UNIVERSIDAD MAYOR DE SAN ANDRÉS

FACULTAD DE CIENCIAS PURAS Y NATURALES

CARRERA DE MATEMÁTICA

PROYECTO DE GRADO

DUALIDAD FUERTE EN OPTIMIZACIÓN NO CONVEXA

PARA OPTAR AL TITULO DE LICENCIATURA EN MATEMÁTICA

POSTULANTE: EDUARDO VINO MACHICADO

TUTOR: M.Sc. WILLY CONDORI EQUICE

LA PAZ-BOLIVIA

2018

Índice general

Agradecimientos

Agradezco a Dios, a mis Padres, a mi tutor M.Sc. Willy Condori Equice que con su guía y paciencia este trabajo salió adelante, al proyecto Dinámicas de Control cuyos profesores encargados el Dr. Efrain Cruz M., M.Sc. Willy Condori E. y M.Sc. Miguel Yucra C. me dieron la oportunidad para desarrollar y realizar las exposiciones de este trabajo en los seminarios que organizan los días miércoles en el campus de Cota Cota. Agradeciendo también la oportunidad que me dio el proyecto Dinámicas de Control de exponer parte de mi trabajo en el COMCA Arica-2017.

Introducción

El presente proyecto de grado estudia la problemática de hallar condiciones que nos permitan caracterizar, en algún sentido, la propiedad de Dualidad Fuerte en optimización no convexa.

Es así que, en el primer capítulo, estudiaremos algunos resultados sobre conjuntos convexos, así como algunas de sus propiedades topológicas que necesitaremos. También realizaremos un pequeño estudio sobre hiperplanos, separación de un convexo y un punto, soporte de conjuntos convexos. Luego estudiaremos algunos conceptos sobre funciones convexas y enunciaremos las condiciones $K.K.T.$ las cuales no la desarrollaremos pues no es nuestro objetivo y solo serán necesarios para el desarrollo de algunos resultados del capítulo 2.

El concepto de dualidad surge con regular frecuencia en matemáticas. Planteado un problema de optimización nos preguntamos si existe otro problema asociado al anterior que permita, entre otras cosas, resolver el primero en forma mas sencilla aprovechando las propiedades que el segundo tiene, como por ejemplo, pasar de un problema con restricciones a uno sin restricciones, éstos son conocidos como los problemas primal y dual.

Entonces, en el segundo capítulo, estudiaremos el concepto de problema primal para luego demostrar una condición de Dualidad Fuerte que afirma que el dual de un problema tiene solución y que éste coincide con el valor optimal del problema primal.

El objetivo principal será presentado en el capítulo 3. El cual es estudiar un resultado de Dualidad Fuerte pero en optimización no convexa y que lo presentamos así:

Sea $C \subset \mathbb{R}^n$ y $P \subseteq \mathbb{R}^m$ cono, convexo, cerrado. Dado $f : C \to \mathbb{R}$ y $g : C \to \mathbb{R}^m$, se considera el problema de minimizacion primal siguiente:

$$\mu = \inf_{g(x) \in -P} f(x)$$
$$x \in C$$

El problema dual Lagrangiano asociado al primal se define como:

$$\upsilon = \sup_{\lambda^* \in P^*} \inf_{x \in C} f(x) + \langle \lambda^*, g(x) \rangle$$

Donde P^* es el cono polar positivo de P. Se dice que el problema primal tiene la propiedad $Z.D.G.$ (Zero duality Gap) si los valores óptimos del primal y dual coinciden, esto es, $\mu = \upsilon$. El problema primal se dice que tiene la propiedad de Dualidad Fuerte si el primal tiene la propiedad $Z.D.G.$ y además el problema dual lagrangiano admite solución.

Capítulo 1

Aspectos de Análisis Convexo

En éste capítulo revisamos los resultados que necesitaremos del análisis convexo para el desarrollo de los capítulos siguientes en el estudio de dualidad. Lo que se hará es desarrollar algunas proposiciones, principalmente sobre los conos convexos, mencionaremos definiciones y resultados de funciones convexas para luego enunciar el teorema de $K.K.T.$ el cual no lo desarrollaremos.

1.1. Conjuntos convexos

Definición 1. Sean $A, P \subseteq \mathbb{R}^n$.

1. A es un conjunto convexo si y solo si para todo $x, y \in A$ y todo $\lambda \in [0, 1]$ se cumple que $\lambda x + (1 - \lambda)y \in A$.

2. Diremos que x de $\mathbb{R}^n$ es combinación lineal convexa de $x_1, x_2, \ldots, x_k \in \mathbb{R}^n$ si existen $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ no negativos con $\sum_{i=1}^k \lambda_i = 1$ y $x = \sum_{i=1}^k \lambda_i x_i$.

3. Sea A no vacío. La *envolvente convexa* de A, $\mathrm{co}(A)$, es el conjunto de puntos de $\mathbb{R}^n$ que pueden escribirse como combinación lineal convexa de una cantidad de puntos de A, es decir;

$$\mathrm{co}(A) = \left\{ \sum_{i=1}^m \lambda_i x_i : \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0, \text{ para todo } i = 1, \ldots, m; x_i \in A; m \in \mathbb{N} \right\}.$$

4. P es un cono si y solo si tP está en P; para todo $t \geq 0$, donde $tP = \{tp; p \text{ en } P\}$.

Es fácil ver que si $A_1 \subset A_2$ entonces $\mathrm{co}(A_1) \subset \mathrm{co}(A_2)$, pues toda combinación convexa de elementos de A_1 es una combinación convexa de A_2. También, si A es convexo entonces $\mathrm{co}(A) = A$. Denotamos por $\overline{\mathrm{co}}(A)$ a la clausura topológica de $\mathrm{co}(A)$.

Proposición 1. Sea $A \subset \mathbb{R}^n$. $\text{co}(A)$ es un conjunto convexo.

Demostración. Sean x, y en $\text{co}(A)$ y λ en $[0,1]$. Luego,

$$x = \sum_{i=1}^{k} \mu_i x_i, \quad y = \sum_{i=1}^{m} \nu_i y_i$$

donde $x_1, \ldots, x_k$, $y_1, \ldots, y_m$ están en A y $\mu_1, \ldots, \mu_2 \geq 0$, $\nu_1, \ldots, \nu_m \geq 0$ tal que $\sum_{i=1}^{k} \mu_i = 1$, $\sum_{i=1}^{m} \nu_i = 1$. Entonces

$$\lambda x + (1-\lambda)y) = \lambda \sum_{i=1}^{k} \mu_i x_i + (1-\lambda) \sum_{i=1}^{m} \nu_i y_i = \sum_{i=1}^{k} \lambda \mu_i x_i + \sum_{i=1}^{m} (1-\lambda) \nu_i y_i.$$

Llamando $z_i = x_i$, $\alpha_i = \lambda \mu_i$, para todo $i = 1, \ldots, k$ y $z_{k+i} = y_i$, $\alpha_{k+i} = (1-\lambda)\nu_i$ para todo $i = 1, \ldots, m$ se tiene que

$$\lambda x + (1-\lambda)y = \sum_{i=1}^{k+m} \alpha_i z_i$$

con $z_i \in A$, $\alpha_i \in [0,1]$, para todo $i = 1, \ldots, k+m$ y

$$\sum_{i=1}^{k+m} \alpha_i = \lambda \sum_{i=1}^{k} \mu_i + (1-\lambda) \sum_{i=1}^{m} \nu_i = \lambda + (1-\lambda) = 1.$$

Luego, por definición, se tiene que $\lambda x + (1-\lambda)y \in \text{co}(A)$. Así $\text{co}(A)$ es un conjunto convexo. □

Proposición 2. Un conjunto $C \subseteq \mathbb{R}^n$ es un cono convexo si y solo si para todo x_1, x_2 de C y para todo $\alpha, \beta \geq 0$ se verifica que $\alpha x_1 + \beta x_2$ pertenece a C.

Demostración. Asumiendo que C es un cono convexo. Sean $x_1, x_2 \in C$ y $\alpha, \beta \geq 0$ entonces

$$\alpha x_1 + \beta x_2 = (\alpha + \beta) \left[\frac{\alpha}{\alpha+\beta} x_1 + \frac{\beta}{\alpha+\beta} x_2 \right]$$

Luego, como $\frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} = 1$ y C es convexo, entonces $\frac{\alpha}{\alpha+\beta} x_1 + \frac{\beta}{\alpha+\beta} x_2 \in C$. Por tanto, como C es un cono, $\alpha x_1 + \beta x_2 \in C$.

Recíprocamente. Sea $x \in C$ y $t \geq 0$. Con $\alpha = t$, $x_1 = x$, $\beta = 0$ en la hipótesis $\alpha x_1 + \beta x_2 = t x_1 \in C$. Por tanto C es un cono. Por otro lado, para $x_1, x_2 \in C$ y $\lambda \in [0,1]$ con $\alpha = \lambda$, $\beta = (1-\lambda)$ tendremos que $\alpha x_1 + \beta x_2 = \lambda x_1 + (1-\lambda) x_2 \in C$. Así C es convexo. □

Lema 1. Sea $\{C_i\}_{i \in I}$, una familia arbitraria de conjuntos no vacíos en $\mathbb{R}^n$. Si cada C_i es cerrado entonces su intersección es cerrada.

Demostración. Sea $a \in \bigcap_{i\in I} C_i$. Entonces para todo $i \in I$ se tendrá $a \in C_i$. Luego como $a \in C_i$ y cada C_i es cerrado, existen sucesiones $(x_{n,i})$, para todo $i \in I$, tal que $x_{n,i} \to a$. Luego $(x_{n,i}) \in \bigcap_{i\in I} C_i$. Por tanto $\bigcap_{i\in I} C_i$ es cerrado. $\square$

Proposición 3. La intersección arbitraria de conjuntos convexos es un conjunto convexo.

Demostración. Sea $\{C_i\}_{i\in I}$ una colección de conjuntos convexos. Si $x, y \in \bigcap_{i\in I} C_i$ entonces, para todo $i \in I$ se tiene que $x, y \in C_i$, y como cada C_i es convexo, el segmento $[x, y] \subset C_i$ para todo $i \in I$. Entonces $[x, y] \subset \bigcap_{i\in I} C_i$, por lo tanto $\bigcap_{i\in I} C_i$ es convexo. $\square$

Proposición 4. Sean $A, K \subset \mathbb{R}^n$.

1. $\overline{\text{co}}(A) = \overline{\text{co}}(\overline{A})$.
2. $\text{co}(A + K) = \text{co}(A) + K$ cuando K es convexo.

Demostración. Veamos que $\overline{\text{co}}(A) = \overline{\text{co}}(\overline{A})$. Como $A \subseteq \overline{A}$ entonces $\text{co}(A) \subset \text{co}(\overline{A})$, luego $\overline{\text{co}}(A) \subseteq \overline{\text{co}}(\overline{A})$. Recíprocamente, sea $a \in \text{co}(\overline{A})$ entonces existen $x_1, x_2, \ldots, x_m \in \overline{A}$ y $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ tal que $\sum_{i=1}^m \lambda_i = 1$ y $\sum_{i=1}^m \lambda_i x_i = a$. Como $x_1, x_2, \ldots, x_m \in \overline{A}$ existen sucesiones $(y_{i,k}) \in A$ tal que $y_{i,k} \to x_i$ para cada $k \in \mathbb{N}$. Entonces

$$a = \sum_{i=1}^{m} \lambda_i x_i = \sum_{i=1}^{m} \lambda_i \lim_{k\to\infty} y_{i,k} = \lim_{k\to\infty} \sum_{i=1}^{m} \lambda_i y_{i,k},$$

donde $\sum_{i=1}^m \lambda_i y_{i,k} \in \text{co}(A)$. Luego $a \in \overline{\text{co}}(A)$. Por tanto

$$\text{co}(\overline{A}) \subset \overline{\text{co}}(A), \tag{1.1}$$

de donde $\overline{\text{co}}(\overline{A}) \subset \overline{\overline{\text{co}}}(A) = \overline{\text{co}}(A)$.

Ahora demostremos que $\text{co}(A + K) = \text{co}(A) + K$. Tomemos $x \in \text{co}(A + K)$, luego existen $a_1, a_2, \ldots, a_k \in A$, $b_1, b_2, \ldots, b_k \in K$, $\lambda_1, \lambda_2, \ldots, \lambda_k \geq 0$ tal que $\sum_{i=1}^k \lambda_i = 1$ y $x = \sum_{i=1}^k \lambda_i(a_i + b_i)$. Así,

$$x = \sum_{i=1}^{k} \lambda_i(a_i + b_i) = \sum_{i=1}^{k} \lambda_i(a_i) + \sum_{i=1}^{k} \lambda_i(b_i) \in \text{co}(A) + \text{co}(K)$$

es decir $x \in \text{co}(A) + \text{co}(K) = \text{co}(A) + K$, pues K es convexo.

Recíprocamente, para todo $x \in \text{co}(A)$ y $y \in K$ existen $a_1, a_2, \ldots, a_k \in A$, $\lambda_1, \lambda_2, \ldots, \lambda_k \geq 0$ tal que $\sum_{i=1}^k \lambda_i = 1$, $x = \sum_{i=1}^k a_i \lambda_i$. Así

$$\begin{aligned} x + y &= \sum_{i=1}^{k} a_i \lambda_i + y \\ &= \sum_{i=1}^{k} a_i \lambda_i + \sum_{i=1}^{k} \lambda_i y \\ &= \sum_{i=1}^{k} \lambda_i (a_i + y) \in \text{co}(A + K). \end{aligned}$$

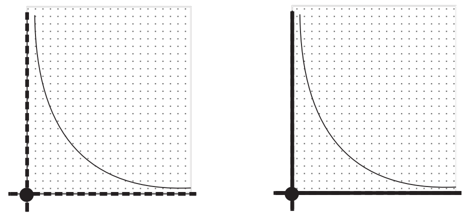


Figura 1.1: $A \cup \{(0,0)\}$

Por lo tanto $x + y \in \mathrm{co}(A + K)$. $\square$

Ejemplo 1. En (1.1) no se da la igualdad. En efecto, sea $A = \left\{a \in \mathbb{R}^2 / a = (x, \frac{1}{x}), x > 0\right\} \cup \{(0,0)\}$. Luego $\mathrm{co}(\overline{A})$ es el primer cuadrante unido al origen sin los ejes, y $\overline{\mathrm{co}}(A)$ es el primer cuadrante unido al origen y a los ejes, ver 1.1. Por tanto $\overline{\mathrm{co}}(A)$ no está contenido en $\mathrm{co}(\overline{A})$.

1.2. Algunas propiedades topológicas de conjuntos convexos

Proposición 5. Sea $A \subset \mathbb{R}^n$ un conjunto convexo, entonces su interior también lo es.

Demostración. Sean $x, y \in \mathrm{int}(A)$ y $\lambda \in [0,1]$. Luego, existen $\epsilon_1 > 0$, $\epsilon_2 > 0$ tal que $B(x, \epsilon_1) \subset A$, $B(y, \epsilon_2) \subset A$. Tomemos $\epsilon = \text{mín}\{\epsilon_1, \epsilon_2\}$. Luego $B(x, \epsilon) \subset A$, $B(y, \epsilon) \subset A$. Sea $z = \lambda x + (1 - \lambda)y$, mostremos que $B(z, \epsilon) \subset A$. Si $w \in B(z, \epsilon)$, entonces existen $\overline{x} = (w - z) + x \in B(x, \epsilon)$, $\overline{y} = (w - z) + y \in B(y, \epsilon)$. Luego

$$\begin{aligned}
\lambda\overline{x} + (1-\lambda)\overline{y} &= \lambda(w - z) + \lambda x + (1 - \lambda)(w - z) + (1 - \lambda)y \\
&= z + w - z = w.
\end{aligned}$$

Por tanto $w \in A$ y $B(z, \epsilon) \subset A$. Además como $\|\overline{x} - x\| < \epsilon$, $\|\overline{y} - y\| < \epsilon$, se tendrá que

$$\begin{aligned}
\|w - z\| &= \|\lambda\overline{x} + (1 - \lambda)\overline{y} - \lambda x - (1 - \lambda)y\| \\
&= \|\lambda(\overline{x} - x) + (1 - \lambda)(\overline{y} - y)\| \\
&\leq \lambda \|\overline{x} - x\| + (1 - \lambda) \|\overline{y} - y\| \\
&< \lambda\epsilon + (1 - \lambda)\epsilon = \epsilon.
\end{aligned}$$

Así el int(A) es convexo. $\square$

Proposición 6. Sea $A \subset \mathbb{R}^n$ un conjunto convexo, entonces su cerradura también es un conjunto convexo.

Demostración. Sean $x, y \in \overline{A}$ y $t \in$[0,1] entonces existen sucesiones $(r_k), (s_k) \in A$ tal que $r_k \to x$ y $s_k \to y$. Luego $tr_k + (1-t)s_k \to tx + (1-t)y$, por lo tanto $tx + (1-t)y \in \overline{A}$ y finalizamos diciendo que $\overline{A}$ es convexo. $\square$

Lema 2. Sea $A \subset \mathbb{R}^n$ convexo y supongamos que $x \in \text{int}(A)$ e $y \in \overline{A}$. Entonces $[x, y) \subset \text{int}(A)$.

Demostración. Sea $\epsilon > 0$ tal que $B(x, \epsilon) \subset A$ y distingamos dos casos. Supongamos primeramente que $y \in A$. Entonces $[x, y] \subset A$ y, dado $z = (1-\lambda)x + \lambda y \in [x, y)$ para $\lambda \in [0, 1)$, mostremos que $B(z, (1-\lambda)\epsilon) \subset A$. Sea $w \in B(z, (1-\lambda)\epsilon)$, entonces existe $\overline{w}$ tal que $\overline{w} = \frac{w - \lambda y}{1-\lambda}$. Luego

$$\begin{aligned}
\|\overline{w} - x\| &= \left\| \frac{w - \lambda y}{1-\lambda} - x \right\| \\
&= \frac{1}{1-\lambda} \|w - \lambda y - (1-\lambda)x\| \\
&= \frac{1}{1-\lambda} \|w - (\lambda y + (1-\lambda)x)\| \\
&= \frac{1}{1-\lambda} \|w - z\| \\
&< \frac{1}{1-\lambda}(1-\lambda)\epsilon = \epsilon.
\end{aligned}$$

así $z \in \text{int}(A)$ y $[x, y) \subset \text{int}(A)$.

Ahora supongamos que $y \notin A$ y mostremos que $[x, y) \subset A$. Es claro que $x \in A$, luego tomemos $z = (1-\lambda)x + \lambda y$ para algún $\lambda \in (0, 1)$. Como $y \in \overline{A}$ entonces para toda $B(y, \frac{(1-\lambda)\epsilon}{\lambda})$ y todo $\frac{(1-\lambda)\epsilon}{\lambda} > 0$, se tiene que $B(y, \frac{(1-\lambda)\epsilon}{\lambda}) \cap A \neq \emptyset$. Existe $u \in B(y, \frac{(1-\lambda)\epsilon}{\lambda}) \cap A$ luego $u = y + \frac{1-\lambda}{\lambda} v$ para un cierto $v \in B(0, \epsilon)$. Entonces

$$\begin{aligned}
z &= (1-\lambda)x + \lambda y \\
&= (1-\lambda)x + \lambda u - (1-\lambda)v \\
&= \lambda u + (1-\lambda)(x - v),
\end{aligned}$$

y observemos que $u \in A$ y que $(x - v) \in B(x, \epsilon) \subset A$ pues $v \in B(0, \epsilon)$. Luego $z = \lambda u + (1-\lambda)(x - v) \in A$ y por tanto $[x, y) \subset A$ como queríamos.

