^*) & \text{si} \quad x^*<0, \\ +\infty & \text{en caso contrario.} \end{cases} \end{aligned}$$

2a) Sea $X^* \in E$. Existe D diagonal y P tales que $PP^t = I$ y $X^* = PDP^t$. Tenemos,

$$\begin{aligned} f^*(X^*) &= \sup_{X\in K} [\,\mathrm{tr}\,(XX^*) + n[\det(X)]^{1/n}\,] \\ &= \sup_{X\in K} [\,\mathrm{tr}\,(XPDP^t) + n[\det(X)]^{1/n}\,] \\ &= \sup_{X\in K} [\,\mathrm{tr}\,(P^tXPD) + n[\det(P^tXP)]^{1/n}\,] \\ &= \sup_{X\in K} [\,\mathrm{tr}\,(XD) + n[\det(X)]^{1/n}\,]. \end{aligned}$$

Si $X^* \notin -K$, entonces existe un elemento d_r de la diagonal de D que es no negativo. Por lo tanto, para X diagonal tal que $x_{jj} = 1$ si $j \neq r$ y $x_{rr} \to +\infty$, se verifica que $\mathrm{tr}\,(XD) + n[\det(X)]^{1/n} \to +\infty$. Se sigue que $f^*(X^*) = +\infty$.

Suponga ahora que $X^* \in -K$. Existe R simétrica definida positiva tal que $X^* = -R^{-2}$. Para cada $X \in K$, considere la matriz simétrica definida positiva $Z = R^{-1}XR^{-1}$. Tenemos

$$\begin{aligned} f^*(X^*) &= \sup_{X\in K} [\,-\mathrm{tr}\,(\mathrm{R}^{-2}\mathrm{X}) + \mathrm{n}[\det(\mathrm{X})]^{1/\mathrm{n}}\,] \\ &= \sup_{Z\in K} [\,-\mathrm{tr}\,(\mathrm{Z}) + \mathrm{n}[\det(\mathrm{Z})\det(\mathrm{R}^2)]^{1/\mathrm{n}}\,]. \end{aligned}$$

A toda matriz Z simétrica definida positiva se le asocia una matriz diagonal definida positiva Δ y una matriz P tales que $PP^t = I$ y $P\Delta P^t = Z$. Note que $\mathrm{tr}\,(\mathrm{Z}) = \sum \delta_{\mathrm{i}}$ y $\det(Z) = \prod \delta_i$. Por lo tanto

$$f^*(X^*) = \sup_{\delta>0} \left[-\sum \delta_i + n[\det(-X^*)]^{-1/n} \prod_{i=1}^{n} \delta_i^{1/n} \right].$$

Por la parte 1a), se cumple que $f^*(X^*) = 0$ si $\det(-X^*) \geq 1$ y $f^*(X^*) = +\infty$ en otro caso. Así, f^* es la función indicatriz del conjunto

$$C = \{X^* \in -K \,:\, \det(-X^*) \geq 1\}.$$

2b) Procediendo como en el caso anterior, tenemos $g^*(X^*) = +\infty$ si $X^* \notin -K$. Si $X^* \in -K$, entonces

$$\begin{aligned} g^*(X^*) &= \sup_{\delta>0}\left[-\sum_{i=1}^{n}\delta_i + \ln(\det(R^2) + \sum_{i=1}^{n}\ln(\delta_i)\right] \\ \text{(por la parte 1b))} &= \ln(\det(-X^*)) - n. \end{aligned}$$

Ejercicio 14 (Valores propios de matrices simétricas) *Denote por $\mathcal{M}$ al conjunto de todas las matrices simétricas de orden n. Para $M \in \mathcal{M}$, ordene sus autovalores de forma decreciente, $\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_n(M)$.*

