DOI: 10.22199/S07160917.1985.0010.00005

REVISTA PROYECCIONES N° 10:57-99
Diciembre 1985 - ISSN 0716-0917
Jornada Matemáticas, Agosto 1985.

ANALISIS CONVEXO Y DUALIDAD EN OPTIMIZACION

Raul Aguila F.* Fernando Paredes C.*

INTRODUCCION.

La noción de convexidad es bastante clásica. Aparentemente la primera noción (17..) de conjunto convexo se encuentra en la definición de equilibrio de un cuerpo sobre un plano horizontal: "Un cuerpo se encuentra en equilibrio sobre un plano horizontal, si la vertical que pasa por el centro de gravedad de dicho cuerpo penetra la envoltura convexa de sus puntos de apoyo". Esta definición ha sido recordada por J.J. Moreau [6], quien ha sido la persona que más ha contribuido al desarrollo de la teoría de las funciones convexas definidas en espacios vectoriales topológicos [7]. Es importante señalar también el texto de R.T. Rockafellar [8], en el cual desarrollo el análisis convexo en los espacios de dimensión finita.

* Profesor Instituto de Matemática, Facultad de Ciencias Básicas y Matemáticas, Universidad Católica de Valparaíso.

En este trabajo se desarrollan los conceptos fundamentales del análisis convexo, los cuales resultan ser muy importantes en el estudio de problemas relacionados con la minimización de las funciones convexas no diferenciables. En la última sección se hace una pequeña introducción a la teoría de la dualidad en optimización.

1. FUNCIONES CONVEXAS.

$\overline{\mathbb{R}}$ denotará la recta real extendida, es decir:

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}.$$

En $\overline{\mathbb{R}}$ consideraremos la relación de orden natural que prolonga el orden en $\mathbb{R}$, es decir:

$$\forall\, a \in \overline{\mathbb{R}} : \quad -\infty \leq a \leq +\infty.$$

Además, se extiende a $\overline{\mathbb{R}}$ la adición de $\mathbb{R}$ de la manera siguiente:

$$\forall\, a \in [-\infty,+\infty[\; : \quad a + (-\infty) = (-\infty) + a = (-\infty) \quad \text{y}$$

$$\forall\, a \in \overline{\mathbb{R}} \quad : \quad a + (+\infty) = (+\infty) + a = (+\infty)$$

Es importante hacer notar que con esta convención se tiene:

$$(+\infty) + (-\infty) = (-\infty) + (+\infty) = (+\infty).$$

Se verifica fácilmente que la adición en $\overline{\mathbb{R}}$ es conmutativa y asociativa, además:

$$\forall\ x,y \in \overline{\mathbb{R}},\ \forall\ \lambda,u \in\ ]0,+\infty[\ :$$

$$\lambda(x + y) = \lambda x + \lambda y$$

$$(\lambda + u)x = \lambda x + ux$$

Se tiene además que esta adición es compatible con el orden en $\overline{\mathbb{R}}$. Si $x,y,z \in \overline{\mathbb{R}}$ entonces $x + y \leqslant z$ si y sólo si $\exists\ x',y' \in \overline{\mathbb{R}}$ con $x' \geqslant x$ y $y' \geqslant y$ tal que $x' + y' = z$.

Si $\{u_i\}_{i\in I}$ y $\{v_j\}_{j\in J}$ son dos familias cualesquiera de elementos de $\overline{\mathbb{R}}$, se tiene que:

$$\operatorname*{Inf}_{(i,j)\in I\times J} (u_i + v_j) = \operatorname*{Inf}_{i\in I} u_i + \operatorname*{Inf}_{j\in J} v_j$$

$$\operatorname*{Sup}_{(i,j)\in I\times J} (u_i + v_j) \leqslant \operatorname*{Sup}_{i\in I} u_i + \operatorname*{Sup}_{j\in J} v_j .$$

Sea E un espacio vectorial sobre el cuerpo $\mathbb{R}$, entonces $\overline{\mathbb{R}}^E$ denotará el conjunto de todas las funciones definidas en E y a valores en $\overline{\mathbb{R}}$.

<u>Definición 1.1.</u>

Sea $C \subseteq E$, C se dirá convexo en E si dados dos puntos cualesquiera en C, entonces el trazo que los une está completamente contenido en C; es decir:

$$x,y \in C \Rightarrow \lambda x + (1 - \lambda)y \in C\ ,\quad \forall\ \lambda \in [0,1].$$

Definición 1.2.

Diremos que una función $f \in \overline{\mathbb{R}}^E$ es convexa si su epígrafo:

$$\text{epi}\ (f) := \{(x,\lambda) \in E \times \mathbb{R}\ /\ f(x) \leq \lambda\}$$

es un conjunto convexo en $E \times \mathbb{R}$. Conv(E) denotará el Conjunto: $\{f \in \overline{\mathbb{R}}^E\ /\ f \text{ convexa}\}$.

Una definición equivalente de función convexa es la siguiente :

Teorema 1.3.

Una función $f \in \overline{\mathbb{R}}^E$ es convexa si y sólo si

$x,y \in E \Longrightarrow f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)\ ,\quad \forall\lambda \in [0,1]$.

Definición 1.4.

Dado $Q \subseteq E$, llamaremos función indicatriz de Q a la función $\chi_Q \in \overline{\mathbb{R}}^E$, definida por:

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

Definición 1.5.

Sea $C \subseteq E$, diremos que una función $f \in \overline{\mathbb{R}}^E$ es convexa en C, si la función: $\tilde{f} = f + \chi_C$ es convexa.

Notemos que el epígrafo de $\tilde{f}$ es igual a la intersección en $E \times \mathbb{R}$ del epígrafo de f con $C \times \mathbb{R}$.

Observación 1.6.

Si $f \in \overline{\mathbb{R}}^E$ es convexa, entonces los conjuntos de nivel $C_\lambda = \{x \in E \,/\, f(x) \leq \lambda\}$ son convexos $\forall\, \lambda \in \overline{\mathbb{R}}$. El recíproco no es verdadero, por ejemplo se puede considerar $f(x) = \sqrt{|x|}$, $\forall\, x \in \mathbb{R}$.

Definición 1.7.

$$\mathrm{dom}(f) : \{x \in E \,/\, f(x) < +\infty\}$$

Se llamará dominio efectivo de f.

Dos propiedades importantes de las funciones convexas son:

1°) Si f y g son dos funciones convexas entonces $f + g$ es una función convexa y λf también lo es si $\lambda \in\,]0,+\infty[$.

2°) Si $\{f_i\}_{i\in I}$ es una familia de funciones convexas entonces la función $f := \sup\limits_{i\in I} f_i$ es una función convexa.

La segunda propiedad es una consecuencia inmediata del hecho que el epígrafo de f es igual a la intersección de los epígrafos de los f_i y al hecho que la intersección de conjuntos convexos es un convexo.

Definición 1.8.

$f \in \overline{\mathbb{R}}^E$ se dirá propia si no es idénticamente igual a $+\infty$ y no toma jamás el valor $-\infty$, es decir, si su dominio efectivo es no vacío y si f es finita sobre este dominio.

Nos interesará trabajar con funciones propias, pues las funcio nes convexas que toman el valor $-\infty$ son de naturaleza bien especial, por ejemplo, sea $E = \mathbb{R}^2$ y $f \in \overline{\mathbb{R}}^E$ definida mediante:

$$f(x_1,x_2) = \begin{cases} -\infty & \text{si } x_1^2 + x_2^2 < 1 \\ 0 & \text{si } x_1^2 + x_2^2 = 1 \\ +\infty & \text{si } x_1^2 + x_2^2 > 1 \end{cases}$$

2. SEMICONTINUIDAD Y CONTINUIDAD DE FUNCIONES CONVEXAS.

En toda esta sección, E será un espacio vectorial topológico separado.

Definición 2.1.

$f \in \overline{\mathbb{R}}^E$ se dirá semicontinua inferior (s.c.i.) en $x_0 \in E$, si para todo k que verifica $k < f(x_0)$, existe una vecindad U de x_0 tal que $\forall\, x \in U$, se tiene $k < f(x)$. $f \in \overline{\mathbb{R}}^E$ se dirá s.c.i., si es s.c.i. en cada punto de E.

Notemos que f es s.c.i. en todo punto x_0 tal que $f(x_0) = -\infty$.

Teorema 2.2.

$f \in \overline{\mathbb{R}}^E$ es s.c.i. si y sólo si $\forall\, \lambda \in \mathbb{R}$, $S_\lambda = \{x \in E \;/\; f(x) \leqslant \lambda\}$ es un conjunto cerrado.