Finalmente, dado $z \in [x, y)$ el párrafo anterior permite aplicar el primer caso a $[x, z)$ luego $[x, z) \subset \text{int}(A)$. Como ésto es cierto para todo $z \in [x, y)$, concluimos que $[x, y) \subset \text{int}(A)$. $\square$

Proposición 7. Dado $A \subset \mathbb{R}^n$ convexo y si $\text{int}(A) \neq \emptyset$, entonces se cumple que $\overline{\text{int}(A)} = \overline{A}$

Demostración. Como $\text{int}(A) \subset A$ entonces $\overline{\text{int}(A)} \subset \overline{A}$. Recíprocamente, sea $a \in \overline{A}$ y tomemos $x_0 \in \text{int}(A) \neq \emptyset$ por el anterior lema $[x_0, a) \subset \text{int}(A)$. Por tanto, a es límite de puntos del interior de A, luego $a \in \overline{\text{int}(A)}$. □

Proposición 8. Dado $A \subset \mathbb{R}^n$ convexo con interior no vacío, se cumple que $\text{int}(\overline{A}) = \text{int}(A)$.

Demostración. Como $A \subset \overline{A}$ entonces $\text{int}(A) \subset \text{int}(\overline{A})$. Veamos la otra inclusión, sea A convexo con interior no vacío y sea $x \in \text{int}(\overline{A})$. Tomemos $a \in \text{int}(A)$. Si $a = x$ no hay nada que probar. Supongamos que $a \neq x$, como $x \in \text{int}(\overline{A})$ existe $\epsilon > 0$ tal que $B(x, \epsilon) \subset \overline{A}$. Consideremos el punto $w = x + \frac{\epsilon}{2}\frac{x-a}{|x-a|}$ pertenece a $B(x, \epsilon) \subset \overline{A}$ y $x \in [a, w)$. Por el anterior lema tenemos que $[a, w) \subset \text{int}(A)$, pues A es convexo. Despejando x, de la definición de w, tendremos que $x = \lambda a + (1 - \lambda)w$ donde $\lambda = \frac{\epsilon}{\epsilon + 2|x-a|}$ en $(0, 1)$ luego $x \in (a, w) \subset [a, w) \subset \text{int}(A)$. □

1.3. Hiperplanos, separación y soporte de conjuntos convexos

Definición 2. Un hiperplano H, es un conjunto de puntos de la forma $H = \left\{x/p^T x = \alpha\right\}$, donde p está en $\mathbb{R}^n$, $p \neq 0$ y α de $\mathbb{R}$. El vector p es el vector normal al hiperplano.

Definición 3. Sean $S_1, S_2 \neq \emptyset$ dos conjuntos en $\mathbb{R}^n$. Se dice que un hiperplano $H = \left\{x/p^T x = \alpha\right\}$ separa a S_1 de S_2 si:

i) $S_1 \subset H^+$ es decir $p^T x \geq \alpha$ para todo $x \in S_1$, y $S \subset H^-$ es decir $p^T x \leq \alpha$ para todo x en S_2.

ii) $S_1 \subset H^-$ y $S_2 \subset H^+$.

Definición 4. Se dice que el hiperplano H separa fuertemente a S_1 de S_2 si existe $\epsilon > 0$, tal que:

i) $p^T x \geq \alpha + \epsilon$ para todo $x \in S_1$; $p^T x \leq \alpha$ para todo $x \in S_2$ ó

ii) $p^T x \leq \alpha - \epsilon$ para todo $x \in S_1$; $p^T x \geq \alpha$ para todo $x \in S_2$.

Teorema 1. Sea $S \subset \mathbb{R}^n$ convexo, cerrado y no vacío. Para cada x de $\mathbb{R}^n$ existe un único $\overline{x}$ en S tal que $|x - \overline{x}| \leq |x - y|$ para todo y en S.

Demostración. Veamos la existencia usando el hecho de que S es cerrado. Dado $x \in \mathbb{R}^n$, tomemos $R > 0$ tal que $K = \overline{B}(x, R) \cap S$. El conjunto K es cerrado y acotado, luego K es compacto. Como la aplicación $y \to |x - y|$ es continua, deducimos que existe $\overline{x} \in K$ tal que $|x - \overline{x}| \leq |x - y|$ para todo $y \in K$. Ahora, para $y \notin K$ se sigue que $|x - y| > R \geq |x - \overline{x}|$, luego $|x - \overline{x}| \leq |x - y|$ para todo $y \in S$.

Probemos la unicidad usando el hecho de que S es convexo. Razonando por reducción al absurdo, supongamos que existen $\overline{x}, \overline{y} \in S$ distintos tal que $|x - \overline{x}| \leq |x - y|$ y $|x - \overline{y}| \leq |x - y|$ para todo $y \in S$, entonces podemos considerar $c = \frac{1}{2}(\overline{x} + \overline{y})$, que pertenece a S, pues S es convexo. Luego

$$\begin{aligned}
\langle \overline{x} - c, x - c \rangle &= \left\langle \overline{x} - \frac{1}{2}\overline{x} - \frac{1}{2}\overline{y}, x - \frac{1}{2}\overline{x} - \frac{1}{2}\overline{}y \right\rangle \\
&= \left\langle \frac{1}{2}(\overline{x} - \overline{y}), \frac{1}{2}(2x - \overline{x} - \overline{y}) \right\rangle \\
&= \frac{1}{4} \langle \overline{x} - \overline{y}, 2x - \overline{x} - \overline{y} \rangle \\
&= \frac{1}{4} \langle (x - \overline{y}) - (x - \overline{x}), (x - \overline{x}) + (x - \overline{y}) \rangle \\
&= \frac{1}{4} \left[|x - \overline{y}|^2 - |x - \overline{x}|^2 \right] = 0
\end{aligned}$$

y los puntos $\overline{x}, x$ y c forman un triángulo rectángulo y el teorema de Pitágoras nos dice que $|x - \overline{x}|^2 = |x - c|^2 + |c - \overline{x}|^2 > |x - c|^2$. Esto prueba que $|x - \overline{x}| > |x - c|$, contradiciendo al hecho de que $\overline{x}$ esta a mínima distancia de x. □

Teorema 2. Sea $S \subset \mathbb{R}^n$ un conjunto convexo y cerrado y sea $y \notin S$. Entonces, existen p en $\mathbb{R}^n$ y α en $\mathbb{R}$ tal que $p^T y > \alpha$ y $p^T x \leq \alpha$, para todo x en S.

Demostración. Por anterior teorema existe un $\overline{x} \in S$ que minimiza la distancia entre y y cualquier punto $x \in S$, es decir; $\|y - \overline{x}\| \leq \|y - x\|$ para todo x en S. Sea $\overline{\overline{x}}$ en S tal que $\overline{\overline{x}} = \lambda x + (1 - \lambda)\overline{x}$, para todo λ en $[0, 1]$. Luego

$$\begin{aligned}
\|y - \overline{x}\|^2 &\leq \left\|y - \overline{\overline{x}}\right\|^2 \\
&= \|y - \overline{x} - \lambda(x - \overline{x})\|^2 \\
&= \|y - \overline{x}\|^2 + \lambda^2 \|x - \overline{x}\|^2 - 2\lambda(x - \overline{x})^T(y - \overline{x}) \\
&= \|y - \overline{x}\|^2 + \lambda^2 \|x - \overline{x}\|^2 + 2\lambda(x - \overline{x})^T(\overline{x} - y).
\end{aligned}$$

Entonces

$$\lambda^2 \|x - \overline{x}\|^2 + 2\lambda(x - \overline{x})^T(\overline{x} - y) \geq 0.$$

A la última desigualdad la dividimos entre λ y tomemos el límite cuando $\lambda \to 0$ del cual se tiene que $(x - \overline{x})^T(y - \overline{x}) \leq 0$ para todo $x \in S$.

De lo anterior decimos que:

$$-x^T(y-\overline{x}) \geq -\overline{x}^T(y-\overline{x}), \tag{1.1}$$

para todo $x \in S$.

Ahora, examinemos lo siguiente:

$$\|y-\overline{x}\|^2 = (y-\overline{x})^T(y-\overline{x}) = y^T(y-\overline{x}) - \overline{x}^T(y-\overline{x}) > 0. \tag{1.2}$$

Sean $p = y - \overline{x} \neq 0$ para y que no está en S, $\overline{x}$ en S y $\alpha = p^T\overline{x}$. De lo obtenido en (1,1) concluimos que $\alpha = p^T\overline{x} \geq p^Tx$ para todo x en S. Y de lo examinado en (1,2) decimos que $p^Ty > p^T\overline{x} = \alpha$. Por lo tanto $p^Tx \leq \alpha < p^Ty$ para todo x en S. □

Definición 5. Sea $S \neq \emptyset$ un conjunto en $\mathbb{R}^n$, y $\overline{x}$ un punto en la frontera de S. Se dice que el hiperplano $H = \left\{x/p^T(x-\overline{x}) = 0\right\}$ soporta a S en $\overline{x}$ si $S \subset H^+$ ó $S \subset H^-$.

Si además, S no es subconjunto de H, se dice que el hiperplano soporta bien a S en $\overline{x}$.

Teorema 3. Sea $S \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$, y $\overline{x}$ un punto sobre la frontera de S. Entonces, existe $p \neq 0$ tal que $H = \left\{x/p^T(x-\overline{x}) = 0\right\}$ soporta a S en $\overline{x}$.

Demostración. Debemos mostrar que existe un p tal que $p^T(x-\overline{x}) \leq 0$ para todo $x \in \overline{S}$. Como $\overline{x}$ está en la frontera de S, entonces $\overline{x}$ no está en el $\mathrm{int}(S)$. Luego por la convexidad de S implica que $\overline{x}$ no está en el $\mathrm{int}(\overline{S})$. De ésta manera decimos que

$$\overline{x} \in \left[\mathrm{int}(\overline{S})\right]^c = \partial\left(\overline{S}\right) \cup \mathrm{Ext}\left(\overline{S}\right),$$

en particular $\overline{x} \in \partial\overline{S}$.

De lo anterior, como $\overline{x} \in \partial\left(\overline{S}\right)$ implicará que

$$\overline{x} \in \overline{\overline{S}} \cap \overline{\mathbb{R}^n \setminus \overline{S}} = \overline{S} \cap \overline{\mathbb{R}^n \setminus \overline{S}}.$$

Entonces, existe una sucesión $\{y_k\} \subset \mathbb{R}^n \setminus \overline{S}$ tal que $y_k \to \overline{x}$. Por el anterior teorema, para toda y_k en la sucesión existe $p_k \neq 0$ tal que

$$(p_k)^T y_k > (p_k)^T x,$$

para todo $k \in \mathbb{N}$ y para toda $x \in \overline{S}$.

Consideremos $\overline{p}_k = \frac{p_k}{\|p_k\|}$. Luego, $\|\overline{p}_k\| = 1$ para toda $k \in \mathbb{N}$. Por lo tanto $\overline{p}_k \in S^{n-1}$, donde S^{n-1} es la esfera unitaria, el cual es compacto. Y por la compacidad de S^{n-1} toda sucesión $\{\overline{p}_k\} \subset S^{n-1}$ posee una subsucesión $\{\overline{p}_k\}_{k\in\mathbb{N}_1}$ tal que $\overline{p}_k \to \overline{p}$, $\overline{p}$ en S^{n-1}.

Además, como $\|\overline{p}_k\| = 1$, para todo k en $\mathbb{N}$ se tiene que $\|\overline{p}\| = 1$ y también por el anterior teorema se satisface la relación:

$$(\overline{p}_k)^T y_k > (\overline{p}_k)^T x,$$

para toda k en $\mathbb{N}_1$ y para toda x en $\overline{S}$. Entonces, tomando el límite, en la última desigualdad cuando $k \to \infty$, obtenemos:

$$\overline{p}^T\overline{x} \geq \overline{p}^T x$$

para toda x en la cerradura de S. □

Observación 1. En la anterior demostración observamos también que si S es convexo implicará que $\partial S = \partial\left(\overline{S}\right)$. En efecto, sea $a \in \partial S$ entonces $a \notin \mathrm{int}(S) = \mathrm{int}\left(\overline{S}\right)$, pues S es convexo. Luego $a \notin \mathrm{int}\left(\overline{S}\right)$. Así, $a \in \left[\mathrm{int}\left(\overline{S}\right)\right]^c = \partial(\overline{S}) \cup \left(\mathrm{Ext}\,\overline{S}\right)$, en particular $a \in \partial\left(\overline{S}\right)$. Recíprocamente, sea $a \in \partial\left(\overline{S}\right)$, entonces para todo $\epsilon > 0$ $B(a,\epsilon) \cap \overline{S} \neq \emptyset$ y $B(a,\epsilon) \cap \left(\overline{S}\right)^c \neq \emptyset$. Tomando en cuenta que $\partial\left(\overline{S}\right) \subset \overline{S}$ entonces $a \in \overline{S} = S \cup \partial(S)$. Luego tenemos los siguientes casos:

i) $a \in \partial(S)$, por tanto $\partial\left(\overline{S}\right) \subseteq \partial(S)$.

ii) $a \in S$ entonces $B(a,\epsilon) \cap S \neq \emptyset$ (*). Y tomemos en cuenta que $B(a,\epsilon) \cap \left(\overline{S}\right)^c \neq \emptyset$, entonces $B(a,\epsilon) \cap (S \cup \partial S)^c \neq \emptyset$, luego $B(a,\epsilon) \cap S^c \cap \partial(S)^c \neq \emptyset$. En particular $B(a,\epsilon) \cap S^c \neq \emptyset$ (**). Luego de (*) y (**) tendremos por definición de frontera que $a \in \partial(S)$; es decir, $\partial\left(\overline{S}\right) \subseteq (S)$.

Corolario 1. Sea $S \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$ y $\overline{x} \notin S$. Entonces existe $p \neq 0$ en $\mathbb{R}^n$ tal que: $p^T(x - \overline{x}) \leq 0$, para toda x en la cerradura de S.

Demostración. Si $\overline{x}$ no está en la cerradura de S entonces por teorema 2 se tiene que existen $p \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ tal que $p^T\overline{x} > \alpha \geq p^T x$ para todo x en la cerradura de S. Luego $0 \geq p^T(x - \overline{x})$ para toda x en la cerradura de S.

Ahora si $\overline{x}$ está en la frontera de S, entonces se tiene el caso del teorema 3 y $p^T(x - \overline{x}) \leq 0$. □

Teorema 4. Separación entre dos conjuntos convexos. Sean S_1, S_2 dos convexos no vacíos de $\mathbb{R}^n$ y supongamos que $S_1 \cap S_2 = \emptyset$. Entonces existe H que separa a S_1 de S_2, es decir; existe p en $\mathbb{R}^n$ tal que $\sup\left\{p^T x : x \in S_1\right\} \leq \inf\left\{p^T x : x \in S_2\right\}$.

Demostración. Definamos el conjunto $S = S_1 - S_2 = \{x = x_1 - x_2 / x_1 \in S_1;\, x_2 \in S_2\}$, observemos que S es convexo. Además, como $S_1 \cap S_2 = \emptyset$ se tiene que el elemento 0 no está en S.

Por el corolario de separación entre un convexo y un punto, existe un vector p en $\mathbb{R}^n$ tal que $p^T(x - 0) = p^T x \leq 0$, para toda x en $\overline{S}$. Por tanto $p^T x = p^T(x_1 - x_2) \leq 0$, para todo x_1 en S_1, x_2 en S_2. Es decir;

$$\sup_{x_1 \in S_1}\left\{p^T x_1\right\} \leq \inf_{x_2 \in S_2}\left\{p^T x_2\right\}.$$

□

1.4. Funciones convexas

Definición 6. Sea $S \subset \mathbb{R}^n$ convexo y $f : S \to \mathbb{R}$ una función. Entonces, se dice que f es una función convexa, si dados $x, y \in S$ se tiene que:

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

para cualquier $\lambda \in [0,1]$.

Si la desigualdad se cumple estrictamente para $\lambda \in (0,1)$ y $x \neq y$, entonces la función es estrictamente convexa; es decir:

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$$

para todo $x \neq y$.

Teorema 5. Sea $S \neq \emptyset$ un conjunto convexo en $\mathbb{R}^n$ y $f : S \to \mathbb{R}$ una función diferenciable sobre S. Entonces, f es convexa sobre S si y sólo si

$$f(y) \geq f(x) + \nabla f(x)^T (y-x)$$

para todo $x, y \in S$. Además f es estrictamente convexa si y sólo si

$$f(y) > f(x) + \nabla f(x)^T (y-x)$$

para todo $x, y \in S$, $x \neq y$.

Demostración. Si f es una función convexa sobre S, por definición se sabe que, dados $x, y \in S$ y $t \in [0,1]$ se cumple la desigualdad siguiente:

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

Es decir:

$$f(x + t(y-x)) \leq f(x) + t\left[f(y) - f(x)\right].$$

De lo anterior decimos que $t\left[f(y) - f(x)\right] \geq f(x + t(y-x)) - f(x)$ y dividiendo por t obtendremos que

$$f(y) \geq f(x) + \frac{f(x + t(y-x)) - f(x)}{t},$$

para cualquier $t \in (0,1)$. Aplicando el límite cuando $t \to 0$ y del hecho de que

$$\lim_{t \to 0} \frac{f(x + t(y-x)) - f(x)}{t} = \nabla f(x)^T (y-x)$$

obtendremos que

$$f(y) \geq f(x) + \nabla f(x)^T (y-x).$$

Recíprocamente, supongamos que $f(y) \geq f(x) + \nabla f(x)^T(y-x)$ para todo $x, y \in S$. Sea $\overline{x} = \lambda x + (1-\lambda)y$; $\lambda \in (0,1)$. Como S es un conjunto convexo se tiene que $\overline{x} \in S$ y se satisface las siguientes relaciones:

$$
\begin{aligned}
f(x) &\geq f(\overline{x}) + \nabla f(\overline{x})^T(x-\overline{x}),\\
f(y) &\geq f(\overline{x}) + \nabla f(\overline{x})^T(y-\overline{x}),
\end{aligned}
$$

Multiplicando la primera desigualdad por λ, y multiplicando la segunda por $(1-\lambda)$ y sumando ambos miembros obtenemos:

$$
\begin{aligned}
\lambda f(x) + (1-\lambda) f(y) &\geq f(\overline{x}) + \nabla f(\overline{x})^T\left[\lambda(x-\overline{x}) + (1-\lambda)(y-\overline{x})\right]\\
&= f(\overline{x}) = f(\lambda x + (1-\lambda)y).
\end{aligned}
$$

Ahora mostremos la segunda parte del teorema. Sean $x, y \in S$, $x \neq y$ y $h = y - x$. Sea también $\lambda \in (0,1)$. Como f es estrictamente convexa, entonces f es convexa. Luego apliquemos la primera parte de éste teorema con x y $x + \lambda h$ en lugar de y. Entonces $\nabla f(x)^T(x+\lambda h - x) \leq f(x+\lambda h) - f(x)$. Luego

$$
\begin{aligned}
\nabla f(x)^T(\lambda h) &\leq f(x+\lambda h) - f(x)\\
&= f(x + \lambda y - \lambda x) - f(x)\\
&= f((1-\lambda)x + \lambda y) - f(x)\\
&< (1-\lambda) f(x) + \lambda f(y) - f(x)\\
&= f(x) - \lambda f(x) + \lambda f(y) - f(x)\\
&= \lambda(f(y) - f(x)).
\end{aligned}
$$

Dividiendo la anterior desigualdad por λ, y tomando el hecho de que $h = y - x$ se tendrá que

$$\nabla f(x)^T(y-x) < f(y) - f(x).$$

Por tanto

$$\nabla f(x)^T(y-x) + f(x) < f(y).$$

Recíprocamente, supongamos que $f(y) > f(x) + \nabla f(x)^T(y-x)$ para todo $x, y \in S$. Sea $\overline{x} = \lambda x + (1-\lambda)y$; $\lambda \in (0,1)$. Como S es un conjunto convexo se tiene que $\overline{x} \in S$ y se satisface las siguientes relaciones:

$$
\begin{aligned}
f(x) &> f(\overline{x}) + \nabla f(\overline{x})^T(x-\overline{x}),\\
f(y) &> f(\overline{x}) + \nabla f(\overline{x})^T(y-\overline{x}),
\end{aligned}
$$

Multiplicando la primera desigualdad por λ, y multiplicando la segunda por $(1-\lambda)$ y sumando ambos miembros obtenemos lo siguiente:

$$\begin{aligned}\lambda f(x)+(1-\lambda)f(y) &> f(\overline{x})+\nabla f(\overline{x})^T\left[\lambda(x-\overline{x})+(1-\lambda)(y-\overline{x})\right]\\ &= f(\overline{x})=f(\lambda x+(1-\lambda)y).\end{aligned}$$

□

Definición 7. Sea $f: S\to\mathbb{R}$ una función definida sobre un conjunto convexo no-vacío $S\subset\mathbb{R}^n$. Se dice que f es cuasi-convexa sobre S, si para cualquier pareja de puntos $x,y\in S$, se tiene que:

$$f(\lambda x+(1-\lambda)y)\leq \max\{f(x),f(y)\}$$

para todo $\lambda\in[0,1]$.