Dado $q \in [1, n]$ entero, considere la función $f_p : \mathcal{M} \to \mathrm{I\!R}$ definida por

$$f_q(M) = \lambda_1(M) + \lambda_2(M) + \cdots + \lambda_q(M).$$

a) Demuestre que $f_1(M) = \sup_x[\,\langle Mx, x\rangle \;:\; x \in \mathrm{I\!R}^n,\; \|x\| = 1\,]$. Deduzca que esta función es convexa en $\mathcal{M}$.

b) Generalice el resultado para f_q, con $q = 1, \cdots, n$. Para ello, considere el conjunto

$$\mathcal{Q}_q \;= \{\, Q \in \mathrm{I\!R}^{n\times q} \;:\; (Q^tQ)^2 = Q^tQ\}.$$

Demuestre que

$$f_q(M) = \sup_Q[\,\mathrm{tr}\,(\mathrm{Q^tMQ}) \;:\; \mathrm{Q} \in \mathcal{Q}_\mathrm{q}\,].$$

Solución. a) Sean $v_1, \cdots, v_n$ autovectores propios ortonormales asociados a los valores propios $\lambda_i(M)$, $i = 1, \cdots, n$. El supremo se alcanza en $x = v_1$. La función f_1 es convexa como supremo de funciones lineales en M.

b) Existe una matriz ortonormal P $(P^tP = I)$ y una matriz diagonal Λ donde los elementos de la diagonal son los $\lambda_i(M)$, tal que $P^t\Lambda P = M$.

Sea $Q \in \mathcal{Q}_q$. Tenemos :

— $Q^tMQ = Q^tP^t\Lambda PQ$, y

— Para $R = PQ$, se cumple $R^tR = Q^tQ$, lo que implica que $R \in \mathcal{Q}_q$.

Note que $Q = P^tR$. Por lo tanto,

$$\begin{aligned} A &:= \sup_Q[\mathrm{tr}\,(\mathrm{Q^tMQ}) : \mathrm{Q} \in \mathcal{Q}_\mathrm{q}], \\ &= \sup_R[\mathrm{tr}\,(\mathrm{R^t\Lambda R}) : \mathrm{R} \in \mathcal{Q}_\mathrm{q}], \\ &= \sup_R\,[\mathrm{tr}\,(\mathrm{RR^t\Lambda}) : \mathrm{R} \in \mathcal{Q}_\mathrm{q}]. \end{aligned}$$

De otro lado, debido a que $(RR^t)^2 - RR^t = 0$, los valores propios de RR^t son 0 ó 1. Además, debido a que esta matriz tiene a lo más q valores propios no nulos, entonces $\operatorname{tr}(RR^t) \leq q$. Sea $\alpha_i = (RR^t)_{ii}$, entonces $0 \leq \alpha_i \leq 1$ para todo i y $\sum \alpha_i \leq q$. Se tiene que

$$A \leq \sum_{i=1}^{q} \lambda_i(M) = f_q(M).$$

Eligiendo R tal que $(RR^t)_{ii} = 1$ para todo i, y $(RR^t)_{ij} = 0$ para los otros i, j, se obtiene $A = f_q(M)$.

La función f_q es convexa como supremo de funciones afines. ■

Ejercicio 15 (Convexificación de funciones cuasiconvexas) *Una función f es cuasiconvexa en un conjunto convexo C si para todo $x, y \in C$ y todo $t \in [0,1]$, se cumple $f(tx + (1-t)y) \leq \max[f(x), f(y)]$. Se puede verificar fácilmente que si f es convexa sci propia sobre C y k una función continua estrictamente creciente, la función $g = k \circ f$ es cuasiconvexa sci. Recíprocamente, una función f cuasiconvexa es llamada convexificable si existe k continua estrictamente creciente tal que la función $g = k \circ f$ es convexa. El problema es saber si una función cuasiconvexa es convexificable.*

1) Demuestre que f es cuasiconvexa si y solo si para todo $\lambda \in \mathbb{R}$, los conjuntos de nivel $S(\lambda) = \{x \in C : f(x) \leq \lambda\}$ son convexos. Lo mismo se cumple para los conjuntos de nivel estricto $\widetilde{S}(\lambda) = \{x \in C : f(x) < \lambda\}$.