Demostración:

i) Demostremos que S_λ es cerrado.

Sea $x_0 \in S_\lambda^c$, luego $f(x_0) > \lambda$ y entonces por hipótesis existe una vecindad U de x_0 tal que $f(x) > \lambda$, $\forall\, x \in U$ $\therefore$ $U \subseteq S_\lambda^c$ y en consecuencia S_λ^c es abierto, es decir, S_λ es cerrado.

ii) Demostremos ahora que f es s.c.i.

Supongamos que S_λ es cerrado $\forall\, \lambda \in \mathbb{R}$.

Sea $k < f(x_0)$, por lo tanto $x_0 \notin S_k$ y como S_k es un cerrado se tiene que $\exists$ U vecindad de x_0 tal que $U \cap S_k = \phi$ y en consecuencia $\forall\, x \in U$ se tiene $f(x) > k$, lo que demuestra que f es s.c.i., pues x_0 es arbitrario. #

A continuación tenemos la siguiente consecuencias inmediata del teorema 2.2.

Corolario 2.3.

$Q \subseteq E$, es cerrado si y sólo si χ_Q es s.c.i.

Teorema 2.4.

$f \in \overline{\mathbb{R}}^E$ es s.c.i. si y sólo si $\mathrm{epi}(f)$ es cerrado en $E \times \mathbb{R}$.

Demostración:

i) Supongamos que $\mathrm{epi}(f) = \{(x,\lambda) \in E \times \mathbb{R} \ / \ f(x) \leqslant \lambda\}$ es cerrado.

Sea $P_u = \{(x,\lambda) \in E \times \mathbb{R} \,/\, \lambda = u\}$, es claro que P es un cerrado en $E \times \mathbb{R}$ (Por ser E separado).

El conjunto: $Epi(f) \cap P_u = \{(x,\lambda) \in E \times \mathbb{R} \,/\, \lambda = u \wedge f(x) \leqslant \lambda\}$ es cerrado y está en correspondencia bicontinua con:

$S_u = \{x \in E \,/\, f(x) \leqslant u\}$. Luego S_u es cerrado $\forall\, u \in \mathbb{R}$ y aplicando el Teorema 2.2. se tiene que f es s.c.i.

ii) Supongamos que f sea s.c.i., es decir, que S_λ es cerrado $\forall \lambda \in \mathbb{R}$.

Sea $(x_0,\lambda_0) \in \overline{epi(f)}$, luego para toda vecindad V_{x_0} de x_0 y para toda vecindad V_{λ_0} de λ_0 , existen $x \in V_{x_0}$ y $\lambda \in V_{\lambda_0}$ tales que: $f(x) \leqslant \lambda$, es decir , $(x,\lambda) \in epi(f)$.

Sea $\varepsilon > 0$ y $V_{\lambda_0} = \,]\lambda_0 - \varepsilon, \lambda_0 + \varepsilon[$. Luego $\forall\, \varepsilon > 0$ y para toda V_{x_0} existen $x \in V_{x_0}$ y λ que verifica $|\lambda - \lambda_0| < \varepsilon$ tales que $f(x) \leqslant \lambda_0 + \varepsilon \;\therefore\; x \in S_{\lambda_0 + \varepsilon} = \overline{S_{\lambda_0 + \varepsilon}}$.

Luego, $\forall\, \varepsilon > 0$, $x_0 \in S_{\lambda_0 + \varepsilon}$, es decir, $\forall\, \varepsilon > 0$, $f(x_0) \leqslant \lambda_0 + \varepsilon \;\therefore\; f(x_0) \leqslant \lambda_0$.

En consecuencia, $(x_0,\lambda_0) \in epi(f)$, lo que demuestra que $epi(f)$ es cerrado en $E \times \mathbb{R}$. #

Observación 2.5.

Si $f \in Conv(E)$ es s.c.i. y f toma el valor $-\infty$ entonces necesariamente los únicos valores que puede tomar f son $+\infty$ y $-\infty$.

En forma más precisa f toma el valor $-\infty$ en un convexo cerrado y $+\infty$ en el complemento de este convexo. Ver ej. a continuación de la definición 1.8.

Definición 2.6.

Llamaremos regularización s.c.i. de una función $f \in \overline{\mathbb{R}}^E$ a la mayor función s.c.i. que minora a f. Se denotará $\bar{f}$.

Observación 2.7.

$\bar{f}$ es la función perteneciente a $\overline{\mathbb{R}}^E$ cuyo epígrafo es el menor cerrado que contiene el epígrafo de f, es decir:

$$\operatorname{epi}(\bar{f}) = \overline{\operatorname{epi}(f)}.$$

Teorema 2.8.

La regularización s.c.i. $\bar{f}$ de $f \in \overline{\mathbb{R}}^E$ es igual a:

$$f(x) = \underline{\lim}_{y \to x} f(y).$$

Se recuerda que se tiene por definición:

$$\underline{\lim}_{y \to x} f(y) = \sup_{V \in F(x)} \ \inf_{y \in V} f(y) \ ,$$ donde $F(x)$ es la familia de todas las vecindades de x.

Demostración:

i) Demostraremos primero que la función: $x \to \underline{\lim}_{y \to x} f(y)$ es s.c.i.

En efecto, sea $k < \underset{y \to x}{\underline{\lim}} f(y)$, entonces $\exists\, V \in F(x)$ tal que $k < \inf_{y \in V} f(y)$. Sea $U \subset V$, U abierto tal que $x \in U$. Entonces $k < \inf_{y \in U} f(y)$ y además se tiene:

$\forall\, x' \in U$: $k < \sup_{V \in F(x')} \inf_{y \in V} f(y) = \underset{y \to x'}{\underline{\lim}} f(y)$, lo que demuestra la s.c.i. de la función: $x \to \underset{y \to x}{\underline{\lim}} f(y)$

ii) Si g es s.c.i. entonces $g(x) \leq \underset{y \to x}{\underline{\lim}} g(y)$, en efecto:

$\forall\, k < g(x)$, existe una vecindad V de x tal que $\forall\, y \in V$ se tiene $k < g(y)$, luego:

$$k \leq \inf_{y \in V} g(y) \leq \sup_{V \in F(x)} \inf_{y \in V} g(y)$$

$\therefore$ $\forall\, k < g(x)$ se tiene: $k \leq \underset{y \to x}{\underline{\lim}} g(y)$, luego $g(x) \leq \underset{y \to x}{\underline{\lim}} g(y)$

iii) Demostraremos por último que si $g(x) \leq f(x)$, $\forall\, x \in E$, y si g es s.c.i. entonces $g(x) \leq \underset{y \to x}{\underline{\lim}} f(y)$, $\forall\, x \in E$, es decir, la función $x \to \underset{y \to x}{\underline{\lim}} f(y)$ es la mayor función s.c.i. que minora a f. En efecto,

$$\inf_{y \in V} g(y) \leq \inf_{y \in V} f(y) \ , \ \forall\, V \in F(x) \ , \ \text{luego:}$$

$$\sup_{V \in F(x)} \inf_{y \in V} g(y) \leq \sup_{V \in F(x)} \inf_{y \in V} f(y)$$

$\therefore$ $\underset{y \to x}{\underline{\lim}} g(y) \leq \underset{y \to x}{\underline{\lim}} f(y)$, $\forall\, x \in E$ y aplicando ii) se tiene que:

$$g(x) \leq \underset{y \to x}{\underline{\lim}} f(y). \qquad \#$$

Corolario 2.9.

$$f \in \overline{\mathbb{R}}^E \quad \text{es s.c.i. si y sólo si} \quad f(x) = \underset{y \to x}{\underline{\text{lím}}}\, f(y).$$

Se define en forma análoga la semicontinuidad superior (s.c.s) de $f \in \overline{\mathbb{R}}^E$. La función $f \in \overline{\mathbb{R}}^E$ es s.c.s. en $x_0 \in E$ si para todo $k > f(x_0)$, existe una vecindad V de x_0 tal que $\forall\, x \in V$, $k > f(x)$. En forma equivalente f es s.c.s. en x_0 si $-f$ es s.c.i. en x_0.

Las siguientes proposiciones son equivalentes:

i) F es s.c.s.

ii) $\forall \lambda \in \mathbb{R}$, $S'_\lambda = \{x \in E \;/\; f(x) < \lambda\}$ es abierto

iii) $f(x) = \overline{\underset{y \to x}{\text{lím}}}\, f(y)$, $\forall\, x \in E$, donde:

$$\underset{y \to x}{\text{lím}}\, f(y) = \underset{V \in F(x)}{\text{Inf}} \;\; \underset{y \in V}{\text{Sup}}\; f(y).$$

Definición 2.10.