Definición 8. Sea $S\subset\mathbb{R}^n$ un conjunto convexo abierto y $f: S\to\mathbb{R}$ una función diferenciable sobre S. Se dice que f es una función pseudoconvexa si y sólo si para todo $x,y\in S$ $(x\neq y)$

$$f(y)<f(x) \text{ entonces } \langle\nabla f(x),(y-x)\rangle<0.$$

Ésto equivale a decir que:

$$\langle\nabla f(x),(y-x)\rangle\geq 0 \text{ entonces } f(y)\geq f(x).$$

1.5. Las condiciones de Karush, Kuhn y Tucker.(K.K.T.)

Definición 9. Sean $f:\mathbb{R}^n\to\mathbb{R}$ y $h:\mathbb{R}^n\to\mathbb{R}^m$ con $n>m$. La función de Lagrange se define por:

$$L(x,u)=f(x)+u^Th(x).$$

Considérese ahora el problema

$$\begin{aligned}&\min && f(x)\\ &\text{sujeto a} && x\in S\end{aligned}\tag{P}$$

Entonces, al conjunto S se le conoce como la Región Factible del problema (P). Además, si $x\in S$, se dice que x es un punto factible.

Definición 10. Supongamos que la región factible se especifica de la manera siguiente:

$$S=\{x/g(x)\leq 0; x\in X\subset\mathbb{R}^n\}.$$

Donde $g : \mathbb{R}^n \to: \mathbb{R}^m$ y $X \neq \emptyset$ es un conjunto abierto. Entonces (P) es un Problema de programación No-Lineal con función objetivo f y restricciones definidas por un conjunto de desigualdades:

$$\begin{aligned} &\min && f(x) \\ &\text{sujeto a} && g(x) \leq 0 \\ & && x \in X. \end{aligned} \tag{PNL}$$

Para esto, recuérdese que la función g está compuesta por funciones $g_i : \mathbb{R}^n \to \mathbb{R}$; $i = 1, 2, \ldots, m$. Entonces, si $\overline{x} \in X$ es un punto factible, defínase los conjuntos de índices siguientes:

$$\begin{aligned} I &= \{i/g_i(\overline{x}) = 0\} \equiv \text{ Conjunto de restricciones activas en } \overline{x} \\ \overline{I} &= \{i/g_i(\overline{x}) < 0\} \equiv \text{ Conjunto de restricciones pasivas en } \overline{x}. \end{aligned}$$

Teorema 6 (Condiciones necesarias de K.K.T.)**.** Sean $f : \mathbb{R}^n \to \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}^m$ funciones continuas sobre el conjunto abierto X. Entonces, si x^* es un mínimo local del problema (PNL) y f, g son diferenciables en x^*, y se cumple el requisito restriccional $\widetilde{\Gamma}(x^*) = \widetilde{D}(x^*)$, existe $\overline{u} \in \mathbb{R}^m$; $\overline{u} \neq 0$, tal que:

$$\begin{aligned} \nabla f(x^*) + \nabla g(x^*)\overline{u} = 0, \\ \overline{u} \geq 0, \\ \overline{u}^T g(x^*) = 0. \end{aligned}$$

Teorema 7 (Condiciones suficientes de K-T)**.** Sea $X \subset \mathbb{R}^n$ un conjunto abierto no-vacío y para (PNL) sean

$$f : \mathbb{R}^n \to \mathbb{R} \text{ una función pseudo-convexa en } \overline{x} \in X$$

$$g : \mathbb{R}^n \to \mathbb{R}^m \text{ compuesta por funciones } g_i : \mathbb{R}^n \to \mathbb{R} \text{ cuasi-convexas y diferenciables en } \overline{x} \in X.$$

Entonces, si se satisfacen las condiciones de $K.K.T.$ en $\overline{x}$, éste es un óptimo global de (PNL).

Demostración. Sea x una solución factible cualquiera de (PNL). Entonces $g_i(x) \leq g_i(\overline{x}) = 0$, para i en I y donde $I = \{i/g_i(\overline{x}) = 0\}$. Ahora, por la cuasi-convexidad de las funciones g_i en $\overline{x}$, para i en I, se tiene que $g_i(\lambda x + (1-\lambda)\overline{x}) = g_i(\overline{x} + \lambda(x - \overline{x})) \leq \max\{g_i(x), g_i(\overline{x})\} = g_i(\overline{x})$, luego

$$g_i(\overline{x} + \lambda(x - \overline{x})) - g_i(\overline{x}) \leq 0$$

para cualquier λ en $(0, 1)$ e i en I.

Aplicando límites cuando $\lambda \to 0$ se tendrá $\nabla g_i(\overline{x})^T(x - \overline{x}) \leq 0$ para i en I y para todo x en S, así $\nabla g_I(\overline{x})^T(x - \overline{x}) \leq 0$ para todo x en $S \ldots (*)$.

Observemos que si i no está en I entonces i está en $\overline{I} = \{i/g_i(\overline{x}) < 0\}$, por lo tanto $\overline{u}_{\overline{I}} = 0$.

Ahora bien, las condiciones de $K.K.T.$ se satisfacen en $\overline{x}$ y por lo tanto existe $\overline{u}_I \geq 0$ tal que

$$\nabla f(\overline{x}) + \nabla g_I(\overline{x})\overline{u}_I = 0,$$

$$\overline{u}_I^T g(\overline{x}) = 0.$$

Luego, de $(*)$ al multiplicar por $\overline{u}_I^T$ se tiene, para todo x en S, que:

$$\begin{aligned} \overline{u}_I^T \nabla g_I(\overline{x})^T (x - \overline{x}) &\leq 0, \\ (\nabla g_I(\overline{x})\overline{u}_I)^T (x - \overline{x}) &\leq 0, \\ -\nabla f(x)^T (x - \overline{x}) &\leq 0. \end{aligned}$$

Así $\nabla f(x)^T (x - \overline{x}) \geq 0$ para todo x en S, y como f es pseudoconvexa en $\overline{x}$ lo anterior implica que $f(x) \geq f(\overline{x})$ para todo x en S y $\overline{x}$ es óptimo global. $\square$

Capítulo 2

Un problema de dualidad en $\mathbb{R}$

En el presente capítulo asociamos un problema auxiliar, llamado dual, al problema de minimización original que será el problema primal. Estudiaremos la relación que existe entre el problema primal y el dual, para luego ver su interpretación geométrica. En la siguiente sección estudiamos un resultado de dualidad débil y la brecha de dualidad que este genera. En la quinta sección estudiaremos los teoremas de inexistencia de la brecha de dualidad y un teorema de dualidad fuerte para el problema general de programación no lineal.

2.1. El dual del problema general de programación no lineal

Considérese el problema general de programación no lineal primal (P):

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeto a} & g(x) \leq 0 \\ & h(x) = 0 \\ & x \in X. \end{array} \tag{P}$$

donde $g : \mathbb{R}^n \to \mathbb{R}^m$; $h : \mathbb{R}^n \to \mathbb{R}^s$; $f : \mathbb{R}^n \to \mathbb{R}$; $X \subset \mathbb{R}^n$. Entonces un problema dual viene dado por el siguiente problema dual lagrangeano (D):

$$\begin{array}{ll} \text{maximizar} & \theta(u, v) \\ \text{sujeto a} & u \geq 0 \end{array} \tag{D}$$

donde: $\theta(u, v) = \text{ínf}\,\{f(x) + \langle u, g(x)\rangle + \langle v, h(x)\rangle\,, x \in X\}\,.$

La función Lagrangeana dual $\theta(u, v)$ puede tomar el valor de $-\infty$ para alguna pareja (u, v) y que en el dual las restricciones $g(x)$ y $h(x)$ han sido incorporadas a la función

objetivo mediante multiplicadores de Lagrange u,v respectivamente. Además los multiplicadores asociados a las restriccciones $g(x) \leq 0$ están restringidos a no-negatividad mientras que los que están asociados a las h(x) no tienen restricciones de signo. Finalmente, nótese que el dual es en realidad un problema maximin ya que se trata de maximizar el ínfimo de una función por ésta última razón también se le conoce a veces como el 'problema dual max-min'.
También a un cierto problema primal pueden haber varios duales lagrangeanos, asociados, dependiendo de cuáles restricciones se manejan como $g(x) \leq 0$; $h(x) = 0$ y cuáles son tomadas en cuenta por el conjunto X. Dicha elección afecta el esfuerzo requerido para evaluar la función θ, al resolver el problema dual. Por lo tanto, una adecuada elección del conjunto X, dependerá de la estructura del problema.

2.2. Interpretación geométrica del problema dual

Por razones de simplificación y sin pérdida de generalidad se considera un ejemplo con una sola restricción de desigualdad $g(x) \leq 0$ y se supondrá que no existen restricciones de estricta igualdad.

Así, el primal será:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0 \\ & x \in X. \end{aligned} \tag{P}$$

Ahora, sea $G = \{(w, z)/w = g(x); z = f(x); x \in X\}$ de manera que G es la imagen de X bajo el mapeo (g, f). El problema primal pide encontrar un punto en G tal que $w \leq 0$ y que minimice z. En el plano (w, z) esto significa encontrar un punto en G a la izquierda del eje z, cuya ordenada al origen sea mínima.

Ahora, supongamos que $u \geq 0$ es un cierto número dado. Para encontrar $\theta(u)$ se requiere el mínimo de $f(x) + ug(x)$ sobre toda $x \in X$. Entonces si $w = g(x)$ y $z = f(x)$ para $x \in X$ se quiere minimizar $z + uw$ sobre cualquier pareja $(w, z) \in G$. Esto equivale a deslizar la línea $z + uw = \alpha$ en forma paralela a sí misma lo mas abajo posible de modo que ésta soporte a G. Así la línea es tangente a G en el punto (w_1, z_1) y $\theta(u)$ es la ordenada al origen donde la línea intersecta al eje z (ahí es donde $W = 0$). Así el problema dual se puede interpretar como encontrar la pendiente del hiperplano de soporte al conjunto G, cuya ordenada z al origen sea máxima. En la figura 2.1 el hiperplano pasa por (w^*, z^*) y su pendiente es $-u^*$ y $\theta(u^*) = z^*$. Además, notemos que las funciones objetivo del dual y el primal tiene el mismo valor en el óptimo.

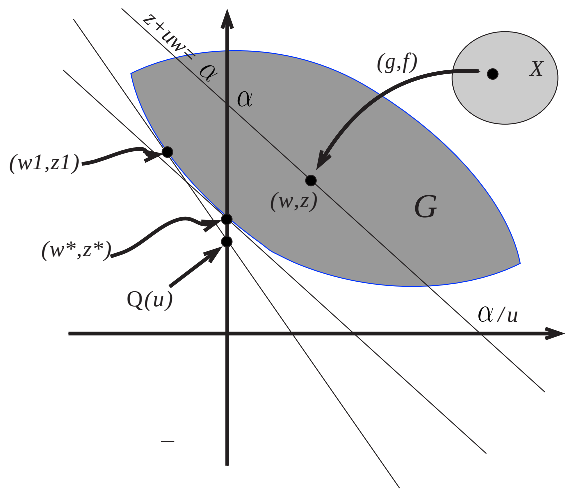


Figura 2.1: Interpretación geométrica

Ejemplo 2. Considérese el problema de programación no lineal siguiente:

$$\begin{aligned} \min \quad & \langle x, x\rangle \\ \text{s.a.} \quad & \langle p, x\rangle \leq \alpha \end{aligned} \tag{P}$$

Usemos Khun y Tucker para resolver el problema, luego se tiene que $f(x) = \langle x, x\rangle$ y $g(x) = \langle p, x\rangle - \alpha$ son convexas, y construimos la función Lagrangeana como: $L(x^*, u) = \langle x^*, x^*\rangle + u\left[\langle p, x^*\rangle - \alpha\right]$ luego

$$\frac{\partial L}{\partial x^*} = \left[2x_1 + u_1p_1 \; 2x_2 + u_2p_2 \ldots 2x_n + u_np_n\right] = \left[0 \; 0 \ldots 0\right].$$

Entonces $2x_i + u_ip_i = 0$ para todo $i = 1, 2, \ldots, n$, es decir; $x_i = -\frac{u_ip_i}{2}$, para todo $i = 1, 2, \ldots, n$.

Por otro lado tenemos que

$$\frac{\partial L}{\partial u} = \langle p, x^*\rangle - \alpha = 0$$

$$\langle p, x^*\rangle = \left\langle p, \frac{-up}{2}\right\rangle = \frac{-u}{2}\langle p, p\rangle = \alpha,$$

Luego

$$u = \frac{-2\alpha}{\|p\|^2}.$$

Y el valor óptimo será:

$$x^* = \frac{\alpha}{\|p\|^2}p.$$

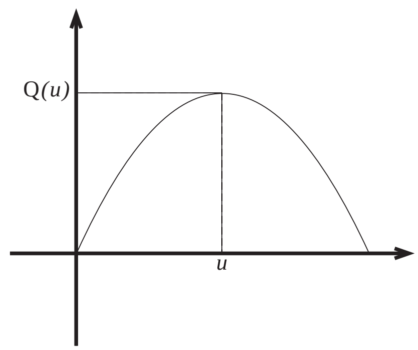


Figura 2.2: Función dual

Ahora veamos cómo se comporta la función dual. Sean $g(x) = \langle p, x\rangle - \alpha$ y $X \equiv \{x/x \geq 0\}$, entonces la función Dual es:

$$\theta(u) = \inf_{x\in X} \{\langle x, x\rangle + u(\langle p, x\rangle - \alpha)\},$$

es decir;

$$\theta(u) = \inf_{x\in X} \{\langle x, x\rangle + u\,\langle p, x\rangle\} - \alpha u,$$

Tomando la derivada obtendremos que $x^* = -\dfrac{u}{2}p$, por tanto

$$\begin{aligned}
\theta(u) &= \left\langle -\frac{up}{2}, -\frac{up}{2} \right\rangle + u \left\langle p, -\frac{up}{2} \right\rangle - \alpha u \\
&= \frac{u^2}{4}\langle p, p\rangle - \frac{u^2}{2}\langle p, p\rangle - \alpha u \\
&= \|p\|^2 \left[\frac{u^2 - 2u^2}{4}\right] - \alpha u \\
&= \|p\|^2 \left(-\frac{u^2}{4}\right) - \alpha u \\
&= -\left(\frac{u}{2}\right)^2 \|p\|^2 - \alpha u.
\end{aligned}$$

Ahora si $u \geq 0$ y $p \leq 0$ entonces $x^* \geq 0$. Además, es importante decir que si $\alpha < 0$ la función dual se comporta así, ver 2.2:

Notemos también que: $\max_{u\geq 0} \theta(u) = \left(\dfrac{\alpha}{\|p\|}\right)^2 = \theta(u^*)$ cuando $u^* = \frac{2|\alpha|}{\|p\|^2}$ y finalmente concluimos que

$$\theta(u^*) = f(x^*).$$

2.3. El resultado débil de dualidad

En ésta sección se investiga la primera relación entre el primal y el dual, que se conoce como el teorema débil. En éste teorema se demuestra formalmente el resultado que se vio intuitivamente en la interpretación geométrica del dual, de que la función objetivo primal acota superiormente la del dual y viceversa. Varios resultados importantes se derivan como corolarios.

Teorema 8. Sea x una solución factible para el primal (ie, $g(x) \leq 0$, $h(x) = 0$; $x \in X$) y sea (u, v) factible para el dual (ie, $u \geq 0$). Entonces $f(x) \geq \theta(u, v)$.

Demostración. Como x es factible para el primal se tiene que:$g(x) \leq 0$ y $h(x) = 0$. Y para $u \geq 0$ tendremos que

$$\langle u, g(x) \rangle \leq 0$$

luego

$$\langle u, g(x) \rangle + \langle v, h(x) \rangle \leq 0,$$

$$f(x) + \langle u, g(x) \rangle + \langle v, h(x) \rangle \leq f(x).$$

Entonces

$$\begin{aligned} \theta(u, v) &= \inf_{y \in X} \{f(y) + \langle u, g(y) \rangle + \langle v, h(y) \rangle\} \\ &\leq f(x) + \langle u, g(x) \rangle + \langle v, h(x) \rangle \\ &\leq f(x). \end{aligned}$$

□

Corolario 2. $\inf \{f(x)/x \in X, g(x) \leq 0, h(x) = 0\} \geq \sup \{\theta(u, v)/u \geq 0\}$.

Demostración. Sean $u \geq 0$, $\{x \in X, g(x) \leq 0, h(x) = 0\}$. Entonces se tiene que cualquier $f(x) \in \{f(x)/x \in X, g(x) \leq 0, h(x) = 0\}$ es cota superior de $\{\theta(u, v), u \geq 0\}$; es decir, el $\sup \{\theta(u, v); u \geq 0\}$ es una cota inferior de $\{f(x)/x \in X, g(x) \leq 0, h(x) = 0\}$, por tanto

$$\inf \{f(x)/x \in X, g(x) \leq 0, h(x) = 0\} \geq \sup \{\theta(u, v), u \geq 0\}.$$

□

Corolario 3. Si $\overline{x}$ y $(\overline{u}, \overline{v})$ son tales que $\overline{x}$ es factible para (P) y $(\overline{u}, \overline{v})$ es factible para (D) y además $f(\overline{x}) \leq \theta(\overline{u}, \overline{v})$, entonces $\overline{x}$, $(\overline{u}, \overline{v})$ son óptimos para (P) y (D) respectivamente.

Demostración. Por el teorema de dualidad débil se cumple que $f(\overline{x}) \geq \theta(\overline{u}, \overline{v})$ y dada la condición $f(\overline{x}) \leq \theta(\overline{u}, \overline{v})$ implicará que $f(\overline{x}) = \theta(\overline{u}, \overline{v})$. Luego, por corolario anterior, tenemos que $\inf\{f(x)/x \in X, g(x) \leq 0, h(x) = 0\} \geq \theta(\overline{u}, \overline{v}) = f(\overline{x})$. Así $\overline{x}$ es óptimo del primal.

Por otro lado $\theta(\overline{u}, \overline{v}) = f(\overline{x}) \geq \sup\{\theta(u, v)/u \geq 0\}$, luego $\theta(\overline{u}, \overline{v}) \geq \sup\{\theta(u, v)/u \geq 0\}$. Por lo tanto $(\overline{u}, \overline{v})$ es óptimo para el dual. □

Corolario 4. Si $\inf\{f(x)/x \in X, g(x) \leq 0, h(x) = 0\} = -\infty$, entonces $\theta(u, v) = -\infty$, para todo $u \geq 0$.

Demostración. Sea $u \geq 0$. Luego $\inf\{f(x)/x \in X, g(x) \leq 0, h(x) = 0\} \geq \theta(u, v)$ y por hipótesis $-\infty \geq \theta(u, v)$. Por tanto $\theta(u, v) = -\infty$. □

Corolario 5. Si $\sup\{\theta(u, v)/u \geq 0\} = +\infty$, entonces el primal no tiene solución factible.