2) Sea f definida en el conjunto $C =]0, \infty[^2$ por $f(x_1, x_2) = x_1^{-1} x_2$. Demuestre que f es cuasiconvexa y que no existe ninguna función k dos veces derivable estrictamente creciente tal que la función $g = k \circ f$ sea convexa, ni siquiera localmente. Es más, tampoco existe una función continua estrictamente creciente que convexifica a f ni siquiera localmente. La demostración para este último caso es más difícil.

Solución. 1) Supongamos f cuasiconvexa. Para todo $\mu \in \mathbb{R}$, $x, y \in \widetilde{S}(\mu)$ y $t \in (0,1)$, se tiene $f(x + t(y-x)) \leq \max[f(x), f(y)] < \mu$, y por lo tanto $x + t(y-x) \in \widetilde{S}(\mu)$. Así, $\widetilde{S}(\mu)$ es convexo para todo μ. De este se deduce la convexidad de $S(\mu)$ debido a que $S(\mu) = \cap_{\lambda>\mu}\widetilde{S}(\lambda)$. Recíprocamente, supongamos ahora que todos los subniveles $S(\lambda)$ son convexos. Sean $x, y \in C$ y $t \in [0,1]$. Para $\lambda = \max[f(x), f(y)]$, tenemos $f(tx+(1-t)y) \leq \max[f(x), f(y)]$.

2) Supongamos f convexificable en una vecindad abierta V de un punto $a \in C$ por medio de una función k estrictamente creciente. Debido a que $g = k \circ f$ es convexa, la función $\theta(t) = k \circ (f(t, a_1)) = k(ta_1^{-1})$ es convexa en una vecindad convexa de a_2. Se deduce que la función k es convexa en una vecindad convexa de $a_2a_1^{-1}$. En una vecindad convexa de (a_1, a_2), se tiene

$$\nabla g(x) = k'(t)) \begin{pmatrix} \frac{-x_2}{x_1^2} \\ \frac{1}{x_1} \end{pmatrix},$$

$$\nabla^2 g(x) = k'(t) \begin{pmatrix} \frac{2x_2}{x_1^3} + r\frac{x_2^2}{x_1^4} & \frac{-1}{x_1^2} - r\frac{x_2}{x_1^3} \\ \\ \frac{-1}{x_1^2} - r\frac{x_2}{x_1^3} & \frac{r}{x_1^2} \end{pmatrix},$$

donde $t = f(x)$ y $r = k''(t)[k'(t)]^{-1}$. Debido a que $\det(\nabla^2 g(x)) = -k'(t)x_1^{-4}$ es estrictamente negativa, la función g no es convexa.

Ejercicio 16 (CNS de convexidad de las funciones cuasiconvexas) *Sea f una función dada y $x^* \neq 0$. Considere las funciones $F_{x^*}, G_{x^*} : \mathbb{R} \to \overline{\mathbb{R}}$ definidas por*

$$F_{x^*}(\lambda) = \sup_y [\, \langle x^*, y\rangle \,:\, f(y) \leq \lambda\,], \quad G_{x^*}(t) = \inf_y [\, f(y) \,:\, \langle x^*, y\rangle \geq t\,].$$

a) *Demuestre que*

$$\{(y,\lambda) \,:\, f(y) \leq \lambda)\} \subset \bigcap_{x^*\neq 0} \{(y,\lambda) \,:\, \langle x^*, y\rangle - F_{x^*}(\lambda) \leq 0\},$$

y

$$f(x) \geq \sup_{x^*\neq 0} G_{x^*}(\langle x^*, x\rangle) \quad \textit{para todo } x \in \mathbb{R}^n.$$

b) *Si f es convexa sci propia, demuestre que la función F_{x^*} es cóncava y la función G_{x^*} convexa.*

c) *Si f es cuasiconvexa sci, demuestre que f es convexa si y solo si para todo $x^* \neq 0$, la función F_{x^*} es cóncava.*

d) *Si f es cuasiconvexa sci, demuestre que f es convexa si y solo si para todo $x^* \neq 0$, la función G_{x^*} es convexa.*