$f \in \overline{\mathbb{R}}^E$ se dice continua en x_0 si f es s.c.i. y s.c.s. en x_0. Además f se dice continua si es continua en todo punto $x_0 \in E$.

Teorema 2.11.

Sea $f \in \text{Conv}(E)$. Si existe un abierto no vacío de E sobre el cual f es acotada entonces f es continua sobre el $\text{Int}(\text{Dom}(f))$, evidentemente no vacío.

Corolario 2.12.

$f \in Conv(E)$ es continua en todo punto del interior (no vacío) de su dominio efectivo si y sólo si existe un punto de este dominio donde es continua (o bien solamente s.c.s.).

Demostración:

i) Supongamos $f \in Con(E)$ y s.c.s. en $x_0 \in dom(f)$. Entonces $f(x_0) < +\infty$ y $\forall\, k > f(x_0)$, existe una vecindad U de x_0 tal que $\forall\, y \in U \;\; k > f(y)$. Luego f es acotada en el abierto $U \neq \phi$. Luego por teorema 2.11. f es continua en todo punto del Int(dom(f)).

ii) Si f es continua en todo punto del Int(dom(f)) supuesto no vacío, entonces existe evidentemente un punto del dom(f) donde f es s.c.s.

#

Ejemplos de funciones convexas que no son s.c.i.

Ejemplo 1:

Si $Q \subsetneq E$ es tal que $\overline{Q} \neq Q$ entonces χ no es s.c.i., ya que $\overline{\chi_Q} = \chi_{\overline{Q}}$.

Ejemplo 2:

Sea $E = \mathbb{R}^2$ y $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ la función definida inmediatamente después de la definición 1.8., entonces

$$\overline{f}(x_1,x_2) = \begin{cases} -\infty & \text{si} \quad x_1^2 + x_2^2 \leq 1 \\ \\ +\infty & \text{si} \quad x_1^2 + x_2^2 > 1. \end{cases}$$

Nota:

En ambos ejemplos para calcular la regularización s.c.i., resulta más fácil calcular la adherencia del epígrafo de la función dada.

3. Γ - REGULARIZADA Y POLAR DE UNA FUNCION.

Consideremos ahora X, Y dos espacios vectoriales (reales) en dualidad mediante una forma bilineal en $X \times Y$ que denotaremos $\langle .,. \rangle$, es decir, dado $x \in X$ no nulo existe $y \in Y$ tal que $\langle x,y \rangle \neq 0$ y dado $y \in Y$ no nulo existe $x \in X$ tal que $\langle x,y \rangle \neq 0$.

Una topología localmente convexa sobre X (resp. sobre Y) se dirá compatible con la dualidad entre X e Y si la función lineal: $x \in X \to \langle x,y \rangle \in \mathbb{R}$ (resp. $y \in Y \to \langle x,y \rangle \in \mathbb{R}$) es continua cualquiera que sea $y \in Y$ (resp. $x \in X$) y si toda función lineal continua de X en $\mathbb{R}$ (resp. de Y en $\mathbb{R}$) se puede representar de esta manera. En este caso diremos que el dual de X es Y, es decir $X^* = Y$, resp. $(Y^* = X)$. En todo lo que sigue X, Y serán dos espacios vectoriales topológicos localmente convexos en dualidad (es decir, sus topologías serán compatibles con la dualidad).

Para simplificar, en todo lo que sigue el lector podrá suponer $X = Y = \mathbb{R}^n$; $\langle .,. \rangle$ el producto punto en $\mathbb{R}^n$ y recordar que sobre $\mathbb{R}^n$ hay una sola topología de espacio vectorial que sea separada, esta es la topología usual (la inducida por la norma euclideana).

De entre todas las posibles topologías localmente convexas que se pueden definir sobre X (resp. sobre Y) y que son compatibles con la dualidad entre X e Y, nos interesará distinguir dos de ellas:

La menos fina, que llamaremos topología débil y que denotaremos $\sigma(X,Y)$ (resp. $\sigma(Y,X)$), y la más fina que llamaremos topología fuerte y que denotaremos $\tau(X,Y)$ (resp. $\tau(Y,X)$). Insistimos en el hecho de que si $X = Y = \mathbb{R}^n$ entonces $\sigma(X,Y) = \tau(X,Y)$, topología usual.

Se ve claramente que las funcionales lineales continuas, las funcionales afines continuas sobre X, es decir, las funciones de la forma $x \in X \to \langle x,y\rangle - r$, con $y \in Y$, $r \in \mathbb{R}$, son las mismas cualesquiera sean las topologías localmente convexas sobre X compatibles con la dualidad entre X e Y. Lo mismo ocurre con los hiperplanos cerrados en X, es decir con los conjuntos de la forma: $H = \{x \in X \,/\, \langle x,y\rangle = r\}$, también lo mismo ocurre con los semiespacios cerrados en X, es decir con los conjuntos de la forma: $D = \{x \in X \,/\, \langle x,y\rangle \leq r\}$, con $y \in Y$, $r \in \mathbb{R}$.

A continuación se tiene un importante teorema de caracterización de los convexos cerrados en X.

Teorema 3.1.

Sean X e Y dos e.v.t.l.c. en dualidad. Todo convexo cerrado de X es igual a la intersección de los semiespacios cerrados que lo contienen.

Demostración:

Sea C un convexo cerrado de X.

Sea $K = \bigcap_{D \supseteq C} D$, la intersección de todos los semiespacios cerrados D que contienen al convexo C. Es evidente que $C \subseteq K$. Para demostrar que $K \subseteq C$, supongamos lo contrario, es decir, que existe $x_0 \in K$ tal que $x_0 \notin C$. Como C es un convexo cerrado, entonces del teorema de Hahn-Banach, sabemos que existe un hiperplano cerrado que separa el punto x_0 del conjunto C, es decir, $\exists\ y \in Y$, $c \in \mathbb{R}$ tal que:

$$\langle x,y\rangle \leq c \quad , \quad \forall\, x \in C$$

$$\langle x_0,y\rangle > c.$$

Vemos entonces que existe un semiespacio cerrado:

$D_C = \{x \in X \ / \ \langle x,y\rangle \leq c\}$ que contiene a C y no contiene al punto $x_0 \in K$, esto constituye una contradicción. #

Observación 3.2.

De este teorema y del hecho que los semiespacios cerrados en X son los mismos cualquiera que sea la topología localmente convexa sobre X compatible con la dualidad, concluímos que los convexos cerrados serán también los mismos y por lo tanto, las funciones $f \in \overline{\mathbb{R}}^X$ convexas y s.c.i. son las mismas cualquiera que sea la topología localmente convexa sobre X compatible con la dualidad (Si $X = Y = \mathbb{R}^n$, está observación carece de sentido).

Definición 3.3.

$\Gamma(X)$ denotará el conjunto de todas las funciones $f \in \overline{\mathbb{R}}^X$ que son envoltura superior de una familia de funciones afines continuas en X. Es decir, debe existir una familia $\{(y_i,r_i)\}_{i\in I}$ en $Y \times \mathbb{R}$, tal que:

$$f(x) = \sup_{i \in I} (\langle x,y_i\rangle - r_i).$$

Puesto que las funciones $w_1(x) \equiv +\infty$ y $w_1(x) \equiv -\infty$ pertenecen a $\Gamma(X)$, será cómodo definir:

$$\Gamma_0(X) = \Gamma(X) \setminus \{w_1, w_2\}.$$

Es importante notar que el conjunto $\Gamma(X)$ sólo depende de la dualidad entre X e Y , es decir, es el mismo cualquiera que sea la topología en X compatible con la dualidad.

Como todo supremo de funciones convexas es una función convexa y todo supremos de funciones s.c.i. es una función s.c.i., entonces las funciones de $\Gamma(X)$ serán convexas y s.c.i.

El siguiente teorema caracteriza fácilmente al conjunto $\Gamma_0(X)$.

Teorema 3.4.

El conjunto $\Gamma_0(X)$ es exactamente el conjunto de las funciones $f \in \overline{\mathbb{R}}^X$ que son convexas, s.c.i. y propias.

Definición 3.5.