Demostración. Por corolario $\sup\{\theta(u, v)/u \geq 0\} \leq \inf\{f(x)/x \in X, g(x) \leq 0, h(x) = 0\}$. Entonces $+\infty \leq \inf\{f(x)/x \in X, g(x) \leq 0, h(x) = 0\}$. Así; (P) no tiene solución factible. □

2.4. Funciones convexas diferenciables

Considérese la pareja de problemas Primal-Dual siguiente:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0 \\ & x \in X. \end{aligned} \tag{Primal}$$

$$\begin{aligned} \max \quad & \theta(u) \\ \text{s.a.} \quad & \theta(u) = \inf L(x, u) \\ & x \in X. \end{aligned} \tag{Dual}$$

Donde $X \subset \mathbb{R}^n$ convexo y abierto y $L(x, u) = f(x) + \langle u, g(x) \rangle$ tal que, $f : \mathbb{R}^n \to \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}^m$ son funciones diferenciables. fvwWWWW Entonces dado u, el ínfimo de $L(x, u)$ respecto a x si existe, se logra en un punto $\overline{x} \in X$ donde desaparece el gradiente del lagrangeano $L(x, u)$; es decir,

$$\nabla f(\overline{x}) + \nabla g(\overline{x})u = 0.$$

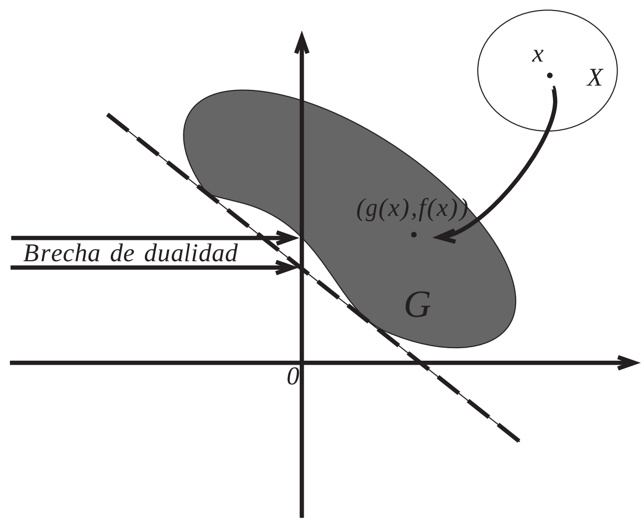


Figura 2.3: Brecha de dualidad

Se sabe que si f y g son convexas, ésta condición es también suficiente. Luego el dual se puede escribir:

$$\begin{aligned} \max_{x,u} \quad & L(x,u) \\ \text{s.a.} \quad & \nabla f(x) + \nabla g(x)u = 0 \\ & u \geq 0 \\ & x \in X. \end{aligned} \tag{DD}$$

2.5. La brecha de dualidad y los teoremas de inexistencia de la brecha y de dualidad conversa.

Con las condiciones del teorema de dualidad débil lo único que se puede demostrar es que la función objetivo del primal y del dual tienen la relación

$$f(x) \geq \theta(u,v),$$

para puntos factibles a sus respectivos problemas. Contrario a lo que sucede en el caso de programación lineal, pueden haber casos en que si, x^* y (u^*, v^*) son óptimos para sus respectivos problemas, la relación se cumpla con estricta desigualdad (ie, $f(x^*) > \theta(u^*, v^*)$). En éste caso se dice que existe una brecha de dualidad entre el primal y el dual.

Geométricamente, la brecha se ilustra en la figura 2.3 para el caso en que se tiene una sola restricción del tipo $g(x) \leq 0$.

El problema surge porque hay una no convexidad inconveniente en el conjunto G. Notemos, sin embargo, que la convexidad del conjunto G no es una condición necesaria para que no exista una brecha de dualidad.

Las condiciones de ausencia de una brecha y de dualidad conversa; es decir, de establecer condiciones bajo las cuales las soluciones óptimas de ambos problemas proporcionan el mismo valor a las dos funciones objetivo y cuando los problemas tienen una misma solución.

Veamos el resultado de Wolfe que fue uno de los primeros obtenidos en tratar éstos problemas.

Teorema 9. Considérese la pareja de problemas primal-dual siguiente:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0 \\ & x \in X. \end{aligned} \tag{Primal}$$

$$\begin{aligned} \max_{x,u} \quad & L(x,u) \\ \text{s.a.} \quad & \nabla f(x) + \nabla g(x) u = 0 \\ & u \geq 0. \end{aligned} \tag{Dual}$$

Donde $X \subset \mathbb{R}^n$, es un conjunto abierto, $L(x,u) = f(x) + u^T g(x)$ y $f : \mathbb{R}^n \to \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}^m$, son funciones convexas y diferenciables. Sea x^* una solución del primal tal que se satisface el requisito restriccional en este punto. Entonces, existe u^* tal que resuelven el problema dual y tal que $f(x^*) = \theta(u^*) = L(x^*, u^*)$.

Demostración. Si x^* es una solución del primal donde se cumple el requisito restriccional, tiene que cumplirse en x^* las condiciones de K.K.T.

Así existe $u^* \in \mathbb{R}^m$ tal que:

$$\begin{aligned} \nabla f(x^*) + \nabla g(x^*) u^* &= 0 \\ \langle u^*, g(x^*) \rangle &= 0 \\ g(x^*) &\leq 0 \\ u^* &\geq 0. \end{aligned}$$

Ahora, consideremos $u \in \mathbb{R}^m$ tal que $u \geq 0$ y examinemos:

$$\begin{aligned}\theta(u^*) - \theta(u) &= f(x^*) + \langle u^*, g(x^*)\rangle - f(x) - \langle u, g(x)\rangle \\ &= f(x^*) - f(x) + \langle u^*, g(x^*)\rangle - \langle u, g(x)\rangle .\end{aligned}$$

Entonces, por convexidad de f y dado que (x^*, u^*) satisfacen las condiciones de K.K.T. se tiene que $\langle u^*, g(x^*)\rangle = 0$ y por lo tanto:

$$\begin{aligned}\theta(u^*) - \theta(u) &= f(x^*) - f(x) - \langle u, g(x)\rangle \\ &\geq \langle \nabla f(x^*), (x^* - x)\rangle - \langle u, g(x)\rangle .\end{aligned}$$

análogamente por convexidad de g se tiene que:

$$\begin{aligned}g(x^*) - g(x) &\geq \langle \nabla g(x), x^* - x\rangle \\ g(x^*) &\geq g(x) + \langle \nabla g(x), x^* - x\rangle \\ -g(x) &\geq \langle \nabla g(x), x^* - x\rangle - g(x^*).\end{aligned}$$

y esto proporciona:

$$\begin{aligned}\theta(u^*) - \theta(u) &\geq \langle \nabla f(x), (x^* - x)\rangle - \langle u, g(x)\rangle \\ &\geq \langle \nabla f(x), (x^* - x)\rangle + \langle u, [\langle \nabla g(x), (x^* - x)\rangle - g(x^*)]\rangle \\ &= \langle \nabla f(x), (x^* - x)\rangle - \langle u, g(x^*)\rangle + \langle u, (\langle \nabla g(x), (x^* - x)\rangle)\rangle \\ &= \langle \nabla f(x), (x^* - x)\rangle - \langle u, g(x^*)\rangle \\ &+ \langle u, \langle \nabla g(x), x^*\rangle\rangle - \langle u, \langle \nabla g(x), x\rangle\rangle \\ &= \langle \nabla f(x), (x^* - x)\rangle - \langle u, g(x^*)\rangle + \langle \nabla g(x)u, x^*\rangle - \langle \nabla g(x)u, x\rangle \\ &= \langle \nabla f(x), (x^* - x)\rangle - \langle u, g(x^*)\rangle + \langle \nabla g(x)u, (x^* - x)\rangle \\ &= \langle \nabla f(x) + \nabla g(x)u, (x^* - x)\rangle - \langle u, g(x^*)\rangle \\ &= -\langle u, g(x^*)\rangle \\ &\geq 0,\end{aligned}$$

pues $u \geq 0$, $g(x^*) \leq 0$ y $\nabla f(x) + \nabla g(x)u = 0$ por ser (x, u) factible para el dual diferenciable, así: $\theta(u^*) \geq \theta(u)$.

Por lo tanto $\theta(u^*) = \max \theta(u)$ sujeto a que $\theta(u) = \inf \{f(x) + \langle u, g(x)\rangle / x \in X\}$, ésto implica que:

$$\max_{u\geq 0} \theta(u) = \max_{u\geq 0} \left\{ \inf_{x\in X} [f(x) + \langle u, g(x)\rangle] \right\} = f(x^*) + \langle u^*, g(x^*)\rangle = L(x^*, u^*).$$

Además, sabemos que $\langle u^*, g(x^*)\rangle = 0$ luego $\theta(u^*) = L(x^*, u^*)$ y $L(x^*, u^*) = f(x^*)$. Por lo tanto: $\theta(u^*) = f(x^*) = L(x^*, u^*)$. $\square$

Observación 2. Como podemos ver, éste teorema da un resultado con condiciones de ausencia de brecha en un caso, donde todas las funciones son diferenciables y convexas y el resultado podría anticiparse fácilmente. Ahora analicemos condiciones bajo las cuales no solo no existe la brecha, sino que los dos problemas tiene una misma solución. Es decir se busca establecer el converso del teorema anterior, en donde la solución del dual también es una solución del primal; de ahí el nombre de dualidad conversa. Para obtener tal teorema hay que pagar un precio como se ilustra en el siguiente resultado atribuido a Mangasarian.

Teorema 10. Sea $X \subset \mathbb{R}^n$ un conjunto abierto y f,g funciones convexas y diferenciables sobre X. Sea $\overline{x}$ una solución del problema primal como se presentó en el teorema anterior, tal que se satisface el requisito restriccional en $\overline{x}$. Si $(\widehat{x}, \widehat{u})$ es una solución dual, y $L(\widehat{x}, \widehat{u}) = f(\widehat{x}) + \langle \widehat{u}, g(\widehat{x}) \rangle$ es una función estrictamente convexa en $\widehat{x}$, entonces $\widehat{x} = \overline{x}$ y por lo tanto, es una solución del problema primal.

Demostración. $L(\widehat{x}, \widehat{u})$ es estrictamente convexa en $\widehat{x}$ si f es estrictamente convexa en $\widehat{x}$ o si para algún índice i, se tiene que el elemento $\widehat{u_i} > 0$ y la correspondiente función g_i de g es estrictamente convexa. Mostremos este hecho, sean x, y en X y λ en $(0,1)$. Considerando el hecho de que f es estrictamente convexa tendremos:

$$\begin{aligned}
L((1-\lambda)x + \lambda y, \widehat{u}) &= f((1-\lambda)x + \lambda y) + \langle \widehat{u}, g((1-\lambda)x + \lambda y) \rangle \\
&< (1-\lambda)f(x) + \lambda f(y) + \sum_{i=1}^{n} u_i g_i((1-\lambda)x + \lambda y) \\
&\leq (1-\lambda)f(x) + \lambda f(y) + \sum_{i=1}^{n} u_i g_i(x) + \lambda \sum_{i=1}^{n} u_i g_i(y) \\
&= (1-\lambda)f(x) + \lambda f(y) + (1-\lambda)\langle \widehat{u}, g(x) \rangle + \lambda \langle \widehat{u}, g(y) \rangle \\
&= (1-\lambda)L(x, \widehat{u}) + \lambda L(y, \widehat{u}),
\end{aligned}$$

Ahora mostremos el caso en el que cada g_i de g es estrictamente convexa: antes notemos el hecho de que

$$g_i((1-\lambda)x + \lambda y) < (1-\lambda)g_i(x) + \lambda g_j(y)$$

para todo $j \neq i$. Ahora, volviendo a la demostración;

$$\begin{aligned}
L((1-\lambda)x + \lambda y, \widehat{u}) &= f((1-\lambda)x + \lambda y) + \langle \widehat{u}, g((1-\lambda)x + \lambda y) \rangle \\
&\leq (1-\lambda)f(x) + \lambda f(y) + \sum u_i g_i((1-\lambda)x + \lambda y) \\
&= (1-\lambda)f(x) + \lambda f(y) + \sum \widehat{u}_k g_k((1-\lambda)x + \lambda y) \\
&< (1-\lambda)f(x) + \lambda f(y) + \sum \widehat{u}_k \left((1-\lambda)g_k(x) + \lambda g_k(y)\right) \\
&= (1-\lambda)f(x) + \lambda f(y) + (1-\lambda)\sum \widehat{u}_k g_k(x) + \lambda \sum \widehat{u}_k g_k(y) \\
&= (1-\lambda)L(x, \widehat{u}) + \lambda L(y, \widehat{u}).
\end{aligned}$$

En ambos casos hemos mostrado que L es estrictamente convexa.

Ahora mostremos el teorema, supongamos $\widehat{x} \neq \overline{x}$.

Como $\overline{x}$ es una solución al primal y satisface el requisito y por teorema anterior existe un $\overline{u} \in \mathbb{R}^m$, tal que la pareja $(\overline{x}, \overline{u})$ es una solución del problema dual, además de que satisface las condiciones de K.K.T.

Así se tiene que: $L(\overline{x}, \overline{u}) = L(\widehat{x}, \widehat{u}) = \theta(\widehat{u}) = \max_u \theta(u)$ sujeto a $\theta(u) = \inf_{x \in X} \{L(x, u)\}$.

Ahora como $(\widehat{x}, \widehat{u})$ resuelven el problema dual entonces satisface: $\nabla_x L(\widehat{x}, \widehat{u}) = \nabla f(\widehat{x}) + \nabla g(\widehat{x})\widehat{u} = 0$, como $L(\widehat{x}, \widehat{u})$ es estrictamente convexa en $(\widehat{x}, \widehat{u})$ se tendrá por teorema que:

$$L(\overline{x}, \widehat{u}) - L(\widehat{x}, \widehat{u}) > \langle \nabla_x L(\widehat{x}, \widehat{u}), (\overline{x} - \widehat{x}) \rangle = 0.$$

Luego $L(\overline{x}, \widehat{u}) > L(\widehat{x}, \widehat{u}) = L(\overline{x}, \overline{u})$, y por definición de Lagrange:

$$\begin{aligned} f(\overline{x}) + \widehat{u}^T g(\overline{x}) &> f(\overline{x}) + \overline{u}^T g(\overline{x}) \\ \widehat{u}^T g(\overline{x}) &> \overline{u}^T g(\overline{x}) = 0 \\ \widehat{u}^T g(\overline{x}) &> 0, \end{aligned}$$

que es una contradicción,pues $\widehat{u} \geq 0$ y $g(\overline{x}) \leq 0$, por lo tanto $\widehat{x} = \overline{x}$.

Además se tiene que: $f(\overline{x}) = f(\overline{x}) + \overline{u}^T g(\overline{x}) = L(\overline{x}, \overline{u}) = L(\widehat{x}, \widehat{u})$, es decir $L(\widehat{x}, \widehat{u})$ es una solución del problema primal. $\square$

Teorema 11. (Generalización del teorema de Hanson y Huard).

Sea $X \subset \mathbb{R}^n$ un conjunto abierto y f, g diferenciables sobre X. Sea $(\widehat{x}, \widehat{u})$ una solución local del dual y supóngase que f es pseudoconvexa y g es cuasiconvexa en $\widehat{x}$. Supóngase que:

a) $L(\widehat{x}, \widehat{u})$ es doblemente diferenciable en $\widehat{x}$ y la matriz hessiana $\nabla_x^2 L(\widehat{x}, \widehat{u})$ es no singular.

b) Existe un conjunto abierto $V \subset \mathbb{R}^m$, tal que $\widehat{u} \in V$ y una función diferenciable $\phi : V \to \mathbb{R}^n$ tal que:

$$\begin{aligned} &i) \widehat{x} = \phi(\widehat{u}), \\ &ii) \phi(u) \in X; \nabla_x L(\phi(u), u) = 0, \text{ para todo } u \in V. \end{aligned}$$

Entonces $\widehat{x}$ es una solución del primal y $f(\widehat{x}) = L(\widehat{x}, \widehat{u})$.

Demostración. Si $(\widehat{x}, \widehat{u})$ es un óptimo local del dual, entonces existe $B[(\widehat{x}, \widehat{u}), \delta]$ tal que $\theta(\widehat{u}) = L(\widehat{x}, \widehat{u}) = \max \theta(u)$ sujeto a $\theta(u) = \inf_{x \in X} L(x, u)$, para todo $(x, u) \subset B[(\widehat{x}, \widehat{u}), \delta] \cap Y = B[(\widehat{x}, \widehat{u}), \delta] \cap (X \times \mathbb{R}^m)$.

Además, como f y g son diferenciables, una condición necesaria para que (x,u) sea una solución del dual es que: $\nabla f(x) + \nabla g(x)u = 0$. Observemos en la hipótesis que $b)$ es consecuencia de $a)$ esto por el teorema de la función implícita, pero $a)$ no sigue del supuesto $b)$, luego por $b)$ se tendrá: $(\widehat{x},\widehat{u}) = [\phi(\widehat{u}),\widehat{u}] \in \{[\phi(u),u]\,/u \in V, u \geq 0\} \subset Y$.

Entonces, como $L[\phi(\widehat{u}),\widehat{u}] = L(\widehat{x},\widehat{u}) = \text{máx}_{(x,u)\in Y} L(x,u)$, se tendrá $L[\phi(\widehat{u}),\widehat{u}] = \text{máx}_{u\in V}\{L[\phi(u),u]\,/u \geq 0\}$, éste último es un problema de optimización es decir,

$$\begin{aligned} \text{máx} \quad & L[\phi(u),u] \\ \text{s.a.} \quad & u \geq 0, \end{aligned} \tag{D}$$

es equivalente al problema primal

$$\begin{aligned} \text{mín} \quad & -L[\phi(u),u], \\ \text{s.a.} \quad & -u \leq 0, \end{aligned} \tag{P}$$

y por K.K.T. si u es una solución de (P) implica que existe $\overline{u} \neq 0$ tal que:

$$\begin{aligned} \nabla_u(-L[\phi(u),u]) + \nabla_u(-u)\cdot\overline{u} = 0 \\ \overline{u} \geq 0 \\ \overline{u}^T(-u) \leq 0 \\ -u \leq 0. \end{aligned}$$

Luego $\nabla_u(-u)\cdot\overline{u} = \langle -1,\overline{u}\rangle = -\overline{u} \leq 0$. Además notemos que: $\nabla_u(-L[\phi(u),u]) = -\nabla_u(L[\phi(u),u])$ y por la regla de la cadena:

$$\begin{aligned} \partial(L[\phi(u),u]) &= \partial(L\circ g)(u) \\ &= \partial L(g(u))\cdot\partial g(u) \\ &= (\nabla_x L(x,u)\nabla_u L(x,u))_{g(u)} \begin{pmatrix} \nabla\phi(u)|_u \\ I \end{pmatrix} \\ &= \nabla_x L(x,u)\big|_{g(u)}\cdot\nabla\phi(u)\big|_u + \nabla_u L(x,u)\big|_{g(u)}\cdot I \\ &= \nabla_x L(x,u)\big|_{g(u)}\cdot\nabla\phi(u)\big|_u + \nabla_u L(x,u)\big|_{g(u)}. \end{aligned}$$

Luego

$$\begin{aligned} -\nabla_u(L[\phi(u),u]) - \nabla_u(u)\cdot\overline{u} = 0, \\ \nabla_u(L[\phi(u),u]) + \nabla_u(u)\cdot\overline{u} = 0, \\ \nabla_x L(x,u)\,|_{g(u)}\cdot\nabla\phi(u)\,|_u + \nabla_u L(x,u)\,|_{g(u)} = -\overline{u} \leq 0. \end{aligned}$$

Así por b)

$$\nabla_u L(x,u)\,|_{g(u)} \leq 0$$

y por teorema de la función implícita se sabe que se satisface la relación

$$\nabla_x L(\widehat{x}, \widehat{u}) \nabla \phi(\widehat{u}) \mid_u = -\nabla_u L(\widehat{x}, \widehat{u})$$

en la solución local del dual $(\widehat{x}, \widehat{u})$ de donde se tiene que

$$\widehat{u}^T \left[\nabla_x L(\widehat{x}, \widehat{u}) \nabla_u \phi(\widehat{u}) + \nabla_u L(\widehat{x}, \widehat{u}) \right] = 0.$$

Por lo tanto $\widehat{u}^T g(\widehat{x}) = 0$ y como $(\widehat{x}, \widehat{u})$ es una solución del dual, se satisfacen las siguientes relaciones:

$$\begin{aligned} \nabla f(\widehat{x}) + \nabla g(\widehat{x}) \widehat{u} &= 0 \\ \widehat{u}^T g(\widehat{x}) &= 0 \\ \widehat{u} &\geq 0 \\ g(\widehat{x}) &\leq 0. \end{aligned}$$

Así $\widehat{x}$ es factible para el primal y se satisfacen las condiciones de K.K.T. con $\widehat{u}$. Finalmente, como f y g son diferenciables y además f es pseudoconvexa y g es cuasi-convexa, por el teorema de condiciones suficientes de K.K.T. $\widehat{x}$ es una solución óptima global para el primal y por lo tanto para todo $x \in X, f(x) \geq f(\widehat{x}) = \theta(\widehat{u}) = L(\widehat{x}, \widehat{u})$. □

Nota 1. Un hecho curioso es que contrariamente a lo que pudiera pensarse, si $\overline{x}$ es una solución del primal bajo las condiciones del teorema anterior no necesariamente existe $\overline{u}$ tal que $(\overline{x}, \overline{u})$ es una solución del dual, ni siquiera cuando g es lineal. Tampoco se satisface la condición débil de dualidad, es decir; que $f(\overline{x}) \geq L(x, u)$ para cualquier pareja (x, u) que sea factible para el dual. Esto se ilustra en el siguiente ejemplo:

Ejemplo 3. sea $f(x) = -e^{-x^2}$ pseudoconvexa.
Así,

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & -e^{-x^2} \\ \text{s.a.} \quad & x \leq -1, \end{aligned} \tag{P}$$

donde $g(x) = x + 1 \leq 0$. El dual de éste problema es:

$$\begin{aligned} \max_{x,u} \quad & -e^{-x^2} + u(x+1) \\ \text{s.a.} \quad & 2xe^{-x^2} + u = 0 \\ & u \geq 0. \end{aligned} \tag{DD}$$

La solución óptima del primal es $\overline{x} = -1$, pues; si $L(x, u) = -e^{-x^2} + (x-1)u = 0$ entonces $\nabla_x L(x, u) = e^{-x^2} + u = 0$ y $-e^{-x^2} = u$. Ahora para $x = -1$, que es el óptimo, se tendrá $u = \frac{2}{e}$.