Solución.

a) Se deducen inmediatamente de la definición.

b) Considere las funciones

$$\theta(y,\lambda) = \begin{cases} \langle x^*, y \rangle & \text{si } f(y) \le \lambda, \\ -\infty & \text{en otro caso.} \end{cases} \qquad \varphi(y,t) = \begin{cases} f(y) & \text{si } \langle x^*, y \rangle \ge t, \\ +\infty & \text{en otro caso.} \end{cases}$$

Por la convexidad de f, la función θ es cóncava y la función φ es convexa. Por lo tanto, F_{x^*} es cóncava como supremo de funciones cóncavas, y G_{x^*} es convexa como ínfimo de funciones convexas.

c) Basta demostrar la condición suficiente. Sea (x,λ) tal que $f(x) > \lambda$. Entonces, x no pertenece al conjunto convexo cerrado $\{y \,:\, f(y) \le \lambda)\}$. Por el teorema de separación, existe $x^* \neq 0$ tal que

$$\langle x^*, x \rangle > \sup_y [\, \langle x^*, y \rangle \,:\, f(y) \le \lambda \,] = F_{x^*}(\lambda).$$

Se deduce que

$$\{(y,\lambda) \,:\, f(y) \le \lambda)\} = \bigcap_{x^* \neq 0} \{(x,\lambda) \,:\, \langle x^*, x \rangle - F_{x^*}(\lambda) \le 0\}.$$

Por lo tanto el epígrafo de f es convexo cerrado como intersección de convexos cerrados.

d) Basta demostrar la condición suficiente. Sea $x \in \mathbb{R}^n$ fijo. Entonces x no pertenece al conjunto convexo $A_x := \{y \,:\, f(y) < f(x)\}$. Por el teorema de separación, existe x^* tal que $\langle x^*, x \rangle > \langle x^*, y \rangle$ para todo $y \in A_x$. Por lo tanto,

$$\begin{aligned} \sup_{x^*} G_{x^*}(\langle x^*, x \rangle) &\le f(x) \\ &\le \inf_y [\, f(y) \,:\, \langle x^*, x \rangle \le \langle x^*, y \rangle \,] \\ &\le \sup_{x^*} G_{x^*}(\langle x^*, x \rangle). \end{aligned}$$

Siendo, para cada $x^* \neq 0$, la función $x \to G_{x^*}(\langle x^*, x\rangle)$ convexa, se deduce que la función f también es convexa como supremo de funciones convexas. ■

Referencias

[1] **Auslender A.,** *Optimisation : Méthodes numériques*. Masson, Paris, 1976.

[2] **Auslender A., Teboulle M.,,** *Asymptotic Cones and Functions in Optimization and Variational Inequalities*. Springer Monographs in Mathematics, Springer-Verlag, New York, 2003.

[3] **Fenchel W.,** *Convex cones, sets and functions*. Mimeographed Notes, Princeton University, 1951.

[4] **Hiriart-Urruty J.-B., Lemaréchal C.,** *Convex analysis and minimization algorithms. I. Fundamentals*. Springer-Verlag, Berlin, 1993.

[5] **Laurent P.J.,** *Approximation and Optimisation*. Herman Editions, Paris, 1972.

[6] **Moreau J.J.,** *Fonctionnelles Convexes*. Séminaire sur les équations aux dérivées partielles. Collège de France, Paris,1966.

[7] **Rockafellar, R.T.,** *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[8] **Vandernberghe J., Boyd S.,** *Semidefinite Programming*. SIAM Review, vol. 38, 49–95, 1996.