Llamaremos Γ- regularización de una función $f \in \overline{\mathbb{R}}^X$ a la mayor función en $\Gamma(X)$ que minora f. La denotaremos f^Γ. Es decir,

$$f^\Gamma(x_0) = \sup_{(y,r)\in I} (\langle x_0, y\rangle - r)$$

donde $I = \{(y,r) \in Y \times \mathbb{R} \ / \ \langle x,y\rangle - r \leq f(x) \ , \ \forall\, x \in X\}$

Como las funciones en $\Gamma(X)$ son s.c.i. se tiene la desigualdad:

$$f^\Gamma(x) \leq \overline{f}(x) \leq f(x) \quad , \quad \forall\, x \in X$$

Insistamos una vez más que todas las definiciones o teoremas que se dan relativos a X pueden ser reformulados de manera análoga relativos a Y.

Definición 3.6.

Dada $f \in \overline{\mathbb{R}}^X$, se define la polar de f (o función dual) que denotaremos f^*, mediante:

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

Es claro que $f^* \in \overline{\mathbb{R}}^Y$. Si f es propia, se tiene que:

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

es decir, f^* es la envoltura superior de una familia (no vacía) de funciones afines continuas sobre Y, lo que muestra que $f^* \in \Gamma_0(Y)$.

Teorema 3.7.

$\forall\, f \in \overline{\mathbb{R}}^X$, se tiene $f^\Gamma = f^{**}$.

Demostración:

Una función afín continua $p : x \in X \to \langle x,y\rangle - r$ minora a f si y sólo si $\langle x,y\rangle - r \leqslant f(x)$, $\forall\, x \in X$, es decir, si y sólo si
$r \geqslant \sup_{x \in X} (\langle x,y\rangle - f(x)) = f^*(y)$

$$\therefore \quad f^\Gamma(x) = \sup_{\substack{y \in Y \\ r \in \mathbb{R} \\ r \geqslant f^*(y)}} (\langle x,y\rangle - r)$$

Supongamos primero que f^* no toma el valor $-\infty$, luego se obtiene el mismo resultado si reemplazamos $r \geq f^*(y)$ por $r = f^*(y)$ y entonces tenemos:

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

Si existe $y_0 \in Y$ tal que $f^*(y_0) = -\infty$, esto significa que se tiene de hecho $f^*(y) \equiv -\infty$, $y \in Y$ y $f(x) \equiv +\infty$, $x \in X$. Pero en este caso también se tiene:

$$f^{\Gamma}(x) \equiv f^{**}(x) \equiv +\infty \ , \ x \in X. \qquad \#$$

<u>Corolario 3.8.</u>

Sea $f \in \overline{\mathbb{R}}^X$ entonces $f^{**} = f$ si y sólo si $f \in \Gamma(x)$.

<u>Demostración</u>:

i) Si $f \in \Gamma(x)$ entonces $f^{\Gamma} = f$, por lo tanto $f = f^{**}$.

ii) Si $f^{**} = f$ entonces $f^{\Gamma} = f$, por lo tanto $f \in \Gamma(x)$.

<u>Teorema 3.9.</u>

Propiedades de la Polar de una función.

a) Si $f_1 \leq f_2$ (es decir : si $f_1(x) \leq f_2(x)$, $\forall\, x \in X$) entonces $f_1^* \geq f_2^*$.

b) Si $\{f_i\}_{i \in I}$ es una familia de funciones pertenecientes a $\overline{\mathbb{R}}^X$ entonces

$$(\inf_{i \in I} f_i)^* = \sup_{i \in I} f_i^*$$

$$(\sup_{i \in I} f_i)^* \leq \inf_{i \in I} f_i^*$$

c) Si $f \in \overline{\mathbb{R}}^X$ y $g(x) = f(x - a)$ con $a \in X$ fijo, $\forall x \in X$ entonces: $g^*(y) = f^*(y) + \langle a,y\rangle$, $y \in Y$

d) Si $f \in \overline{\mathbb{R}}^X$ y $g(x) = f(\lambda x)$, $\lambda \in \mathbb{R} \setminus \{0\}$ fijo, $\forall x \in X$ entonces $g^*(y) = f^*(y/\lambda)$, $y \in Y$.

e) Si $f \in \overline{\mathbb{R}}^X$ y $g(x) = uf(x)$, $u \in \mathbb{R}^+$ fijo entonces:

$$g^*(y) = uf^*(y/u) \ , \ y \in Y$$

Sea $f_u : x \in X \to uf(x/u)$, entonces:

$$(uf)^* = (f^*)_u \quad \text{y}$$

$$(f_u)^* = uf^*$$

Definición 3.10.

Se dice que una función f es positivamente homogénea si:

$$f(ux) = uf(x) \quad \forall x \in X \quad \text{y} \quad \forall u > 0 .$$

Si f es positivamente homogénea entonces $f^* = uf^*$, $\forall u > 0$. Vemos en este caso que f^* puede tomar sólo los valores $-\infty, 0$ y $+\infty$.

Si f^* toma el valor $-\infty$ entonces $f^* \equiv -\infty$ y esto corresponde al caso en que $f \equiv +\infty$. Si se exceptúa este caso entonces f^* puede tomar sólo los valores 0 y $+\infty$. Como $f^* \in \Gamma(Y)$, entonces f^* es la función indicatriz de un convexo cerrado de Y.

De igual modo, si f es la función indicatriz χ_A de una parte cualquiera A de X, se tiene $f = uf$, $\forall\, u > 0$ y en consecuencia $(f^*)_u = f^*$, $\forall\, u > 0$, lo que significa, que $f^* \in \Gamma(Y)$ es una función positivamente homogénea.

Luego tenemos : $\chi_A^*(y) = \sup\limits_{x \in A} \langle x,y\rangle \cdot \chi_A^*$ se llama la función de apoyo de A. Esta función juega un rol importante en el análisis convexo.

Ejemplos:

a) $X = Y = \mathbb{R}$; $f(x) = 1/2\, x^2$, $\forall\, x \in \mathbb{R}$,

$$f^*(y) = \sup_{x \in \mathbb{R}} (xy - 1/2\, x^2) = \sup_{x \in \mathbb{R}} (1/2\, y^2 - 1/2(x-y)^2)$$

$$= 1/2\, y^2 \quad , \quad y \in \mathbb{R}.$$

b) $X = Y = \mathbb{R}$; $f(x) = \frac{1}{p} |x|^p$, $x \in \mathbb{R}$ (con $1 < p < +\infty$).

Se obtiene $f^*(y) = \frac{1}{q} |y|^q$, $y \in \mathbb{R}$ (con $\frac{1}{p} + \frac{1}{q} = 1$).

c) $X = Y = \mathbb{R}$:

$$f(x) = \begin{cases} -\sqrt{r^4 - x^4} & \text{si } |x| \leq r \quad (\text{donde } r > 0) \\ +\infty & \text{si } |x| > r. \end{cases}$$

Se obtiene: $f^*(y) = r\sqrt{1 + y^2}$, $y \in \mathbb{R}$.

d) $X = Y = \mathbb{R}$; $f(x) = |x|$, $x \in \mathbb{R}$

Como f es positivamente homogénea entonces f^* será una función indicatriz de un convexo cerrado de Y:

$$f^*(y) = \begin{cases} 0 , & \text{si } |y| \leq 1 \\ +\infty , & \text{si } y > 1. \end{cases}$$

Terminaremos esta sección mostrando la relación importante que existe entre la continuidad en un punto de una función convexa y una cierta propiedad de inf-compacidad de su función polar.

Definición 3.11.

Una función $g \in \overline{\mathbb{R}}^Y$ se dirá inf-compacta si $\forall\, \lambda \in \mathbb{R}$ los conjuntos : $\{y \in Y \,/\, g(y) \leq \lambda\}$ son compactos.

Teorema 3.12.

Una función $f \in \Gamma_0(X)$ es finita y continua en $0 \in X$ para la topología $\tau(X,Y)$ si y sólo si la polar f^* es inf-compacta para la topología débil $\sigma(Y,X)$.

Observación 3.13.

Si $f \in \mathrm{Con}(X)$ es finita y $\tau(X,Y)$- continua en $0 \in X$ entonces f^* es $\sigma(Y,X)$- inf-compacta.

Definición 3.14.

$g \in \overline{\mathbb{R}}^Y$ se dirá inf-compacta para la pendiente $x_0 \in X$ si la función g_0 definida por $g_0(y) := g(y) - \langle x_0, y\rangle$ es inf-compacta, es decir, si $\forall\, \lambda \in \mathbb{R}$, los conjuntos: $\{y \in Y \,/\, g(y) - \langle x_0, y\rangle \leq \lambda\}$ son compactos.

Corolario 3.15.