Luego $\overline{x} = -1$ y $u = \frac{2}{e}$ satisfacen las condiciones K.K.T.

$$\begin{aligned} \nabla f(\overline{x}) + \nabla g(\overline{x})\overline{u} &= 0 \\ \overline{u}^T g(\overline{x}) &= 0, \\ \overline{u} &\geq 0, \\ g(\overline{x}) &\leq 0. \end{aligned}$$

Notemos que el dual exige que $u = -2xe^{-x^2}$. Así el dual puede escribirse de la siguiente manera:

$$\begin{aligned} \text{máx} \quad & -e^{-x^2} - 2xe^{-x^2}(x+1) \\ \text{s.a.} \quad & x \leq 0. \end{aligned} \tag{D}$$

es decir; $\text{máx}_{x\leq 0} - [1 + 2x + 2x^2]\, e^{-x^2}$. Ahora, notemos también que $1 + 2x + 2x^2 \geq -\frac{1}{2}$, pues el mínimo de ésta función está en $x = -\dfrac{1}{2}$. Además $-[1 + 2x + 2x^2]\, e^{-x^2} \to 0$ cuando $e^{-x^2} \to 0$.

Finalmente consideremos el punto: $(\widetilde{x}, \widetilde{u}) = (-5, 10e^{-25})$ que es factible para el dual, notese que:

$$L(\widetilde{x}, \widetilde{u}) = -e^{-25} - 2(-5)e^{-25}(-4) = -e^{-25} - 40e^{-25} = -41e^{-25}.$$

y que $f(\overline{x}) = -e^{-(-1)^2} = -e^{-1}$. Así $f(x) < L(\widetilde{x}, \widetilde{u})$.

Y en éste ejemplo se ve que ni siquiera se satisface la condición débil de dualidad. Lo que sucede en éste caso que al no ser f y g funciones convexas,la condición $\nabla f(x) + \nabla g(x)u = 0$ es necesaria mas no suficiente para un mínimo y por lo tanto, la forma equivalente que se ha dado al dual tampoco es suficiente en éste caso.

Lema 3. Sea $X \subset \mathbb{R}^n$ un conjunto convexo no vacío. Sean $f : \mathbb{R}^n \to \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}^m$ convexas y sea $h : \mathbb{R}^n \to \mathbb{R}^s$ una función afín; es decir, h es de la forma $h(x) = Ax - b$, donde A es una matriz de $(n \times s)$ y b es un vector de dimensión s. Considérese los siguientes dos sistemas:

1) $f(x) < 0$, $g(x) \leq 0$, $h(x) = 0$, para alguna $x \in X$.

2) $u_0 f(x) + u^T g(x) + v^T h(x) \geq 0$, tal que $x \in X$, $(u_0, u) \geq 0$, $(u_0, u, v) \neq 0$.

Entonces, si el sistema 1) no tiene una solución x, el sistema 2) tiene una solución (u_0, u, v). Lo contrario es cierto si $u_0 > 0$.

Demostración. Supongamos que el sistema 1) no tiene solución x, considérese el siguiente conjunto: $C = \{(p, q, r)/p > f(x), q \geq g(x), r = h(x), \text{para alguna } x \in X\}$.

Probemos que el conjunto C es convexo. Sean (p_1, q_1, r_1) y (p_2, q_2, r_2) en C, por demostrar que para $\lambda \in (0,1)$ se tendrá $(p,q,r) = \lambda(p_1,q_1,r_1)+(1-\lambda)(p_2,q_2,r_2) \in C$. En efecto; como $(p_1,q_1,r_1) \in C$ y $(p_2,q_2,r_2) \in C$ se tendrá: $p_1 > f(x), q_1 \geq g(x), r_1 = h(x), x \in X$ y $p_2 > f(y), q_2 \geq g(y), r_2 = h(y), y \in X$. Sea

$$\begin{aligned}(p,q,r) &= \lambda(p_1,q_1,r_1) + (1-\lambda)(p_2,q_2,r_2)\\ &= (\lambda p_1 + (1-\lambda)p_2, \lambda q_1 + (1-\lambda)q_2, \lambda r_1 + (1-\lambda)r_2),\end{aligned}$$

luego,

$$\begin{aligned}
p &= \lambda p_1 + (1-\lambda)p_2 > \lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y).\\
q &= \lambda q_1 + (1-\lambda)q_2 \geq \lambda g(x) + (1-\lambda)f(y) \geq g(\lambda x + (1-\lambda)y).\\
r &= \lambda r_1 + (1-\lambda)r_2\\
&= \lambda h(x) + (1-\lambda)h(y)\\
&= \lambda(Ax-b) + (1-\lambda)(Ay-b); \text{ ésto por definición de } h.\\
&= \lambda Ax - \lambda b + Ay - b - \lambda Ay + \lambda b\\
&= A\lambda x - A\lambda y + Ay - b\\
&= A(\lambda x - \lambda y + y) - b\\
&= A(\lambda x + (1-\lambda)y) - b\\
&= h(\lambda x + (1-\lambda)y).
\end{aligned}$$

Y como X es convexo $\lambda x + (1-\lambda)y \in X$ y tendremos $(p,q,r) \in C$. Por lo tanto C es convexo.

Ahora, si el sistema 1) no tiene solución entonces $0 = (0,0,0) \notin C$. Por ser C convexo y no vacío existe (u_0, u, v), tal que el hiperplano correspondiente separa 0 de C, es decir; se satisface que $(u_0,u,v)^T(0-(p,q,r)) \leq 0$ para cada $(p,q,r) \in \widetilde{C}$. Lo cual implica que $(u_0,u,v)^T(p,q,r) \geq 0$ para cada $(p,q,r) \in \widetilde{C}$, esto quiere decir que por teorema existe un hiperplano que soporta a $\widetilde{C}$ en $(0,0,0)$. Luego $u_0 p + u^T q + v^T r \geq 0$ para cada $(p,q,r) \in \widetilde{C}$.

Sea $\overline{x} \in X$, dado que p y q pueden hacerse arbitrariamente grandes, la anterior expresión se cumple si y solo si $u_0 \geq 0$ y $u \geq 0$.

Además, si $(p,q,r) \in C$ se tendrá $(p,q,r) = (f(x), g(x), h(x))$ luego

$$u_0 f(x) + u^T g(x) + v^T h(x) \geq 0, (u_0,u,v) \neq 0, (u_0,u) \geq 0.$$

Dado que lo anterior se cumple para cada $x \in X$. Por lo tanto el sistema 2) tiene solución.

Recíprocamente, supongamos que el sistema 2) tiene una solución (u_0,u,v) tal que $u_0 > 0$ y $u \geq 0$ y se satisface $u_0 f(x) + u^T g(x) + v^T h(x) \geq 0$ para cada $x \in X$. Sea $x \in X$

tal que $g(x) \leq 0$ y $h(x) = 0$ dado que $u \geq 0$, se concluye que $u_0 f(x) \geq 0$ pues $u^T g(x) \leq 0$ y $h(x) = 0$.

Así,

$$\begin{aligned} u_0 f(x) + u^T g(x) &\geq 0, \\ 0 \geq u^T g(x) &\geq -u_0 f(x), \\ 0 &\leq u_0 f(x), \end{aligned}$$

lo que implica $f(x) \geq 0$. Por lo tanto 1) no tiene solución. □

Teorema 12 (Dualidad fuerte)**.** Sea $X \subset \mathbb{R}^n$ un conjunto convexo no-vacío y sean $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ funciones convexas y $h : \mathbb{R}^n \to \mathbb{R}^s$ una función afín. Supóngase además que se satisface el requisito restriccional siguiente:

i) $\{x/g(x) < 0; h(x) = 0\} \neq \emptyset$.

ii) $0 \in \operatorname{int}\{h(x)/x \in X\}$.

Entonces

a) $\text{ínf}\{f(x)/x \in X; g(x) \leq 0, h(x) = 0\} = \sup\{\theta(u,v)/u \geq 0\}$.
Además si el ínfimo es finito, entonces existen $(\overline{u}, \overline{v})$ con $\overline{u} \geq 0$ tales que
$\sup\{\theta(u,v)/u \geq 0\} = \theta(\overline{u}, \overline{v})$.

b) Finalmente, si existe $\overline{x} \in X$, tal que $f(\overline{x}) = \text{ínf}\{f(x)/x \in X; g(x) \leq 0; h(x) = 0\}$ entonces $\overline{u}^T g(\overline{x}) = 0$.

Demostración. Sea $\alpha = \text{ínf}\{f(x)/x \in X; g(x) < 0, h(x) = 0\} \neq \emptyset$. Entonces, si $\alpha = -\infty$ y por corolario 3 se tiene que $\sup\{\theta(u,v)/\mu \geq 0\} = -\infty$.

Ahora, si el ínfimo es finito, ocurre $f(x) \geq \alpha$, $g(x) \leq 0$, $h(x) = 0$ y examinemos el siguiente sistema $f(x) - \alpha < 0$, $g(x) \leq 0$, $h(x) = 0$, $x \in X$.

Por la forma en el que se tiene α éste último sistema no tiene solución y por el lema anterior existe $(u_0, u, v) \neq 0$ con $(u_0, u) \geq 0$ tal que

$$u_0 [f(x) - \alpha] + u^T g(x) + v^T h(x) \geq 0, \tag{2.1}$$

para cualquier $x \in X$. El recíproco es cierto si $u_0 > 0$. Supongamos por contradicción que $u_0 \leq 0$. Como $(u_0, u) \geq 0$ se tiene que $u_0 = 0$, como se cumple que el requisito restriccional, existe $\widehat{x} \in X$, tal que $g(\widehat{x}) < 0$, $h(\widehat{x}) = 0$ y además $0 \in \operatorname{int}(h(X))$.

Entonces al substituir en (2.1) sucede que:

$$\begin{aligned} u_0 [f(\widehat{x}) - \alpha] + u^T g(\widehat{x}) + v^T h(\widehat{x}) &\geq 0, \\ u^T g(\widehat{x}) &\geq 0 \end{aligned}$$

además $u^T g(\widehat{x}) \leq 0$ es decir; $0 \geq u^T g(\widehat{x}) \geq 0$ lo que implica $u^T g(\widehat{x}) = 0$ si $u = 0$. De esto último al tener que $u_0 = 0$ y $u = 0$ nuevamente de (2.1) $v^T h(x) \geq 0$, para todo $x \in X$.

Ahora bien como $0 \in \text{int}(h(X))$, existe $\epsilon > 0$ tal que $B(0, \epsilon) \subset h(X)$. Por lo tanto para $\lambda > 0$ tal que $-\lambda v \in B(0, \epsilon)$, existe $\widetilde{x} \in X$ tal que $h(\widetilde{x}) = -\lambda v$. Luego

$$0 \leq v^T h(\widetilde{x}) = v^T(-\lambda v) = -\lambda v^T v = -\lambda \|v\|^2 \leq 0,$$

luego $-\lambda \|v\|^2 = 0$ entonces $v = 0$ ya que $\lambda > 0$ lo que nos lleva a una contradicción pues $(u_0, u, v) \neq 0$.

Ahora dividamos (2.1) entre $u_0 > 0$

$$u_0 \left[f(x) - \alpha\right] + u^T g(x) + v^T h(x) \geq 0,$$
$$\left[f(x) - \alpha\right] + \left[\frac{u}{u_0}\right]^T g(x) + \left[\frac{v}{u_0}\right]^T h(x) \geq 0.$$

y definamos: $\overline{u} = \dfrac{u}{u_0}$ y $\overline{v} = \dfrac{v}{v_0}$ con lo cual se obtiene

$$f(x) + \overline{u}^T g(x) + \overline{v}^T h(x) \geq \alpha, \tag{2.2}$$

para todo $x \in X$. Luego $\inf\limits_{x \in X} \left\{f(x) + \overline{u}^T g(x) + \overline{v}^T h(x)\right\} \geq \alpha$, esto nos muestra que

$$\theta(\overline{u}, \overline{v}) = \inf \left\{f(x) + \overline{u}^T g(x) + \overline{v}^T h(x)\right\} \geq \alpha = \inf \left\{f(x)/x \in X, g(x) < 0, h(x) = 0\right\}$$

y por el resultado débil de dualidad $\theta(\overline{u}, \overline{v}) = \alpha$ es decir, $(\overline{u}, \overline{v})$ es una solución del dual.

Finalmente, sea $\overline{x} \in X$, $g(\overline{x}) \leq 0$, $h(\overline{x}) = 0$; $f(\overline{x}) = \alpha$ donde $\overline{x}$ es una solución óptima del primal, de (2.2) se tiene

$$\overline{u}^T g(\overline{x}) \geq 0. \tag{2.3}$$

Por otro lado como $\overline{u} \geq 0$ y $g(\overline{x}) \leq 0$ implica $\overline{u}^T g(\overline{x}) \leq 0$.

Finalmente, de (2.3) tenemos que $0 \geq \overline{u}^T g(\overline{x}) \geq 0$. Por tanto $\overline{u}^T g(\overline{x}) = 0$. $\square$

Capítulo 3

Dualidad fuerte en optimización no convexa.

En éste capítulo estudiamos dualidad fuerte pero en el caso no convexo, estudiando, antes, propiedades sobre los conos las cuales necesitaremos para demostrar un teorema de dualidad fuerte, luego veremos dualidad fuerte en un caso general y el caso general con una sola restricción.

3.1. Conos, pointed y sus propiedades

En la presente sección estudiamos de manera detallada el menor cono que contiene a un conjunto, la definición de cono pointed y algunas de sus propiedades que necesitaremos, para luego estudiar algunas propiedades de los conos asintóticos.

Definición 11. Sea $A \subseteq \mathbb{R}^n$

1. Se define $\operatorname{cone}(A)$, al menor cono que contiene a A, esto es: $\operatorname{cone}(A) = \bigcup_{t \geq 0} tA$, $\overline{\operatorname{cone}}(A)$ denota al menor cono cerrado que contiene a A, $\left(\overline{\operatorname{cone}}(A) = \overline{\operatorname{cone}(A)}\right)$.

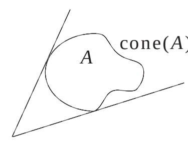
2. $\operatorname{cone}_+(A) = \bigcup_{t>0} tA$ observemos que: $\operatorname{cone}(A) = \operatorname{cone}_+(A) \bigcup \{0\}$

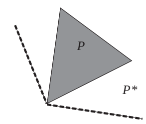
3. Entenderemos por cono polar (positivo) de P como:

$$P^* = \{p^* \in \mathbb{R}^n : \langle p^*, p \rangle \geq 0, \text{ para todo } p \in P\}$$

4. Dado un conjunto $M \subseteq \mathbb{R}^n$, Denotaremos por$M^\perp$ al subespacio ortogonal de M en donde:

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





Definición 12. Un cono $P \subset \mathbb{R}^m$ es pointed si y solo si la igualdad $y_1 + y_2 + \ldots + y_n = 0$ no se cumple para elementos de P a menos que $y_1 = y_2 = \ldots = y_n = 0$.

Es posible caracterizar a un cono con la propiedad pointed cuando éste es a su vez es convexo, es decir, si $P \subset Y$, cono convexo.

Lema 4. Sea $K \subseteq \mathbb{R}^n$ un cono convexo. Entonces K es pointed si y solo si $K \cap (-K) = \{0\}$.

Demostración. Razonando por contradicción, suponemos que existe $y \neq 0$ tal que $y \in K \cap (-K)$ esto implica que $y \in K$ y $y \in -K$ luego $y \in K$ y $-y \in K$ así $y + (-y) = 0$ y como K es pointed se tiene que $y = 0$ lo que es una contradicción. Por lo tanto $K \cap (-K) = \{0\}$.

Recíprocamente supongamos que $K \cap (-K) = \{0\}$. Existen y_i en K con $i = 1, 2, \ldots, n$ tal que $y_1 + y_2 + \ldots + y_n = 0$. Ahora, supongamos que existe $j \in \{1, 2, \ldots, n\}$ tal que $y_j \neq 0$, entonces $y_j = -\sum_{i=1_{i \neq j}}^{n} y_i \in K$ y $y_j \in -K$. Luego $y_j \in K \cap (-K)$ para todo $j \in \{1, 2, \ldots, n\}$ lo cual es una contradicción. Por tanto K es pointed. $\square$

Lema 5. Sea $A \subset \mathbb{R}^n$ no vacío. Luego $\text{cone}(\overline{\text{cone}}(A)) \subset \overline{\text{cone}}(A)$.

Demostración. Sea $a \in \text{cone}(\overline{\text{cone}}(A))$. Luego, por definición de cone, existe $t_0 \geq 0$ tal que $a \in t_0 \overline{\text{cone}}(A)$. Así $a = t_0 c$ con $c \in \overline{\text{cone}}(A)$. Entonces para $a = t_0 c$ existe una sucesión $(x_n) \in \text{cone}(A)$ tal que $x_n \to c$. Lo cual implica que $t_0 x_n \to t_0 c$. Poniendo $(z_n) = t_0(x_n)$ y $a = t_0 c$ tendremos que $z_n \to a$. Por tanto $a \in \overline{\text{cone}}(A)$ y así $\text{cone}(\overline{\text{cone}}(A)) \subset \overline{\text{cone}}(A)$ como queríamos. $\square$

Proposición 9. Sean $A, K \subseteq \mathbb{R}^n$, conjuntos no vacíos.

a) $\overline{\text{cone}}(A) = \overline{\text{cone}}(\overline{A})$

b) Si A es abierto entonces $\text{cone}_+(A)$ es abierto.

c) $\text{cone}(\text{co}(A)) = \text{co}(\text{cone}(A))$, $\text{cone}_+(\text{co}(A)) = \text{co}(\text{cone}_+(A))$.

d) $\text{cone}_+(A+K) = \text{cone}_+(A) + K$, cuando K tiene la propiedad que para todo $t > 0$; $tK \subseteq K$.

e) $\overline{A+K} = \overline{A+\overline{K}}$.

f) $K \subseteq \overline{\text{cone}}(A+K)$ cuando K es un cono.

g) $\text{cone}(A+K) \subseteq \text{cone}(A)+K \subseteq \overline{\text{cone}}(A+K)$, cuando K es un cono.] Adicionalmente si $0 \in A$, entonces $\text{cone}(A+K) = \text{cone}(A) + K$.