$f \in \Gamma_0(X)$ es finita y $\tau(X,Y)$-còntinua en $x_0 \in X$ si y sólo si la función polar f^* es $\sigma(Y,X)$-inf-compacta para la pendiente x_0.

Demostración:

Definamos $\tilde{f}(x) := f(x + x_0)$, luego por teorema 3.9. c) tenemos que:

$$\tilde{f}^*(y) = f^*(y) - \langle x_0, y\rangle$$

Como $\tilde{f} \in \Gamma_0(X)$ es finita y $\tau(X,Y)$-continua en cero entonces por teorema 3.12. esto es equivalente a que la polar $\tilde{f}^*$ es $\sigma(Y,X)$-inf-compacta, es decir, f^* es $\sigma(Y,X)$-inf-compacta para la pendiente x_0. #

Observación 3.16.

a) Si $f \in \mathrm{Conv}(X)$ es finita y $\tau(X,Y)$-continua en $x_0 \in X$, entonces f^* es $\sigma(Y,X)$-inf-compacta para la pendiente x_0.

b) Sea g es convexa y s.c.i. y supongamos que $\exists\, \lambda_0 \in \mathbb{R}$ tal que $S_{\lambda_0} = \{y \in Y \,/\, g(y) \leq \lambda_0\}$ es compacto e $\inf_{y \in Y} g(y) < \lambda_0$ entonces g es inf-compacta.

4. SUB-DIFERENCIAL Y DERIVADA DIRECCIONAL.

Definición 4.1.

Dada $f \in \overline{\mathbb{R}}^X$, se dirá que $y_0 \in Y$ es un subgradiente de f en un punto $x_0 \in X$, si $f(x_0)$ es finito y

$$f(x) \geqslant f(x_0) + \langle x - x_0, y_0 \rangle , \quad \forall\, x \in X.$$

En forma equivalente, y_0 será un subgradiente de f en x_0 si la función afín:

$$p(x) = f(x_0) + \langle x - x_0, y_0 \rangle , \quad x \in X ,$$ minora a la función f.

Definición 4.2.

El conjunto (eventualmente vacío) de subgradientes de f en x_0 se llamará el subdiferencial de f en x_0 y será denotado por $\partial f(x_0)$.

El subdiferencial de f es entonces una multiplicación que a cada $x \in X$ le asocia el subconjunto de Y, $\partial f(x)$.

Supongamos $f(x_0)$ finito. El elemento $y_0 \in Y$, es un subgradiente de f en x_0 si y sólo si se tiene:

$$f(x) - \langle x, y_0 \rangle \geqslant f(x_0) - \langle x_0, y_0 \rangle , \quad \forall\, x \in X ,$$ es decir si y sólo si $f(x_0) - \langle x_0, y_0 \rangle = \underset{x \in X}{\mathrm{Min}} \, (f(x) - \langle x, y_0 \rangle)$, es decir si y sólo si la función $x \in X \to f(x) - \langle x, y_0 \rangle \in \mathbb{R}$ realiza el mínimo sobre X en x_0.

Teorema 4.3.

$y_0 \in Y$ es un subgradiente de $f \in \overline{\mathbb{R}}^X$ en $x_0 \in X$ si y sólo si

$$f(x_0) + f^*(y_0) = \langle x_0, y_0 \rangle.$$

Demostración:

Como $\langle x_0, y_0 \rangle \in \mathbb{R}$, es claro que $f(x_0)$ y $f^*(y_0)$ son cantidades finitas.

$y_0 \in \partial f(x_0)$ si y sólo si $\langle x, y_0 \rangle - f(x) \leqslant \langle x_0, y_0 \rangle - f(x_0)$, $\forall\, x \in X$, es decir, si y sólo si

$$\sup_{x \in X} (\langle x, y_0 \rangle - f(x)) \leqslant \langle x_0, y_0 \rangle - f(x_0)$$

es decir, $f^*(y_0) \leqslant \langle x_0, y_0 \rangle - f(x_0)$; esto termina la demostración pues siempre se verifica que:

$$f^*(y_0) \geqslant \langle x_0, y_0 \rangle - f(x_0), \text{ por definición de } f^*. \qquad \#$$

Observación 4.4.

$f(x_0) + f^*(y_0) = \langle x_0, y_0 \rangle$ si y sólo si $f(x_0) + f^*(y_0) \leqslant \langle x_0, y_0 \rangle$.

Luego, se tiene la siguiente caracterización del subdiferencial:

$$\partial f(x_0) = \{y_0 \in Y \;/\; f^*(y_0) - \langle x_0, y_0 \rangle \leqslant -f(x_0)\ ,\ f(x_0) \text{ finito}\}$$

Además como $f^* \in \Gamma(Y)$, la función $y_0 \in Y \to f^*(y_0) - \langle x_0, y_0 \rangle$ es convexa y s.ċ.i., por lo tanto $\partial f(x_0)$ es un convexo cerrado (eventualmente vacío) de Y.

Teorema 4.5.

Dada la función $f \in \overline{\mathbb{R}}^X$, se tiene:

i) $f(x_0) \neq \phi \Longrightarrow f^{**}(x_0) = f(x_0)$.

ii) $f^{**}(x_0) = f(x_0) \Longrightarrow \partial f^{**}(x_0) = \partial f(x_0)$.

iii) $y_0 \in \partial f(x_0) \Longrightarrow x_0 \in \partial f^*(y_0)$.

Demostración:

i) Sea $y_0 \in \partial f(x\ )$, entonces la función afín continua:

$p(x) = f(x_0) + \langle x - x_0, y_0 \rangle$ minora a f y es tal que $p(x_0) = f(x_0)$. Como f^Γ es la envoltura superior de todos los minorantes afines continuos de f , es evidente que $f^\Gamma(x_0) = f(x_0)$ y en consecuencia $f^{**}(x_0) = f(x_0)$ (Teorema 3.7.).

ii) Puesto que $f^\Gamma = f^{**}$, se tiene que f y f^* tienen los mismos minorantes afines continuos. Si además $f^{**}(x_0) = f(x_0)$ (finito) , los minorantes afines continuos de f y f^{**} que toman este valor en x_0 son los mismos. Luego $\partial f^{**}(x_0) = \partial f(x_0)$. Si $f(x_0)$ no es finito, es claro que ambos subdiferenciales son vacíos.

iii)
$$
\begin{aligned}
y_0 \in \partial f(x_0) &\Longleftrightarrow f^*(y_0) + f(x_0) = \langle x_0, y_0 \rangle \\
&\Longrightarrow f^*(y_0) + f^{**}(x_0) = \langle x_0, y_0 \rangle , \quad \text{por i)} \\
&\Longleftrightarrow x_0 \in \partial f^*(y_0). \qquad \#
\end{aligned}
$$

Observación 4.6.

Es importante notar que si $f \in \Gamma_0(X)$ se tiene $f = f^{**}$ (Corolario 3.8.) y en consecuencia, el punto iii) del teorema 4.5. resulta ser una equivalencia.

A continuación daremos un teorema que justifica la importancia del subdiferencias en la teoría de optimización:

Teorema 4.7.

Sea $f \in \overline{\mathbb{R}}^X$, un punto $\overline{x} \in X$ con $f(\overline{x})$ finito es solución del problema: (1) $\underset{x \in X}{\text{Min}}\ f(x)$ si y sólo si $0 \in \partial f(\overline{x})$.

Si además, $f \in \Gamma_0(X)$, el conjunto $\{\overline{x} \in X \ / \ \overline{x}$ es solución de (1)$\}$ es igual a $\partial f^*(0)$.

Demostración:

Un punto $\overline{x} \in X$ es solución de (1) si y sólo si $f(x) \geq f(\overline{x}) + \langle x - \overline{x}, 0\rangle$, $\forall\ x \in X$, es decir, si y sólo si $0 \in \partial f(\overline{x})$.

Si $f \in \Gamma_0(X)$, entonces se tiene la equivalencia:

$$0 \in \partial f(\overline{x}) \iff \overline{x} \in \partial f^*(0).$$

El teorema siguiente nos da un caso importante de subdiferencial para funciones convexas.

Teorema 4.8.

Si $f \in Conv(X)$ es finita y continua en x_0 entonces el subdiferencial de f en x_0 es un convexo no vacío en Y y es además compacto para la topología $\tilde{\sigma}(Y,X)$ (la topología menos fina sobre Y compatible con la dualidad).

Demostración:

Por observación 4.4. sólo debemos demostrar que $\partial f(x_0)$ es no vacío y $\sigma(Y,X)$-compacto. Probemos en primer lugar que $\partial f(x_0)$ es no vacío. Puesto que f es finita y continua en $x_0 \in X$, entonces epi(f) tiene interior no vacío (Corolario 2.12.). Del teorema de Hahn-Banach sabemos que se puede separar el convexo abierto int(epi(f)) y el punto $(x_0, f(x_0))$ en $X \times \mathbb{R}$, es decir, existe $(y,s) \in Y \times \mathbb{R}$ (no nulo) y $c \in \mathbb{R}$ tal que:

$$\begin{cases} \langle x,y\rangle + rs \leq c \quad , \quad \forall\ (x,r) \in epi(f) \\ \\ \langle x_0,y\rangle + f(x_0)s = c \end{cases}$$

es claro que $s \leq 0$, más aún $s < 0$. Multiplicando ambas relaciones por $\dfrac{1}{|s|}$ (podemos suponer $s = -1$, sin pérdida de generalidad) y en consecuencia tenemos:

$$\begin{cases} \langle x,y\rangle - c \leq f(x) \quad , \quad \forall\ x \in dom(f) \\ \\ \langle x,y\rangle - c \leq f(x_0) \quad , \end{cases}$$

lo que demuestra que $y \in \partial f(x_0)$.

En segundo lugar demostremos que $\partial f(x_0)$ es $\sigma(Y,X)$-compacto.

Sea $g_0(y) = f^*(y) - \langle x_0,y\rangle$, $y \in Y$. Como f es finita y continua en x_0 entonces (Corolario 3.15.) se tiene que g_0 es $\sigma(Y,X)$-inf-compacta. Por lo tanto, $\exists$ $y_0 \in Y$ tal que:

$$f^*(y_0) - \langle x_0,y_0\rangle = \min_{y \in Y} (f^*(y) - \langle x_0,y\rangle). \qquad (2)$$

Además por la hipótesis que satisface f, se tiene que f^* es propia y en consecuencia $f^*(y_0)$ es finito. Más aún $x_0 \in X$ e $y_0 \in Y$ satisfacen (2) si y sólo sí $x_0 \in \partial f^*(y_0)$, es decir, si y sólo sí $y_0 \in \partial f^{**}(X_0)$. Luego hemos probado que $\partial f^{**}(x_0)$ es el conjunto formado por los elementos $y_0 \in Y$ que satisfacen (2), el cual es un conjunto $\sigma(Y,X)$-compacto. #

A continuación veremos la relación que existe entre la noción de subgradiente y la derivada direccional. Una propiedad importante de las funciones convexas, es que en un punto x_0 donde $f(x_0) \in \mathbb{R}$, su derivada en x_0 , en la dirección d verifica:

$$f'(x_0;d) = \lim_{\lambda \to 0^+} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda} = \inf_{\lambda > 0} \frac{f(x_0 + \lambda d) - f(x_0)}{\lambda}$$

y la función $h(d) = f'(x_0;d)$ es convexa y positivamente homogénea.

La función $h(d) = f'(x_0;d)$ juega un rol importante en teoría de la optimización. El siguiente teorema relaciona la función h con el conjunto $\partial f(x_0)$.

Teorema 4.9.

Sea $f \in Conv(X)$, finita en x_0 , se tiene entonces:

$$h^* = \chi_{\partial f(x_0)}$$

$$h^{**}(d) = \sup_{y \in \partial f(x_0)} \langle d,y\rangle .$$

Si f es además continua en $x_0 \in X$, entonces h es finita y continua en todo X y se tiene que:

$$h(d) = \max_{y \in \partial f(x_0)} \langle d,y\rangle .$$

Demostración:

Por teorema 3.9. c), d), e) se tiene que la polar de la función:

$$x \in X \to \frac{1}{\lambda}\,[f(x_0 + \lambda x) - f(x_0)] \qquad \text{es la función:}$$

$$y \in Y \to \frac{1}{\lambda}\,[f(x_0) + f^*(y) - \langle x_0,y\rangle]$$

Luego, como $h(x) = \inf\limits_{\lambda > 0} \frac{1}{\lambda}\,[f(x_0 + \lambda x) - f(x_0)]$

entonces $h^*(y) = \sup\limits_{\lambda > 0} \dfrac{f(x_0) + f^*(y) - \langle x_0,y\rangle}{\lambda}$ (teorema 3.9. b))

Como siempre se tiene que: $f(x_0) + f^*(y) - \langle x_0,y\rangle \geq 0$ y se satisface la igualdad cuando $y \in \partial f(x_0)$ entonces tenemos:

$h^*(y) = \chi_{\partial f(x_0)}(y)$, $\forall\, y \in Y$. Si calculamos polar a ambos lados (ya se vio antes), tenemos que:

$$h^{**}(d) = \chi^{*}_{\partial f(x_0)}(d) = \sup_{y \in \partial f(x_0)} \langle d,y\rangle.$$

Si $f \in \mathrm{Conv}(X)$ es finita y continua en $x_0 \in X$, se puede demostrar que h además de ser convexa, es acotada superiormente sobre una vecindad del origen y por lo tanto es continua en el interior de su dominio efectivo y como es positivamente homogénea debe ser finita y por lo tanto continua en todo X. Esto demuestra que $h^{**} = h$.

Por otra parte el supremo se puede cambiar por máximo puesto que si f es convexa, finita y continua en $x_0 \in X$ el conjunto $\partial f(x_0)$ es $\sigma(Y,X)$-compacto. ∎

La relación:

$$h(d) = \max_{y \in \partial f(x_0)} \langle d,y\rangle \,, \quad \text{cuando } f \text{ es una función con-}$$

vexa, continua y finita en $x_0 \in X$, juega el mismo papel que la relación:

$$h(d) = \langle d,f'(x_0)\rangle \,, \quad \text{cuando } f \text{ es una función diferencia-}$$

ble en $x_0 \in X$.

En orden de importancia, las funciones convexas positivamente homogéneas finitas en el origen, vienen inmediatamente después de las funciones lineales. Toda función convexa, positivamente homogénea s.c.i. y que no toma el valor $-\infty$, puede escribirse como el supremo de una familia de funciones lineales continuas. Siguiendo esta idea algunos autores definen el subdiferencial de una función f en $x_0 \in X$, como el conjunto:

$$\partial f(x_0) = \{y \in Y \ / \ f'(x_0;d) \geq \langle d,y\rangle \ , \ \forall\, d \in X\}.$$

Un resultado importante en el llamado cálculo subdiferencial es el siguiente:

Teorema 4.10.

Sea $f_1, f_2, \ldots, f_p \in \Gamma_0(X)$. Si existe un punto $x_0 \in X$ donde todas estas funciones son finitas y todas, salvo eventualmente una, son continuas, se tiene:

$$\partial\left(\sum_{i=1}^{p} f_i\right)(x) = \sum_{i=1}^{p} \partial f_i(x).$$

Con este resultado podemos generalizar el teorema de Kuhn - Tucker, clásico en teoría de la optimización.

Sean $C_1, C_2, \ldots, C_{p-1}$ convexos cerrados con interior no vacío en X, C_p convexo cerrado en X. Denotemos:

$D = \bigcap_{i=1}^{p} C_1$. Sea $f \in \Gamma_0(X)$ continua en D. Supongamos además que: $\bigcap_{i=1}^{p-1} \overset{0}{C}_i \quad C_p \neq \phi$.

Bajo estas hipótesis tenemos el siguiente teorema:

Teorema 4.11.

Una condición necesaria y suficiente para que $\bar{x} \in D$ sea un mínimo de f sobre D es que existan elementos $y_0, y_1, \ldots, y_p \in Y$ tales que: $y_0 \in \partial f(\bar{x})$; $y_i \in \partial_{\chi_{C_i}}(\bar{x})$, $i = 1, 2, \ldots\ldots, p$;

$\sum_{i=1}^{p} y_i = 0$, es decir, $0 \in \partial f(\bar{x}) + \sum_{i=1}^{p} \partial \chi_{C_i} (\bar{x})$.