Demostración.

a) Como $A \subset \overline{A}$ implica que $\text{cone}(A) \subset \text{cone}(\overline{A})$ y de ésta manera $\overline{\text{cone}}(A) \subset \overline{\text{cone}}(\overline{A})$.

Recíprocamente, como $A \subset \text{cone}(A)$ implica que $\overline{A} \subset \overline{\text{cone}}(A)$ y se tiene que $\text{cone}(\overline{A}) \subset \text{cone}(\overline{\text{cone}}(A)) \subset \overline{\text{cone}}(A)$.

Por lo mostrado en el lema 4, obtenemos que $\text{cone}(\overline{A}) \subset \overline{\text{cone}}(A)$. Lo cual implica que $\overline{\text{cone}}(\overline{A}) \subset \overline{\overline{\text{cone}}}(A) = \overline{\text{cone}}(A)$. Finalmente tenemos que $\overline{\text{cone}}(\overline{A}) \subset \overline{\text{cone}}(A)$.

Y así hemos mostrado que $\overline{\text{cone}}(\overline{A}) = \overline{\text{cone}}(A)$ como queríamos.

b) Por hipótesis, como A es abierto entonces todo $a \in A$ es interior si existe $\epsilon > 0$ tal que $B(a,\epsilon) \subset A$ y además $\text{cone}_+(A) = \bigcup_{t>0} tA = \bigcup_{t>0} \{ta/a \in A\}$.

Solo nos queda por mostrar que tA es abierto. En efecto, sea $x \in tA$ entonces existe un $t_0 > 0$ tal que $x = t_0 a$ con $a \in A$. Como $a \in A$ entonces existe $\epsilon > 0$ tal que $B(a,\epsilon) \subset A$. Luego $t_0 B(a,\epsilon) \subset t_0 A$ entonces $B(t_0 a, t_0\epsilon) \subset t_0 A$. Tomando $t_0\epsilon = \epsilon'$ tendremos $B(t_0 a, \epsilon') \subset t_0 A$. Ahora si $y \in B(t_0 a, \epsilon')$ se tendrá que $d(t_0 a, y) < \epsilon'$. Luego $t_0 d(a,y) < \epsilon'$ y para algún $z \in B(t_0 a, \epsilon')$ obtendremos que $d(t_0 a, z) < \epsilon'$ y concluimos que tA es abierto.

Luego $\text{cone}_+(A) = \bigcup_{t>0} tA$ es abierto.

c) Primero mostremos que $\text{cone}(\text{co}(A)) = \text{co}(\text{cone}(A))$. Sea $a \in \text{cone}(\text{co}(A))$ y por definición existe $t_0 \geq 0$ tal que $a \in t_0\, \text{co}(A)$. Por definición de $\text{co}(A)$ el elemento a puede escribirse como una combinación lineal convexa de elementos de A, es decir;

$$a = t_0 \sum_{i=1}^{m} \lambda_i a_i = \sum_{i=1}^{m} \lambda_i t_0 a_i$$

tal que $\sum_{i=1}^{m} \lambda_i = 1$ para $\lambda_i \geq 0$, $a_i \in A$. Observemos que $t_0 a_i$ está en $t_0 A$ y como existe $t_0 \geq 0$ decimos que $t_0 a_i \in \bigcup_{t \geq 0} tA$. Por lo anterior, a está en la envoltura convexa de $\bigcup_{t\geq 0} tA$ es decir, $a \in \text{co}(\bigcup_{t\geq 0} tA)$ y concluimos diciendo que $a \in \text{co}(\text{cone}(A))$.

Recíprocamente, sea $a \in \text{co}(\text{cone}(A))$. Entonces a puede escribirse como una combinación lineal convexa de elementos del $\text{cone}(A)$, es decir;

$$a = \sum_{i=1}^{m} \alpha_i a_i,$$

tal que $\sum_{i=1}^{m} = 1$, $\alpha_i \geq 0$, $a_i \in \text{cone}(A)$. Como a_i está en el cone(A), entonces existe $t_i \geq 0$ de tal manera que $a = \sum_{i=1}^{m} \alpha_i t_i b_i$, donde $\sum_{i=1}^{m} \alpha_i = 1$, $\alpha_i \geq 0$, $b_i \in A$. Así,

$$a = \sum_{i=1}^{m} \alpha_i t_i b_i = \sum_{i=1}^{m} \alpha_i t_i \sum_{i=1}^{m} \frac{\alpha_i t_i}{\sum_{i=1}^{m} \alpha_i t_i} b_i,$$

donde $\sum_{i=1}^{m} \frac{\alpha_i t_i}{\sum_{i=1}^{m} \alpha_i t_i} = 1$, $\frac{\alpha_i t_i}{\sum_{i=1}^{m} \alpha_i t_i} \geq 0$, $b_i \in A$. Notemos que $\sum_{i=1}^{m} \frac{\alpha_i t_i}{\sum_{i=1}^{m} \alpha_i t_i} b_i \in \text{co}(A)$, por tanto $a \in \text{cone}(\text{co}(A))$.

La demostración de $\text{cone}_+(\text{co}(A)) = \text{co}(\text{cone}_+(A))$ es similar a la anterior.

d) Comencemos mostrando que $\text{cone}_+ K = K$ si $tK \subseteq K$. Sea $a \in \text{cone}_+ K$ entonces $a \in \bigcup_{t>0} tK$. Luego existe $t_0 > 0$ tal que $a \in t_0 K$. Como $t_0 K \subseteq K$, se tiene que $a \in K$ y de ésta manera $\text{cone}_+ K \subseteq K$. Recíprocamente, sea $a \in K$. Entonces existe $t_0 > 0$ tal que $a = t_0 \frac{1}{t_0} a \in \bigcup_{t>0} tK$ y por definición $a \in \text{cone}_+ K$, así $K \subseteq \text{cone}_+ K$. Por tanto $\text{cone}_+(K) = K$.

Ahora mostremos $\text{cone}_+(A + K) = \text{cone}_+(A) + K$. Sea $x \in \text{cone}_+(A + K)$ entonces existe $t_0 > 0$ tal que $x = t_0(a + k)$ donde, $a \in A$ y $k \in K$. Luego $x = t_0 a + t_0 k$ y por hipótesis tenemos que $t_0 k \in t_0 K \subseteq K$. Así, $x \in \text{cone}_+(A) + K$. Recíprocamente, sea $x \in \text{cone}_+(A) + K$. Entonces existe $t_0 > 0$ tal que $x = t_0 b + k$, donde $t_0 b \in \text{cone}(A)$ y $k \in K$. Notemos que $x = t_0(b + \frac{1}{t_0} k)$ donde $\frac{1}{t_0} k \in K$, por tanto $x \in \text{cone}_+(A + K)$.

e) Primero empecemos mostrando que $\overline{A + K} \subset \overline{A + \overline{K}}$. Si $A \subset \overline{K}$ entonces $A + K \subset A + \overline{K}$. Luego $\overline{A + K} \subset \overline{A + \overline{K}}$.

Ahora veamos que $\overline{A + K} \supset \overline{A + \overline{K}}$. Sea $a + k \in A + \overline{K}$ con $a \in A$ y $k \in \overline{K}$. Entonces, existe $x_n \in K$ tal que $x_n \to k$. Luego $a + x_n \to a + k$ cuando $n \to +\infty$, lo que muestra que $a + k \in \overline{A + K}$. Así afirmamos que $A + \overline{K} \subset \overline{A + K}$ y $\overline{A + \overline{K}} \subset \overline{\overline{A + K}} = \overline{A + K}$. Por tanto $\overline{A + K} = \overline{A + \overline{K}}$.

f) Sea $x \in K$ y para cada elemento $a \in A$ la sucesión $\frac{a}{n} + x \to x$ cuando $n \to +\infty$. Luego $\frac{1}{n}(a + nx) \to x$ cuando $n \to +\infty$. Como K es un cono entonces $nx \in k$ y $\frac{1}{n}(a + nx) \in \text{cone}(A + K)$. Luego tendremos que $x \in \overline{\text{cone}}(A + K)$ y así $K \subseteq \overline{\text{cone}}(A + K)$.

g) Por e) obtenemos la primera inclusión.

Tomando en cuenta $t_0 \geq 0$. Sea $x \in \text{cone}(A) + K$ tal que $x = a + k$ donde $a \in \text{cone}(A)$ y $k \in K$. De ésta manera, si $a \in \text{cone}(A)$ implica que $a = t_0 b$ con $b \in A$, $t_0 \geq 0$ y $k = t_1 c$ con $t_1 \geq 0$ y $c \in K$. Entonces podemos escribir $x = t_0 b + t_1 c$. Ahora para $x = t_0 b + t_1 c$ analicemos los siguientes casos:

i) si $t_0 = t_1 = 0$ entonces $x = 0 \in \overline{\text{cone}}(A + K)$.

ii) Si $t_0 > 0$ y $t_1 = 0$ entonces $x = t_0 b + 0 = t_0(b + 0) \in \overline{\text{cone}}(A + K)$.

iii) Si $t_0 = 0$ y $t_1 > 0$, entonces existe $x_n = \dfrac{a}{n} + t_1 c = \dfrac{1}{n}(a + n t_1 c)$ tal que $x_n \to x$ cuando $n \to +\infty$ por lo tanto $x \in \overline{\text{cone}}(A + K)$.

iv) Si $t_1, t_2 > 0$ entonces $x \in \text{cone}_+ A + K$ y de acuerdo con e) obtenemos que

$$\text{cone}_+ A + K = \text{cone}_+(A + K) \subseteq \overline{\text{cone}}(A + K)$$

.

□

3.2. Conos Asintóticos y sus propiedades.

Definición 13. Dado un conjunto $C \subseteq \mathbb{R}^n$, se define al cono asintótico de C, denotado como C^∞ al conjunto

$$C^\infty = \{v : \text{ existe } t_k \downarrow 0, \text{ existe } (x_k) \text{ en } C, t_k x_k \to v\}.$$

Ejemplo 4. El cono asintótico de una recta que pasa por el origen es la misma recta.

Sea $R = \{(x, y) \in \mathbb{R}^2 : x = y\}$ y la sucesión $a\,(n, n)$ en R para todo a en $\mathbb{R}$ y $t_n \to 0$ tal que $\{t_n\} = \left\{\frac{1}{n}\right\}$ luego $\lim a\,(n, n)\, t_n = \lim a\,(n, n)\,\dfrac{1}{n} = (a, a)$ así (a, a) en C^∞. Ahora mostremos que para cualquier (x, y) en $\mathbb{R}^2$ con $x \neq y$ entonces (x, y)no está en R^∞. En efecto; si x está en R^∞ entonces existe algún (z_k) en R y $t_k \to 0$ tal que $t_k z_k \to (x, y)$ como (z_k) en R implica que $z_k = (x_k, x_k)$. Luego $t_k(x_k, x_k) \to (x, y)$ ó $t_k x_k \to x$, $t_k x_k \to y$ entonces $x = y$ lo que es una contradicción, pues $x \neq y$. Por lo tanto (x, y) no está en R^∞.

Ejemplo 5. El cono asintótico de la parábola es el eje vertical no negativo.

Sean el conjunto $C = \{(x, y) \in \mathbb{R}^2 : x = y^2\}$, la sucesión $\{x_n\} = \{(an, a^2 n^2)\} \subset C$ y $\{t_n\} = \left\{\frac{1}{n^2}\right\}$ tal que $t_n \downarrow 0$. Luego $\lim \frac{1}{n^2}(an, a^2 n^2) = (0, a^2)$, donde $a \in \mathbb{R}$ y $n \to \infty$.

Ahora mostremos que para cualquier (a, b) en $\mathbb{R}^2$ implica que (a, b) no está en C^∞. En efecto, sea (a, b) en C^∞ tal que $a \neq 0$ y $a > 0$. Entonces existen $t_n \downarrow 0$, $\{x_n\} = \{(z_n, z_n^2)\} \subset C$ tal que $t_n x_n \to (a, b)$; es decir $t_n(z_n, z_n^2) \to (a, b)$. De lo anterior decimos que $t_n z_n \to a$ y como $t_n \downarrow 0$ implica que $z_n \to +\infty$, luego $t_n z_n^2 = (t_n z_n) z_n \to +\infty$ lo que es una contradicción. Ahora, tomando el caso $a < 0$ en la anterior demostración, tendremos que si $t_n z_n \to a$ implicará que $t_n z_n^2 \to -\infty$ que nos da otra contradicción. Finalmente observemos que cualquier elemento que tomemos en la segunda coordenada obtendremos su cuadrado el cual es positivo.

Ejemplo 6. El cono asintótico de un conjunto no convexo no necesita ser convexo.

Sea $C = \left\{(x,y) \in \mathbb{R}^2 : y = \frac{1}{x}, x > 0\right\}$ luego C es no convexo y C^∞ es no convexo. En efecto; el cono asintótico de C es la unión de los ejes no negativos x e y. Sean $\{x_n\} = \left\{(\frac{1}{an}, an)\right\} \subset C$ y $t_n = \frac{1}{n} \to 0$ con $n \to \infty$, luego

$$\lim x_n t_n = \lim \left(an, \frac{1}{an}\right)\frac{1}{n} = \lim \left(a, \frac{1}{n^2}\right) = (a,0)$$

para todo $a \geq 0$; es decir, tenemos el eje x.

Por otro lado, sean $\{x_n\} = \left\{(\frac{1}{an}, an)\right\} \subset C$ y $\frac{1}{n} \to 0$ con $n \to \infty$, luego

$$\lim x_n t_n = \lim \left(\frac{1}{an}, an\right)\frac{1}{n} = (0,a)$$

para todo $a \geq 0$, es decir tenemos el eje y.

Ahora mostremos que para cualquier punto (x,y) en $\mathbb{R}^2$ entonces (x,y) no está en C^∞. En efecto, supongamos que (x,y) en C^∞ implica que existen $\{z_k\} \subset C$ y $t_k \downarrow 0$ tal que $t_k z_k \to (x,y)$, $x > 0$, $y > 0$. Como $\{z_k\} \subset C$ entonces $\{z_k\} = \left\{(x_k, \frac{1}{x_k})\right\}$. Luego $t_k z_k = t_k \left(x_k, \frac{1}{x_k}\right) \to (x,y)$ ó $t_k x_k \to x$, $t_k \left(\frac{1}{x_k}\right) \to y$. Por lo tanto $t_k x_k t_k \frac{1}{x_k} \to xy > 0$ y $(t_k)^2 \to xy > 0$ lo cual es una contradicción, pues $(t_k)^2 \to 0$.

Proposición 10. Sea $C \subseteq \mathbb{R}^n$.

a) C^∞ es un cono cerrado.

b) C es un cono cerrado si y solo si $C^\infty = C$.

c) $C^\infty = \{0\}$ si y solo si C es acotado.

d) $C_1 \subseteq C_2 \Rightarrow (C_1)^\infty \subseteq (C_2)^\infty$.

e) $C^\infty = (C + x_0)^\infty$ para todo x_0 de $\mathbb{R}^n$.

f) $C^\infty = \left(\overline{C}\right)^\infty$.

g) Si C es no vacío, cerrado y convexo, x_0 en C entonces

$$C^\infty = \left\{u \text{ en } \mathbb{R}^n : x_0 + tu \text{ en } C, \text{ para todo } t > 0\right\}.$$

h) Sea $\{C_i\}$, $i = 1,2,3,\ldots,m$ una familia finita de conjuntos no vacíos en $\mathbb{R}^n$, entonces

$$\left(\bigcup C_i\right)^\infty = \bigcup (C_i)^\infty.$$

i) Sea $\{C_i\}$, $i \in I$, una familia arbitraria de conjuntos no vacíos en $\mathbb{R}^n$, entonces

$$\left(\bigcap_{i\in I} C_i\right)^{\infty} \subseteq \bigcap_{i\in I} (C_i)^{\infty}.$$

Adicionalmente, si cada C_i es cerrado y convexo y $\bigcap_{i\in I} C_i \neq \emptyset$ entonces

$$\left(\bigcap_{i\in I} C_i\right)^{\infty} = \bigcap_{i\in I} (C_i)^{\infty}.$$

Demostración.

a) C^∞ es cono. En efecto, sea $x \in tC^\infty$, entonces existe $t_0 \geq 0$ tal que $x = t_0 p$ con $p \in C^\infty$ y como $p \in C^\infty$ existen $(y_k) \in C$ y $t_k \to 0$ tal que $x_k t_k \to p$ es decir $\lim x_k t_k = p$ con $k \in \mathbb{N}$ entonces $x = t_0 p = t_0 \lim x_k t_k = \lim x_k t_0 t_k$. Además notemos que $t_0 t_k \to 0$ lo cual implica que $x \in C^\infty$ y asi hemos demostrado que C^∞ es un cono.

C^∞ es cerrado. En efecto, sea $x_n \in C^\infty$ con $x_n \to x$ y para cada m existe una sucesión $x_{n,m}$ con $\lambda_{n,m} x_{n,m} \to x_n$, $\lambda_{n,m} \to 0$ y con $m \to \infty$ y cada $\lambda_{n,m} > 0$. Entonces para cada k existe un N_k tq para todo $n \geq N_k$, $\|x_n - x\| < \frac{1}{k}$ y para cada k existe un M_k tal que para todo $m \geq M_k$,$\|\lambda_{N_k,m} x_{N_k,m} - x_{N_k}\| < \frac{1}{k}$ tal que para todo $m \geq L_k$ con $|\lambda_{N_k,m}| < \dfrac{1}{k}$. Sea el conjunto $P_k =$máx$\{M_k, L_k\}$ y $y_k = x_{N_k,P_k}$, $\lambda = \lambda_{N_k,P_k}$, entonces cada $\lambda_k > 0$, $\lambda_k \to 0$ de tal manera que:

$$\begin{aligned}
\|\lambda_k y_k - x\| &\leq \|\lambda_k y_k - x_{N_k}\| + \|x_{N_k} - x\| \\
&\leq \|\lambda_{N_k,P_k} x_{N_k,P_k} - x_{N_k}\| + \|x_{N_k} - x\| \\
&< \frac{1}{k} + \frac{1}{k} \\
&= \frac{2}{k},
\end{aligned}$$

luego $x \in C^\infty$ por lo tanto C^∞ es cerrado.

b) Sea C cono cerrado. Primero mostremos que $C^\infty \subseteq C$; sea $x \in C^\infty$ entonces existen $t_k \downarrow 0$ y $(x_k) \in C$ tal que $t_k x_k \to x$ y como C es cerrado $(t_k x_k) \in C$, por lo tanto $x \in C$. Ahora mostremos que $C^\infty \supseteq C$; sea $c \in C$ y como C es un cono cerrado existe una sucesión $x_n \to c$, notemos que $c \to c$ ésto implica que $\frac{1}{n} nc \to c$ cuando $n \to \infty$, por lo tanto $c \in C^\infty$.

El recíproco es inmediato.

c) Sea la hipótesis $C^\infty = \{0\}$. Supongamos por contradicción que C es no acotado, entonces existe una sucesión (x_k) en C tal que $\|x_k\| \to \infty$ luego $t_k = \left\|\dfrac{1}{x_k}\right\| \to 0$ además

$$\|t_k x_k\| = |t_k| \, \|x_k\| = \left\|\frac{1}{x_k}\right\| \|x_k\| = 1$$

de donde $(t_k x_k) \in S$ (S es la esfera unitaria). Como S es compacto, existe una subsucesión de $(t_k x_k)$ que converge en S. Luego existe $x \in C^\infty$ y $x \in S$ con $x \neq 0$ lo que es una contradicción pues $C^\infty = \{0\}$. Por lo tanto C es no acotado.