Demostración:

$$f(\bar{x}) = \underset{x \in D}{\text{Min}}\ f(x) \text{ si y sólo si } 0 \in \partial(f + \sum_{i=1}^{p} \chi_{C_i})(\bar{x}) .$$

Por otra parte, existe $x_0 \in \overset{o}{C}_i$, $\forall\, i = 1, \ldots, p-1$, $x_0 \in C_p$, y en este punto las funciones f y χ_{C_i} son claramente finitas, además son continuas salvo eventualmente χ_{C_p}. Del teorema 4.10, se tiene la igualdad: $\partial(f + \sum_{i=1}^{p} \chi_{C_i})(\bar{x}) = \partial f(\bar{x}) + \sum_{i=1}^{p} \partial \chi_{C_i}(\bar{x})$, lo que termina la demostración. #

Si los convexos C_i están definidos por funciones f_i en $\Gamma_0(X)$: $C_i = \{x \in X \,/\, f_i(x) \leq 0\}$, $i = 1, 2, \ldots, p$, y existe un punto $x_0 \in X$ tal que $f_i(x_0) < 0$, $\forall\, i = 1, 2, \ldots, p$, se puede verificar que los conjuntos $\partial \chi_{C_i}(\bar{x})$ son conos de la forma:

$$\partial \chi_{C_i}(\bar{x}) = \begin{cases} \bigcup_{\lambda \geq 0} \{\lambda \partial f_i(\bar{x})\} & \text{si } f_i(\bar{x}) = 0 \\ \{0\} & \text{si } f_i(\bar{x}) < 0 \\ \phi & \text{si } f_i(\bar{x}) > 0 \end{cases}$$

y en este caso una condición necesaria y suficiente para que $\bar{x}$ sea mínimo de f en $D = \sum_{i=1}^{p} C_i$, es que existan escalares $\lambda_i \geq 0$ para $i \in I(\bar{x}) = \{i \,/\, f_i(\bar{x}) = 0\}$ tales que:

$$0 \in \partial f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \partial f_i(\bar{x}).$$

Para terminar este panorama introductorio al análisis convexo, daremos un último resultado importante.

A menudo el las aplicaciones una función convexa f es dada como el supremo de una familia de funciones convexas:

$$f(x) = \operatorname{Sup}_{\alpha \in A} f_\alpha(x).$$

El teorema siguiente da una expresión simple para el subdiferencial de f en términos de los subdiferenciales de las f_α , $\alpha \in A$.

Teorema 4.12.

Sea $\{f_\alpha\}_{\alpha \in A}$ una familia de funciones convexas en $\overline{\mathbb{R}}^X$, donde A es un compacto cualquiera. Supongamos que existe un abierto U de X tal que la aplicación:

$$(\alpha, x) \in A \times U \to f_\alpha(x) \quad \text{sea finita y continua en} \quad A \times U.$$

Entonces, $\forall\, x_0 \in U$ se tiene:

i) La función $f = \operatorname{Sup}_{\alpha \in A} f_\alpha$ es continua.

ii) $\partial f(x_0) = \overline{co}\left(\bigcup_{\alpha \in F(x_0)} \partial f_\alpha(x_0)\right)$, donde

$F(x_0) = \{\alpha \in A \ / \ f_\alpha(x_0) = f(x_0)\}$ y donde $\overline{co}\,(E)$: denota la envoltura convexa cerrada de E.

5. PERTURBACION DE UN PROBLEMA DE MINIMIZACION Y PROBLEMA DUAL ASOCIADO.

Sean X e Y dos e.v.t.l.c. en dualidad por la forma bilineal $<\cdot,\cdot>$.

Consideremos el siguiente problema de minimización:

$$P)\quad \alpha = \operatorname*{Inf}_{x \in X} f(x) \quad , \quad \text{donde} \quad f \in \Gamma(X)$$

Denotaremos A al conjunto (eventualmente vacío) de soluciones del problema P)).

Si α es finito, se tiene:

$$A = \{\bar{x} \in X \,/\, \alpha = f(\bar{x})\} = \partial f^*(0).$$

El espacio X será llamado el espacio de las variables para el problema P).

Sean U y V dos e.v.t.l.c. en dualidad por la forma bilineal $(\cdot,\cdot)$. El espacio U será el espacio de las perturbaciones para el problema P).

Sea φ una función en $\Gamma_0(X \times U)$ que verifica:

$$\varphi(x,0) = f(x) \quad , \forall\, x \in X.$$

Para la perturbación $u \in U$, nosotros consideramos el problema perturbado:

$$P_u)\quad h(u) = \operatorname*{Inf}_{x \in X} \varphi(x,u).$$

Es claro que $\alpha = h(0)$.

Notemos que h es una función convexa en $\overline{\mathbb{R}}^U$, puesto que su epígrafo es justamente la proyección sobre $U \times \mathbb{R}$ del epígrafo de f, sin embargo, no está necesariamente en $\Gamma_0(U)$.

Tenemos la siguiente consecuencia:

$$-h^{**}(0) = -\sup_{v \in V} (\langle 0,v\rangle - h^*(v)) = \inf_{v \in V} h^*(v) \quad , \quad \text{y como}$$

$h^{**} = h^{\Gamma} \leqslant h$, se tiene la siguiente desigualdad:

$$-\alpha \leqslant -h(0) \leqslant -h^{**}(0) = \inf_{v \in V} h^*(v).$$

Los espacios $X \times U$ e $Y \times V$ son dos e.v.t.l.c. en dualidad por medio de la forma bilineal:

$$\langle (x,u), (y,v)\rangle := \langle x,y\rangle + (u,v).$$

Sea $\psi \in \Gamma_0(Y \times V)$ la polar (dual) de la función φ :

$$\psi(y,v) = \sup_{\substack{x \in X \\ u \in U}} (\langle x,y\rangle + (u,v) - \varphi(x,u)) \quad , \quad \text{y por lo tanto}$$

$$h^*(v) = \sup_{u \in U} \left((u,v) - \inf_{x \in X} \varphi(x,u) \right)$$

$$= \sup_{\substack{x \in X \\ u \in U}} \left(\langle x,0\rangle + (u,v) - \varphi(x,u) \right)$$

$$= \psi(0,v) \quad , \quad \forall\, v \in V.$$

Denotemos $g(v) = \psi(0,v)$, $\forall\, v \in V$. La desigualdad anterior:

$$-\inf_{x \in X} f(x) \leq \inf_{v \in V} h^*(v) \quad ,$$

nos conduce a definir el problema de minimización:

$$Q) \quad \beta = \inf_{v \in V} g(v) \quad , \quad \text{donde} \quad g \in \Gamma(V).$$

El problema Q) se llamará problema dual de P) con respecto a la familia de problemas perturbados P_u) (la dual está enteramente definida por la función φ).

Notaremos B al conjunto (eventualmente vacío) de solución del problema Q). Si β es finito, se tiene:

$$B = \{\overline{v} \in V \;/\; \beta = g(\overline{v})\} = \partial g^*(0).$$

Para este problema Q), el espacio de las variables es V y si lo consideramos como espacio de las perturbaciones, y ψ la función de perturbación, obtenemos el problema perturbado

$$Q_y) \quad k(y) = \operatorname{Inf}_{v \in V} \psi(y,v).$$

El problema dual (Q) corresponde al valor $y = 0$: $\beta = k(0)$. Se observa que la función $k \in \mathrm{Conv}(Y)$ pero en general no pertenece a $\Gamma_0(Y)$. Además siempre se tiene la desigualdad siguiente:

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

Siguiendo el mismo desarrollo que hicimos para el problema P), se demuestra que el dual de Q) con respecto a la perturbación ψ es precisamente el problema P).

En efecto, por simetría entre P) y Q) se tiene:

$$-\beta = -k(0) \leqslant -k^{**}(0) = \operatorname*{Inf}_{x \in X} k^*(\dot{x}).$$ Un cálculo directo nos muestra que:

$$k^*(x) = \varphi(x,0) = f(x) \quad , \quad \forall\, x \in X.$$ Por lo tanto:

$$-\beta \leqslant \operatorname*{Inf}_{x \in X} \varphi(x,0) = \operatorname*{Inf}_{x \in X} f(x) = \alpha.$$

A continuación desarrollaremos un ejemplo sencillo.

Consideremos el problema siguiente:

$$(P) \quad \alpha = \operatorname*{Inf}_{x \in C} f_0(x) \quad , \quad \text{donde} \quad f_0 \in \Gamma_0(X) \quad y$$

$$C = \{x \in X \ / \ f_i(x) \leqslant 0 \ , \ i = 1,\dots, m\} \quad \text{con}$$
$$f_i \in \Gamma_0(X) \quad , \quad \forall\, i = 1,\dots, m.$$

Luego C es un convexo cerrado en X.

Se recuerda que en el caso general nosotros debemos definir:

$$f(x) = \begin{cases} f_0(x) & \text{si} \quad x \in C \\ +\infty & \text{si} \quad x \notin C. \end{cases}$$

Notar que tendremos que perturbar el segundo miembro de las desigualdades que definen el convexo C.