Recíprocamente, sea la hipótesis C acotado. Mostremos que $C^\infty \subset \{0\}$, entonces tomemos un $x \in C^\infty$, luego existen $t_k \downarrow 0$ y $(x_k) \in C$ tal que $t_k x_k \to x$, de tal manera que $t_k x_k \to 0$ pues $(x_k) \in C$ y C es acotado, así

$$\|t_k x_k\| = |t_k| \, \|x_k\| \leq |t_k| \, M \to 0$$

luego $t_k x_k \to 0$, de ésta manera $x = 0 \in \{0\}$ por lo tanto $x \in \{0\}$.

Para la inclusión $C^\infty \supset \{0\}$; existe $x \in C$ tal que $x_k = x$ con $t_k x_k \to 0$ para cualquier $t_k \to 0$ entonces $0 \in C^\infty$.

d) Sea $x \in (C_1)^\infty$ entonces existen $(x_k) \in C_1$ y $t_k \to 0$ tal que $t_k x_k \to x$ luego $(x_k) \in C_2$ y $t_k \to 0$ tal que $t_k x_k \to x$ por lo tanto $x \in C_2^\infty$.

e) Sea $y \in (C + x_0)^\infty$ luego existe $t_k \downarrow 0$, $(x_k + x_0) \in C + x_0$, $(x_k) \in C$ tal que $t_k (x_k + x_0) \to y$ implica $t_k x_k + t_k x_0 \to y$ así, $t_k x_k \to y$ pues $t_n x_k \to 0$ con $(x_k) \in C$, por lo tanto $y \in C^\infty$. Y así hemos probado que $(C + x_0)^\infty \subseteq C^\infty$.

Por otro lado sabemos que $C \subseteq C + x_0$ entonces $C^\infty \subseteq (C + x_0)^\infty$.

f) Como $C \subseteq \overline{C}$ entonces $C^\infty \subseteq \left(\overline{C}\right)^\infty$.

Recíprocamente, sea $x \in (\overline{C})^\infty$ por definición existen sucesiones $x_n \in \overline{C}$ y $t_n \downarrow 0$ tal que $x_n t_n \to x$ donde $n \in \mathbb{N}$. Como $(x_n) \in \overline{C}$ existen sucesiones $\{y_{n,m}\} \subset C$ tal que $y_{n,m} \to x_n$. Ahora de la hipótesis tenemos: para cada k existe un M_k tal que para todo $n \geq M_k$ se tiene que $\|x_n t_n - x\| < \frac{1}{k^2}$, para cada k existe un N_k tal que para todo $m \geq N_k$ se tiene que $\|y_{N_k,m} - x_{M_k}\| < \frac{1}{k}$ y también para todo $n \geq M_k$ se cumple que $\|t_n\| < \frac{1}{k}$. Sea $P_k = \max\{N_k, L_k\}$ y $p_k = y_{M_k,N_k}$, entonces $\|t_n p_k - x\| \leq \|t_n p_k - x_{M_k} t_n\| + \|x_{M_k} t_n - x\| = \|t_n\| \, \|p_k - x_{M_k}\| + \|x_{M_k} t_n - x\| < \frac{2}{k^2}$ y así $x \in C^\infty$.

g) Sean $z \in C^\infty$, $x \in C$ y $t_0 \geq 0$ por demostrar que $x + t_0 z \in C$. En efecto, por la definición de C^∞ existen $(z_k) \in C$ y $t_k \to 0$ tal que $t_k z_k \to z$ con $t_k \geq 0$. Luego para algún k suficientemente grande y $t \in [0,1]$ se tiene $0 \leq t t_k \leq 1$ y como C es convexo se tendrá $(1 - t t_k)\, x + t t_k z_k \in C$, entonces para $k \to \infty$ el

$$\begin{aligned}
\lim \left[(1 - t t_k)\, x + t t_k z_k\right] &= \lim (1 - t t_k)\, x + \lim t t_k z_k \\
&= \lim x - t x \lim t_k z_k \\
&= x + t z
\end{aligned}$$

y como C es cerrado implica $x + t z \in C$ como queríamos.

Recíprocamente, sea $z \in \{u \in \mathbb{R}^n; x_0 + t u \in C\}$ entonces $\frac{1}{n}(x_0 + n z) \to z$, así $z \in C^\infty$.

h) Para $i = 1, 2, \ldots, m$ sea x en $(\bigcup C_i)^\infty$, entonces existen $\{x_k\} \subset \bigcup C_i$ y $t_k \downarrow 0$ tal que $x_k t_k \to x$. Luego, algún C_i tal que $i = 1, 2, \ldots, m$, contiene infinitos índices x_k en C_i. Llamemos a tal conjunto de índices $\mathbb{N}^1$ notemos que $\{x_k\} \subset C_i$ con k en $\mathbb{N}^1$. Como $\{t_k\} \downarrow 0$ y $x_k t_k \to x$ para k en $\mathbb{N}^1$ implica que x está en $(C_i)^\infty$.

Por otro lado, sea $x \in \bigcup (C_i)^\infty$ para $i = 1, 2, \ldots, m$, luego existe $i = 1, 2, \ldots, m$ tal que $x \in (C_i)^\infty$ entonces existen sucesiones $(x_{k,i})$ en C_i y $t_k \to 0$ tal que $x_{k,i} t_k \to x$ y por definición de cono asintótico obtenemos que $x \in (\bigcup C_i)^\infty$.

i) Primero probemos que $\left(\bigcap_{i\in I} C_i\right)^\infty \subseteq \bigcap_{i\in I} (C_i)^\infty$. En efecto, sea $a \in \left(\bigcap_{i\in I} C_i\right)^\infty$ entonces existen $(x_k) \in \bigcap_{i\in I} C_i$ y $t_k \to 0$ tal que $t_k x_k \to a$. Entonces para todo $i \in I$ y $(x_k) \in C_i$ concluimos que $a \in (C_i)^\infty$. Por lo tanto para todo $i \in I$, $a \in \bigcap_{i\in I} (C_i)^\infty$.

Adicionalmente, si cada C_i es cerrado y convexo y $\bigcap_{i\in I} C_i \neq \emptyset$ entonces solo nos queda demostrar $\bigcap_{i\in I} (C_i)^\infty \subset \left(\bigcap_{i\in I} C_i\right)^\infty$ y obtendremos la igualdad de la proposición. Sean $a \in \bigcap_{i\in I} (C_i)^\infty$, $x_0 \in \bigcap_{i\in I} C_i$ y $t_0 > 0$. Entonces, para todo i en I, a está en $(C_i)^\infty$ y x_0 en C_i. Como cada C_i es cerrado y convexo, por la demostración en g) tendremos que $x_0 + t_0 a$ está en C_i para todo i en I. Luego $x_0 + t_0 a$ se encuentra en $\bigcap_{i\in I} C_i$. Así, para $t_0 \to +\infty$ tenemos que $\frac{1}{t_0}(x_0 + t_0 a) \to a$, por lo tanto a está en $\left(\bigcap_{i\in I} C_i\right)^\infty$ como queríamos. $\square$

Observación 3. En i) no se cumple la igualdad directamente.

En efecto; sean $E_1 = \{0, 2, 4, ...\}$ y $E_2 = \{1, 3, 5, ...\}$ implica que $E_1 \bigcap E_2 = \emptyset$ y $(E_1 \bigcap E_2)^\infty = \emptyset$ pero $(E_1)^\infty = \mathbb{R}_+$ y $(E_2)^\infty = \mathbb{R}_+$.

3.3. Dualidad Fuerte: El caso general

En ésta sección estudiaremos Dualidad Fuerte para un caso general. Sea $C \subseteq \mathbb{R}^n$ no vacío, $f : C \to \mathbb{R}$. Consideremos el siguiente problema de minimización:

$$\mu = \inf_{x\in K} f(x). \tag{P}$$

donde $g : C \to \mathbb{R}^m$, $P \subseteq \mathbb{R}^m$ cono convexo, $K = \{x \in C, g(x) \in -P\}$ conjunto de restricciones, $P^* = \{y \in \mathbb{R}^m / \langle y, p\rangle \geq 0, \text{ para todo } p \in P\}$.

Observación 4. Como el interior topológico de P puede ser vacío, en éste caso el conjunto de restricciones es descrito por desigualdades e igualdades.

Sea el lagrangiano: $L(\gamma^*, \lambda^*, x) = \gamma^* f(x) + \langle \lambda^*, g(x)\rangle$, donde $\gamma^* \geq 0$, $\lambda^* \in P^*$, $x \in \mathbb{R}^n$.

Si $x \in K$, $\lambda^* \in P^*$, $\gamma^* \geq 0$, entonces: $-g(x) \in P$, luego $\langle \lambda^*, -g(x)\rangle \geq 0$ o equivalentemente $\langle \lambda^*, -g(x)\rangle \leq 0$; es decir, $\langle \lambda^*, g(x)\rangle \leq 0$, $x \in K$, $\lambda^* \in P^*$.

Luego de lo anterior decimos que $L(\gamma^*, \lambda^*, x) \leq \gamma^* f(x)$. Entonces

$$\inf_{x\in C} L(\gamma^*, \lambda^*, x) \leq \inf_{x\in K} L(\gamma^*, \lambda^*, x) \leq \inf_{x\in K} \gamma^* f(x) = \gamma^* \inf_{x\in K} f(x),$$

por lo tanto

$$\inf_{x\in C} L\left(\gamma^*, \lambda^*, x\right) \leq \gamma^* \inf_{x\in K} f(x),$$

para todo $\gamma^* \geq 0$ y para todo $\lambda^* \in P^*$.

La propiedad de Dualidad Fuerte asociado al problema (P), se cumple cuando además, se tiene la desigualdad contraria, es decir, cuando existen $\lambda_0^* \in P^*$ y $\gamma^* > 0$ de modo que

$$\gamma^* \inf_{x\in K} f(x) = \inf_{x\in C} L\left(\gamma^*, \lambda_0^*, x\right). \tag{3.1}$$

Lema 6. Si $A \subset \mathbb{R}^n$ es convexo, entonces $\operatorname{cone}(A)$ es convexo.

Demostración. Sean x, y en $\operatorname{cone}(A)$ y λ en $[0,1]$. Luego, para $t_0, t_1 \geq 0$ y x_1, y_1 en A, se tendrá que

$$\begin{aligned}\lambda x + (1-\lambda)y &= \lambda t_0 x_1 + (1-\lambda)t_1 y_1\\ &= (\lambda t_0 + (1-\lambda)t_1)\left[\lambda\frac{t_0}{\lambda t_0 + (1-\lambda)t_1}x_1 + (1-\lambda)\frac{t_1}{\lambda t_0 + (1-\lambda)t_1}y_1\right]\end{aligned}$$

se encuentra en el $\operatorname{cone}(A)$, pues $\lambda\frac{t_0}{\lambda t_0+(1-\lambda)t_1} + (1-\lambda)\frac{t_1}{\lambda t_0+(1-\lambda)t_1} = 1$ y $\lambda t_0 + (1-\lambda)t_1 \geq 0$. Por lo tanto $\operatorname{cone}(A)$ es convexo. □

Teorema 13. Consideremos el problema (P), en donde

$$F(C) = \{(f(x), g(x))\,; x \in C\}$$

suponiendo que $\mu \in \mathbb{R}$ y $\operatorname{int}\left(\operatorname{co}\left(F(C)\right) - (\mu, 0) + \mathbb{R}_+ \times P\right) \neq \emptyset$. Las afirmaciones siguientes son equivalentes:

a) Existe multiplicador de lagrange

$$(\gamma_0^*, \lambda_0^*) \in \mathbb{R}_+ \times P^*, (\gamma_0^*, \lambda_0^*) \neq (0,0)$$

tal que

$$\gamma_0^* \inf_{x\in K} f(x) = \inf_{x\in C} L\left(\gamma_0^*, \lambda_0^*, x\right).$$

b) $\operatorname{cone}(\operatorname{int}(\operatorname{co}(F(C)) - (\mu,0) + (\mathbb{R}_+ \times P)))$ es pointed.

c) $(0,0) \notin \operatorname{int}(\operatorname{co}(F(C) - (\mu, 0) + (\mathbb{R}_+ \times P)))$.

d) $\operatorname{cone}(\operatorname{co}(F(C) - (\mu,0) + \operatorname{int}\left(\mathbb{R}_+ \times P\right)))$ es pointed cuando $\operatorname{int} P \neq \emptyset$.

Demostración. Veamos a) implica b). Como P^* es el cono polar positivo de P y por hipótesis se cumple $\gamma_0^* \inf_{x\in K} f(x) = \inf_{x\in C} L(\gamma_0^*, \lambda_0^*, x)$ para $(\gamma_0^*, \lambda_0^*) \in \mathbb{R}_+ \times P^*$,$(\gamma_0^*, \lambda_0^*) \neq (0,0)$. Entonces

$$\gamma_0^* f(x) + \langle \lambda_0^*, g(x)\rangle \geq \gamma_0^* \inf_{x\in C} f(x)$$

y para $\mu = \inf_{x\in C} f(x)$ podemos reescribir la anterior desigualdad de la siguiente manera:

$$\langle (\gamma_0^*, \lambda_0^*), (f(x) - \mu, g(x))\rangle \geq 0;$$

para todo $x \in C$.

Sea $A = F(C) - (\mu, 0)$, dado que el cone(int(co$(A) + (\mathbb{R}_+ \times P)$)) es convexo queda por demostrar que cualquiera sea $x, -x \in$ cone(int(co$(A)+(\mathbb{R}_+ \times P)$)) se tendrá que $x = 0$. En efecto, supongamos que $x \neq 0$ entonces podemos escribir: $x = t_1\epsilon_1$ y $-x = t_2\epsilon_2$; $t_1t_2 > 0$ y $\epsilon_i \in$ int((co$(A) + (\mathbb{R}_+ \times P)$)) para $i = 1, 2$. Entonces si $\epsilon = (f(x) - \mu, g(x)) + (r, p)$ donde $(f(x) - \mu, g(x)) \in A$ y $r \in \mathbb{R}_+$, $p \in P$. Luego, se tendrá para $(\gamma_0^*, \lambda_0^*) \neq (0,0)$ en $\mathbb{R}_+ \times P$ que

$$\langle (\gamma_0^*, \lambda_0^*), \epsilon\rangle = \langle (\gamma_0^*, \lambda_0^*), (f(x) - \mu + r, g(x) + p)\rangle \geq 0$$

Ahora si $\epsilon \in$ co$(A)+(\mathbb{R}_+ \times P$ entonces $\epsilon = \sum \alpha_i\epsilon_i + (r, p)$ donde $\epsilon_i \in [0, 1]$, $\epsilon_i \in A$,$\sum \alpha_i = 1$, luego $\epsilon = \sum \alpha_i(\epsilon_i + (r, p))$ donde $\epsilon_i + (r, p) \in A + (\mathbb{R}_+ \times P)$, entonces

$$\langle (\gamma_0^*, \lambda_0^*), \epsilon\rangle = \sum \alpha_i \langle (\gamma_0^*, \lambda_0^*), \epsilon_i + (r, p)\rangle \geq 0.$$

Por otro lado existe $\delta > 0$ tal que $\epsilon_i+\lambda \in$ co$(A)+(\mathbb{R}_+\times P)$, $0 < \lambda < \delta$, $\forall y \in \mathbb{R}\times\mathbb{R}^n$,$\|y\| = 1$ (y vector unitario)(δ depende de y), luego se tiene que $0 \leq \langle (\gamma_0^*, \lambda_0^*), \epsilon_i + \lambda y\rangle$. Como $i = 1, 2$ y poniendo $\beta^* = (\gamma_0^*, \lambda_0^*)$ tendremos $\langle \beta^*, \epsilon_1 + \lambda y\rangle \geq 0$ y $\langle \beta^*, \epsilon_2 + \lambda y\rangle \geq 0$ multiplicando por t_1 y t_2 respectivamente se obtiene:

$$\langle \beta^*, t_1\epsilon_1 + \lambda y t_1\rangle \geq 0, \tag{3.2}$$

$$\langle \beta^*, t_2\epsilon_2 + \lambda y t_2\rangle \geq 0. \tag{3.3}$$

Sumando (3.2) y (3.3) tenemos que $\langle \beta^*, t_1\epsilon_1 + \lambda y t_1 + t_2\epsilon_2 + \lambda y t_2\rangle \geq 0$, pues $x - x = t_1\epsilon_1 + t_2\epsilon_2 = 0$. Así obtenemos que $\langle \beta^*, (t_1 + t_2)\lambda y\rangle \geq 0$. Luego

$$\langle \beta^*, y\rangle \geq 0 \tag{3.4}$$

para todo $y \in \mathbb{R} \times \mathbb{R}^m$. Y si $y = -\beta^*$ se tendrá $\langle \beta^*, -\beta^*\rangle \geq 0$, $\langle \beta^*, \beta^*\rangle \leq 0$ y que

$$\|\beta^*\| \leq 0. \tag{3.5}$$

De (3.4) y (3.5) concluimos que $\beta^* = 0$; es decir; $(\gamma_0^*, \lambda_0^*) = 0$ lo que es una contradicción, pues $(\gamma_0^*, \lambda_0^*) \neq (0, 0)$.

Ahora veamos b) implica c). Asumiendo la hipótesis; supongamos por contradicción que $(0,0) \in \operatorname{int}(\operatorname{co}(F(C)) - (\mu, 0) + (\mathbb{R}_+ \times P))$ entonces cone $[\operatorname{int}(\operatorname{co}(F(C)) - (\mu, 0) + \mathbb{R}_+ \times P)] = \mathbb{R} \times \mathbb{R}^m$ que es una contradicción, pues cone $[\operatorname{int}(\operatorname{co}(F(C)) - (\mu, 0) + (\mathbb{R}_+ \times P)]$ es pointed.

Probemos la equivalencia entre b) y d). Primero, dado un cono convexo Q con interior no vacío y un conjunto K, se cumple por propiedades que: $K + \operatorname{int} Q = \operatorname{int}(K + Q)$, $\operatorname{cone}(\operatorname{co}(K)) = \operatorname{co}(\operatorname{cone}(K))$, $\operatorname{co}(K + Q) = \operatorname{co}(K) + Q$. Ahora notemos que para P cono convexo con interior no vacío y A un conjunto cualquiera se tendrá:

$$\operatorname{cone}(\operatorname{int}(\operatorname{co}(A) + P)) = \operatorname{cone}(\operatorname{co}(A) + \operatorname{int} P),$$

donde $\operatorname{int} P \neq \emptyset$. Tomando en cuenta los anteriores hechos, sea $A = F(C) - (\mu, 0)$; luego

$$\begin{aligned}
\operatorname{cone}(\operatorname{int}(\operatorname{co}(F(C)) - (\mu, 0) + (\mathbb{R}_+ \times P))) &= \operatorname{cone}(\operatorname{int}(\operatorname{co}(F(C)) - (\mu, 0) + (\mathbb{R}_+ \times P))) \\
&= \operatorname{cone}(\operatorname{int}(\operatorname{co}(A)) + (\mathbb{R}_+ \times P)) \\
&= \operatorname{cone}(\operatorname{co}(A) + \operatorname{int}(\mathbb{R}_+ \times P))
\end{aligned}$$

si $\operatorname{int} P \neq 0$.