El espacio de las perturbaciones será aquí $U = \mathbb{R}^m$. Para una perturbación $u \in \mathbb{R}^m$, se define:

$C_u = \{ x \in X \ / \ f_i(x) \leq u_i \ , \quad i = 1, 2, \ldots, m \}$ y consideremos el problema perturbado:

$$P_u) \quad h(u) = \operatorname*{Inf}_{x \in C_u} f_0(x).$$

Siguiendo la notación de antes, debemos definir:

$$\varphi(x,u) = \begin{cases} f_0(x) \ , & \text{si} \quad f_i(x) \leq u_i \ , \quad i = 1, \ldots, m \\ +\infty \ , & \text{en caso contrario.} \end{cases}$$

Calculemos la polar $\psi = \varphi^*$ de la funcional :

$$\psi(y,v) = \operatorname*{Sup}_{\substack{x \in X \\ u \in \mathbb{R}^m}} \left(\langle x,y \rangle + \sum_{i=1}^{m} u_i v_i - (x,u) \right)$$

$$= \operatorname*{Sup}_{\substack{x \in X \\ u \in \mathbb{R}^m \\ f_i(x) \leq u_i}} \left(\langle x,y \rangle + \sum_{i=1}^{m} u_i v_i - f_0(x) \right) ,$$

de donde finalmente se tiene:

$$\psi(y,v) = \begin{cases} \operatorname*{Sup}_{x \in X} \left(\langle x,y \rangle + \sum_{i=1}^{m} v_i f_i(x) - f_0(x) \right) \\ \qquad\qquad \text{Si} \quad v_i \leq 0 \ , \quad i = 1, \ldots, m \\ +\infty \ , \quad \text{en caso contrario.} \end{cases}$$

El problema dual Q) de P) (respecto de las perturbaciones que hemos introducido) se escribe entonces (para la perturbación $y = 0$) :

$$Q) \quad = \inf_{\substack{v \in \mathbb{R}^m \\ v_i \leq 0 \,,\; i = 1,\ldots,m}} \left(\sup_{x \in X} \left(\sum_{i=1}^{m} v_i f_i(x) - f_0(x) \right) \right)$$

El problema dual perturbado correspondiente a la perturbación $y \in Y$, se escribe:

$$Q_y) \quad k(y) = \inf_{\substack{v \in \mathbb{R}^m \\ v_i \leq 0 \,,\; i = 1,\ldots,m}} \left(\sup_{x \in X} \left(\langle x,y \rangle + \sum_{i=1}^{m} v_i f_i(x) - f_0(x) \right) \right).$$

Simplifiquemos este ejemplo suponiendo que: $X = Y = \mathbb{R}^n$ y que las f_i son funciones afines.

Sea $\langle x,y \rangle_n = \sum_{i=1}^{n} x\, y$, el producto escalar usual en $\mathbb{R}^n$, se supone además que se tiene:

$$f_0(x) = \langle c,x \rangle_n \,, \quad x \in \mathbb{R}^n$$

$$f_i(x) = \langle a_i,x \rangle_n - b_i \,, \quad x \in \mathbb{R}^n \,, \quad i = 1,\ldots, m \,,$$

donde $c, a_i, i = 1,\ldots, m$ son elementos fijos de $\mathbb{R}^n$ y $b_i \in \mathbb{R}$, $i = 1,\ldots, m$.

Escribamos explícitamente el problema P) y el problema perturbado $P_u)$ correspondiente a la perturbación $u \in \mathbb{R}^m$:

$$\text{P}') \quad \alpha = \inf_{x \in \mathbb{R}^n} \langle c,x \rangle_n$$

$$\langle a_i,x \rangle - b_i \leq 0$$

$$i = 1,\ldots, m$$

$$P'_u) \quad h(u) = \operatorname*{Inf}_{x \in \mathbb{R}^n} \; \langle c,x\rangle_n$$

$$\langle a_i,x\rangle_n - b_i \leqslant u_i$$

$$1 = 1 \;,\dots,\; m$$

Explicitemos ψ en este caso particular:

$$\psi(y,v) = \begin{cases} \displaystyle\operatorname*{Sup}_{x \in X} \left(\langle x,y\rangle_n + \sum_{i=1}^{m} v_i(\langle a_i,x\rangle_n - b_i) - \langle c,x\rangle_n\right) \\ \qquad\qquad \text{si} \quad v_i \leqslant 0 \;, \quad i = 1,\dots, m \\ +\infty \;, \quad \text{en el caso contrario} \;, \end{cases}$$

finalmente se tiene:

$$\psi(y,v) = \begin{cases} \displaystyle -\sum_{i=1}^{m} v_i b_i \,, \quad \text{si} \quad c - \sum_{i=1}^{m} v_i a_i = y \;, \quad \text{y} \quad v_i \leqslant 0 \;, \\ \qquad\qquad\qquad\qquad i = 1,\dots, m \\ +\infty \quad , \quad \text{en el caso contrario.} \end{cases}$$

Se obtiene entonces para el problema dual Q') y el problema dual perturbado correspondiente a la perturbación $y \in \mathbb{R}^n$:

$$Q') \quad \beta = \operatorname*{Inf}_{v \in \mathbb{R}^m} \left(-\langle b,v\rangle_m \right) \quad ,$$

$$v_i \leqslant 0 \;, \quad i = 1,\dots, m$$

$$c - \sum_{i=1}^{m} v_i a_i = 0$$

$$Q'_y) \quad k(y) = \operatorname*{Inf}_{v \in \mathbb{R}^m} \; (-\langle b,v\rangle_m) \; .$$

$$v_i \leq 0 \; , \quad i = 1,\ldots, m$$

$$c - \sum_{i=1}^{m} v_i a_i = y$$

COMENTARIOS:

En estas notas hemos visto que en el caso convexo no diferenciable, el subdiferencial de f en x_0 está ligado a la derivada direccional mediante la relación: $f'(x_0;d) = \max_{y \in \partial f(x_0)} \langle d,y \rangle$. Esta relación es muy importante desde el punto de vista del desarrollo de algoritmos.

En el año 1973, Clarke F.H. [1] origina un salto importante en la teoría de la optimización no diferenciable no (caso no convexo) pues, logra definir una derivada direccional generalizada denotada por $f^0(x_0;d)$ (en el caso convexo coincide con $f'(x_0;d)$) y que resultó también ser la función soporte de este nuevo subdiferencial llamado gradiente generalizado. Esta generalización al caso localmente lipschitziano puede verse también en [2]. Posteriormente Rockafellar R.T. [9], lo generaliza no necesariamente al caso localmente lipschitziano.

Un libro que contiene un estudio sistemático de los últimos resultados en la diferenciación de funciones convexas es [4].

Un buen texto para iniciarse en el estudio de la optimización no diferenciable es [3].

Un buen artículo para ver una síntesis del desarrollo que ha tenido la optimización no diferenciable entre los años 1974 - 1984 y ver su estado actual es [5].

Al finalizar estos comentarios no podemos dejar de manifestar nuestro reconocimiento al Dr. Rafael Correa F. quién dirige desde hace tres años en la Universidad Federico Santa María un grupo de investigación en optimización en el cual hemos tenido la suerte de participar. A través de todo este tiempo hemos podido constatar el gran apoyo que hemos tenido de este profesor para iniciarnos en esta gran línea de investigación de las Matemáticas Aplicadas.

REFERENCIAS:

[1] CLARKE, F.H. "Necessary conditions for nonsmooth problems in optimal Control and teh calculus of variations", Ph. D. Thesis , Department of Mathematics, University of Washington, Seattle, 1973.

[2] CLARKE, F.H. "Generalized Gradients and Applications", American Math. Soc., Vol. 205, 1975.

[3] CLARKE, F.H. "Optimization and Nonsmooth Analysis", J. Wiley, 1973.

[4] GILES, J.R. "Convex analysis with application in diferentiation of convex functions", Pitman Adv. Publ. Program., 1982.

[5] HIRIART-URRUTY, J.B. "Miscellanies on Nonsmooth Analysis and Optimization", Charla en Sopron (Hungría), 1984.

[6] MOREAU, J.J. "La convexité en statique", Lectures Notes in Economics and Mathematical Systems, 102, 1974.

[7] MOREAU, J.J. "Fonctionnelles Convexes", Séminaire sur les Equàtions aux Dérivées Partielles, College de France, París, 1966-67.

[8] ROCKAFELLAR, R.T. "Convex Analysis", Princeton Univ. Press, 1972.

[9] ROCKAFELLAR, R.T. "The Theory of Subgradients and its Applications to Problems of Optimization", Haldermann Verlag, 1981.