Finalmente mostramos c) implica a). Por separación entre un convexo y un punto tendremos que $\langle (\gamma_0, \lambda_0), (f(x) - \mu, g(x)) \rangle \geq 0$, esto para todo x en C. Como se cumple para todo x en C y por la definición de P^* (cono polar positivo) implica que existe $(\gamma_0^*, \lambda_0^*)$ en $\mathbb{R}_+ \times P^* \setminus \{(0,0)\}$. Así, $\langle (\gamma_0^*, \lambda_0^*), (f(x) - \mu, g(x)) \rangle \geq 0$. Luego, para todo x en C tendremos que $\gamma_0^* f(x) + \langle \lambda_0^*, g(x) \rangle \geq \gamma_0^* \mu$ lo que a su vez implica que $\inf_{x \in C} L(\gamma_0^*, \lambda_0^*, x) \geq \gamma_0^* \mu$. Y por definición de K tendremos que $\inf_{x \in C} L(\gamma_0^*, \lambda_0^*, x) = \gamma_0^* \inf_{x \in K} f(x)$. □

Corolario 6. Consideremos el problema (P). Suponiendo que $\mu \in \mathbb{R}$,

$$\operatorname{int}(\operatorname{co}(F(C) - (\mu, 0) + (\mathbb{R}_+ \times P))) \neq \emptyset$$

y se cumple la condición de Slater Generalizado:

$$\overline{\operatorname{cone}}\,(g(C) + P) = \mathbb{R}^n.$$

Entonces son equivalentes:

a) Existe multiplicador de lagrange $\lambda_0^* \in P^*$, tal que:

$$\inf_{x \in K} f(x) = \inf_{x \in C} L(1, \lambda_0^*, x).$$

b)

$$\inf_{x \in K} f(x) = \max_{\lambda^* \in P^*} \inf_{x \in C} L(1, \lambda^*, x).$$

c) $\text{cone}(\text{int}(\text{co}(F(C)) - (\mu, 0) + \text{int}(\mathbb{R}_+ \times P)))$ es pointed.

Demostración. Primero veamos b) implica a). Por hipótesis existe $\lambda_0^* \in P^*$ tal que: $\inf_{x\in K} f(x) = \inf_{x\in C} L(1, \lambda_0^*, x)$, luego de ésta manera tenemos a).

Ahora mostramos a) implica b). Sea $\mu = \inf_{x\in K} f(x)$ y por a) tenemos $\inf_{x\in C} L(1, \lambda^*, x) \leq \mu$, para todo $\lambda^* \in P^*$, luego $\max_{\lambda^*\in P^*} \inf_{x\in C} L(1, \lambda^*, x) = \mu = \inf_{x\in C} L(1, \lambda_0^*, x)$.

Notemos que a) implica c) es inmediato por el teorema anterior.

Probamos c) implica a). Si $\gamma_0^* = 0$ entonces $0 \neq \lambda_0^* \in P^*$ y $\langle \lambda_0^*, y\rangle \geq 0$ para todo $y \in \mathbb{R}^m$, por demostrar que $\langle \lambda_0^*, y\rangle \geq 0$. En efecto, sea $y \in \overline{\text{cone}}\,(g(C) + P)$, entonces existe $z_n = t(g(x_n) + P)$ tal que $z_n \to y$, $(t(g(x_n) + P) \to y)$ luego $\lim_{n\in\mathbb{N}} t(g(x_n) + P, \lambda_0^*) = \langle y, \lambda_0^*\rangle$ donde

$$\begin{aligned} \langle \lambda_0^*, y\rangle &= \lim_{n\in\mathbb{N}} t\,\langle g(x_n), \lambda_0^*\rangle + t\,\langle P, \lambda_0^*\rangle \\ &\geq \left\langle \lim_{n\in\mathbb{N}} tg(x_n), \lambda_0^* \right\rangle \\ &\geq 0, \end{aligned}$$

para todo $y \in \mathbb{R}^m$. Si $y = -\lambda_0^*$ entonces $\langle \lambda_0^*, -\lambda_0^*\rangle \geq 0$. Luego $\|\lambda_0^*\| \leq 0$ y por la condición de Slater generalizado $\lambda_0^* = 0$ que nos da una contradicción, pues $0 \neq \lambda_0^*$. Ahora suponiendo $\gamma_0^* = 1$ tendremos que:

$$\inf_{x\in K} f(x) \geq \inf_{x\in C} L(1, \lambda_0^*, x). \tag{3.6}$$

Y por otro lado como $\text{cone}(\text{int}(\text{co}(F(C)) - (\mu, 0) + (\mathbb{R}_+ \times P)))$ es pointed por teorema anterior $(0, 0) \notin \text{cone}(\text{int}(\text{co}(F(C)) - (\mu, 0) + (\mathbb{R}_+ \times P)))$ luego $\gamma_0^* f(x) + \langle \lambda_0^*, g(x)\rangle \geq \gamma_0^*\mu$, para todo $x \in C$ y como $\gamma_0^* = 1$ se tendrá:

$$\inf_{x\in C} L(1, \lambda_0^*, x) \geq \inf_{x\in K} f(x). \tag{3.7}$$

De (3.6) y (3.7) concluimos que

$$\inf_{x\in K} f(x) = \inf_{x\in C} L(1, \lambda_0^*, x).$$

□

3.4. Caso general con una restricción: Formulación del problema.

En ésta sección, abarcamos dualidad fuerte (3.1) de acuerdo al enfoque planteado en 5), en el cual a partir de una completa descripción que se obtuvo de la condición de pointed del conjunto:

$$\operatorname{cone}(F(C)-(\mu,0)+\operatorname{int}\mathbb{R}^2_+)$$

mencionado en el anterior teorema. A raíz de esto, una nueva caracterización de $(D.F.)$, la cual entre otras, evidencia la no dependencia de la condición de Slater. En éste caso al tratarse de una restricción, el problema (P) queda,

$$\begin{aligned}\mu = \inf_{g(x)\leq 0} f(x)\\ x\in C.\end{aligned} \tag{P}$$

El Lagrangiano usual asociado a (P) es

$$L(\gamma^*,\lambda^*,x)=\gamma^* f(x)+\lambda^* g(x),$$

donde $\gamma^*\geq 0$ y $\lambda^*\geq 0$ se conocen como multiplicadores de Lagrange. Notemos además que $K=\{x\in C: g(x)\leq 0\}$ y luego obtenemos

$$\gamma^*(f(x)-\mu)+\lambda^* g(x)\geq 0, \qquad\qquad \text{para todo } x\in C,$$

lo cual implica Dualidad Fuerte cuando $\gamma^*>0$. Dados $\gamma^*\geq 0$, $\lambda^*\geq 0$ y haciendo $F(C)=\{(f(x),g(x)) : x\in C\}$ con $\rho=(\gamma^*,\lambda^*)$, la anterior desigualdad la podemos escribir como:

$$\langle\rho,b\rangle\geq 0 \qquad\qquad \text{para todo } b\in F(C)-(\mu,0)$$

lo que a su vez podemos reescribirla como:

$$\langle\rho,a\rangle\geq 0 \qquad\qquad \text{para todo } a\in F(C)-(\mu,0)+\mathbb{R}^2_+.$$

Ahora, antes de caracterizar ésta relación con el siguiente teorema, veamos algunos resultados:

Lema 7. Si P es un cono, $a\in\operatorname{int}(P)$ entonces para todo $t>0$, $ta\in\operatorname{int}(P)$.

Demostración. Existe $r>0$, $B(a,r)\subset P$. Sea $t>0$, tomando $s=tr$ afirmamos que $B(ta,s)\subset P$. En efecto; si $b\in B(ta,s)$ entonces

$$\begin{aligned}|b-ta|&<s\\ t\left|\frac{b}{t}-a\right|&<s\\ \left|\frac{b}{t}-a\right|&<\frac{s}{t}=r\end{aligned}$$

entonces $\frac{b}{t}\in B(a,r)\subset P$ por lo tanto $b=t\frac{b}{t}\in P$. □

Lema 8. Sea $P \subseteq \mathbb{R}^2$ un cono, convexo, cerrado de modo tal que $\operatorname{int} P \neq \emptyset$ y $A \subset \mathbb{R}^2$ un conjunto no vacío. Entonces

$$A \cap (-\operatorname{int} P) = \emptyset \text{ si y solo si } \overline{\operatorname{cone}}(A+P) \cap (-\operatorname{int} P) = \emptyset.$$

Demostración. Primero notemos que $-\operatorname{int} P = \operatorname{int}(-P)$ y como $A \cap (-\operatorname{int} P) = \emptyset$ se tendrá que $A \subset [\operatorname{int}(-P)]^c$.

Afirmamos que el $[\operatorname{int}(-P)]^c$ es un cono cerrado. En efecto; $[\operatorname{int}(-P)]^c$ es cerrado pues $\operatorname{int}(-P)$ es abierto. Ahora veamos que $[\operatorname{int}(-P)]^c$ es un cono. En efecto; sean $a \in [\operatorname{int}(-P)]^c$, $t \geq 0$ y supongamos que $ta \notin [\operatorname{int}(-P)]^c$. Entonces $ta \in \operatorname{int}(-P)$ y por hipótesis $a \notin \operatorname{int}(-P)$. Como $-P$ es un cono se tendrá que $t > 0$, pues si $t = 0$ entonces $ta = 0 \notin \operatorname{int}(-P)$ (salvo si $-P = \mathbb{R}^2$ que no es el caso). Por tanto $a = \dfrac{1}{t}(ta) \in \operatorname{int}(-P)$, lo cual es una contradicción.

Como $A \subset [\operatorname{int}(-P)]^c$ se tendrá que $A + P \subset [\operatorname{int}(-P)]^c$. Para mostrar este hecho tomemos $a + p \in A + P$, donde $a \in A$ y $p \in P$. Supongamos que $a + p \notin [\operatorname{int}(-P)]^c$. Luego $a + p \in \operatorname{int}(-P)$ del cual deducimos que $-(a + p) \in \operatorname{int} P \subset P$. Entonces existe $r > 0$ tal que $B(-a - p, r) \subset \operatorname{int} P$. Ahora, notemos que $a + p \in \operatorname{int}(-P) \subset -P$, luego $a = a + p - p \in -P$. Por otro lado $B(a + p - p, r) \subset P$, luego $B(a + p - p, r) \subset -P - p$. Por tanto $a \in \operatorname{int}(-P)$ lo cual es una contradicción ya que $A \cap (\operatorname{int}(-P)) = \emptyset$.

Y por lo mostrado en el párrafo anterior tenemos que $\operatorname{cone}(A + P) \subset [\operatorname{int}(-P)]^c$. Así $\overline{\operatorname{cone}}(A+P) \subset [\operatorname{int}(-P)]^c$, por tanto $\overline{\operatorname{cone}}(A+P) \cap \operatorname{int}(-P) = \emptyset$. Recíprocamente; como $A \subset A+P$ entonces $A \subset \overline{\operatorname{cone}}(A+P)$ luego $A \cap (-\operatorname{int} P) \subset \overline{\operatorname{cone}}(A+P) \cap (-\operatorname{int} P) = \emptyset$, por tanto $A \cap (-\operatorname{int} P) = \emptyset$. □

Teorema 14. Sea $P \subseteq \mathbb{R}^2$ un cono, convexo, cerrado de modo tal que $\operatorname{int} P \neq \emptyset$ y $A \subset \mathbb{R}^2$ un conjunto no vacío tal que $A \cap -\operatorname{int} P = \emptyset$. Entonces las siguientes proposiciones son equivalentes:

a) Existe $\lambda^* \in P^* \setminus 0 : \langle \lambda^*, a \rangle \geq 0$, para todo $a \in A$;

b) $\overline{\operatorname{cone}}(A + P)$ es convexo;

c) $\operatorname{cone}_+(A + \operatorname{int} P)$ es convexo;

d) $\operatorname{cone}(A + \operatorname{int} P)$ es convexo;

e) $\operatorname{cone}(A + \operatorname{int} P)$ es pointed;

f) $\operatorname{co}(A) \cap -\operatorname{int} P = \emptyset$

Demostración. Mostremos a) implica b).Por hipótesis existe $\lambda^* \in P^* \setminus 0$ tal que $\langle \lambda^*, x \rangle \geq 0$ para todo $x \in A$ y claramente $x \in \overline{\operatorname{cone}}(A + P)$. Sean $u \in \operatorname{int} P$, $y, z \in A$ entonces

$\text{cone}(\{y\}) + \text{cone}(\{u\}) = \{sy + tu; s, t \geq 0\}$ es un cono convexo y cerrado que contiene a u y y, y está contenido en $\text{cone}(A + P)$. Lo mismo es cierto para $\text{cone}(\{z\}) + \text{cone}(\{u\}) = \{\alpha z + \beta u; \alpha, \beta \geq 0\}$ es un cono convexo y cerrado. Ambos conos tienen en común a $\text{cone}(\{u\})$ por lo tanto la unión de ambos conos está contenida en

$$\text{cone}(\{A + P\}) \subseteq \left\{x \in \mathbb{R}^2 : \langle \lambda^*, x \rangle \geq 0\right\}. \tag{3.8}$$

Antes mostremos(3.8). Sea $a \in \text{cone}(\{A + P\})$ entonces $a = t(c+k)$, donde $c \in A$ y $k \in P$. Luego

$$\begin{aligned}
\langle \lambda^*, a \rangle &= t \langle \lambda^*, c + k \rangle \\
&= t \langle \lambda^*, c \rangle + t \langle \lambda^*, k \rangle \\
&\geq 0 + t \langle \lambda^*, k \rangle \\
&\geq 0
\end{aligned}$$

Por tanto $a \in \{x \in \mathbb{R}^2 : \langle \lambda^*, x \rangle \geq 0\}$.

Volviendo a la idea anterior, definimos

$$B = \text{cone}(\{y\}) + \text{cone}(\{u\}) \bigcup \text{cone}(\{z\}) + \text{cone}(\{u\}).$$

B es un cono convexo. En efecto; es claro que la unión de dos conos es un cono. La convexidad la obtenemos de lo siguiente. Sean $C_1 = \text{cone}\{u\} + \text{cone}\{y\}$ y $C_2 = \text{cone}\{u\} + \text{cone}\{z\}$ y analizamos los siguientes casos:

i) si $z \in \text{cone}\{y\}$ ó $z \in \text{cone}\{u\}$ entonces $C_1 = C_2$ luego $C_1 \bigcup C_2 = C_1$ donde C_1 es convexo.

ii) Si $z \in \text{int}(C_1)$ entonces $C_2 \subset C_1$, pues tomando $a \in C_2$ tendremos que:

$$\begin{aligned}
a &= tu + pz \\
&= tu + p(mu + ny) \\
&= (t + pm)u + pny \\
&= \alpha u + \beta y \in C_1
\end{aligned}$$

por lo tanto $C_1 \bigcup C_2 = C_1$ convexo.

iii) Afirmamos que si $z \notin C_1$ entonces $C_1 \bigcup C_2 = \text{cone}\{z\} + \text{cone}\{y\}$. En efecto; Si $u \in \text{cone}\{z\} + \text{cone}\{y\}$ entonces $u = \alpha_1 z + \beta_1 y \in C_1 \bigcup C_2$. Recíprocamente, sea $\alpha u + \beta y \in C_1$ entonces

$$\begin{aligned}
\alpha(\alpha_1 z + \beta_2 y) + \beta y &= \alpha\alpha_1 z + \alpha\beta_1 y + \beta y \\
&= \alpha\alpha_1 z + (\alpha\beta_1 + \beta)y \in \text{cone}\{z\} + \text{cone}\{y\}.
\end{aligned}$$

es decir, $\alpha u + \beta z \in C_2$. Luego

$$\begin{aligned}\alpha(\alpha_1 z + \beta_1 y) + \beta z &= \alpha\alpha_1 z + \alpha\beta_1 y + \beta z\\ &= (\alpha\alpha_1 + \beta)z + \alpha\beta_1 y \in \operatorname{cone}\{z\} + \operatorname{cone}\{y\}.\end{aligned}$$

Con lo anterior hemos mostrado que B es un cono convexo.

Finalmente, dados $y, z \in B$ se deduce que: $[y, z] \subseteq B \subseteq \operatorname{cone}(A + P)$ y además como $y, z \in A$ se tendrá $\operatorname{co}(A) \subseteq \overline{\operatorname{cone}}(A + P)$ por lo tanto, como P es cono convexo tendremos $\overline{\operatorname{cone}}(A + P)$ es convexo.

Veamos b) implica a). Esto es una consecuencia de argumentos de separación de convexos debido a que

$$A \cap (-\operatorname{int} P) = \emptyset \text{ si y solo si } \overline{\operatorname{cone}}(A + P) \cap (-\operatorname{int} P) = \emptyset.$$

Luego, existe $\lambda^* \neq 0$ tal que $\langle \lambda^*, p\rangle \leq 0$ para todo $p \in \operatorname{int}(-P)$ y $\langle \lambda^*, x\rangle \geq 0$ para todo $x \in \overline{\operatorname{cone}}(A + P)$. Además tenemos que $P \subset \overline{\operatorname{cone}}(A + P)$. Si $p \in P$, tendremos que $p \in \operatorname{int}(P)$ ó $p \in \partial P$, analicemos éstos casos:

i) Si $p \in \operatorname{int} P$ entonces $-p \in -\operatorname{int}(P)$, luego $\langle \lambda^*, -p\rangle \leq 0$ y por lo tanto $\langle \lambda^*, p\rangle \geq 0$

ii) Si $p \in \partial P$ entonces, como $\overline{\operatorname{int}(P)} = P$, existe una sucesión $\{p_n\} \subset \operatorname{int}(P)$ tal que $p_n \to p$. Luego decimos que $\{-p_n\} \subset -\operatorname{int}(P)$, y que $\langle \lambda^*, -p_n\rangle \leq 0$. Entonces $\langle \lambda^*, p_n\rangle \geq 0$ y $\langle \lambda^*, p\rangle \geq 0$. Por tanto $\lambda^* \in P^*$.

Es decir; en cualquier caso λ^* está en P^* y como a pertenece a A y $A \subset \overline{\operatorname{cone}}(A + P)$ concluimos $\langle \lambda^*, a\rangle \geq 0$ para todo a en A.

Veamos c) implica b). Primero mostremos que $\overline{\operatorname{cone}}_+(A + \operatorname{int} P) = \overline{\operatorname{cone}}(A + P)$. En efecto; como $A + \operatorname{int} P \subset A + P$ entonces $\operatorname{cone}_+(A + \operatorname{int} P) \subset \operatorname{cone}_+(A + P) \subset \operatorname{cone}(A + P)$. Por tanto $\overline{\operatorname{cone}}_+(A + \operatorname{int} P) \subset \overline{\operatorname{cone}}(A + P)$. Recíprocamente; $\operatorname{cone}(A + P) \subset \overline{\operatorname{cone}}_+(A + \operatorname{int} P)$, pues tomando $t(a + p) \in \operatorname{cone}(A + P)$ y $t \geq 0$ existen $\{t_n\}$ tal que $t_n \to t^+$ y $\{p_n\} \subset \operatorname{int}(P)$ tal que $(p_n) \to p$. Luego $t_n(a + p_n) \in \operatorname{cone}_+(A + \operatorname{int} P)$ y además $t_n(a + p_n) \to t(a + p)$. Así; $t(a + p) \in \overline{\operatorname{cone}}_+(A + \operatorname{int} P)$ y por tanto $\overline{\operatorname{cone}}(A + P) \subset \overline{\operatorname{cone}}_+(A + \operatorname{int} P)$.

Ahora, por propiedades de convexos sabemos que si $\operatorname{cone}_+(A + \operatorname{int}(P))$ es convexo, entonces $\overline{\operatorname{cone}}_+(A + \operatorname{int}(P))$ es convexo. Y por lo mostrado anteriormente tenemos que $\overline{\operatorname{cone}}_+(A + \operatorname{int} P) = \overline{\operatorname{cone}}(A + P)$, con lo cual decimos que $\overline{\operatorname{cone}}(A + P)$ es convexo.

Ahora veamos b) implica c). Notemos antes que $\overline{\operatorname{cone}}(A + P) = \overline{\operatorname{cone}}_+(A + P)$ y además convexo y por propiedades de conjuntos convexos implica que $\operatorname{int}\left(\overline{\operatorname{cone}_+(A + P)}\right)$

es convexo. Así,

$$\begin{aligned}
\operatorname{int}(\overline{\operatorname{cone}_+((A)+P}) &= \operatorname{int}(\overline{\operatorname{cone}_+(A)+P}) \\
&= \operatorname{int}(\operatorname{cone}_+(A)+P) \\
&= \operatorname{cone}_+(A)+\operatorname{int} P \\
&= \operatorname{cone}_+(A+\operatorname{int} P),
\end{aligned}$$

Usando el hecho de que $\operatorname{cone}_+(A)+P$ es convexo. Por tanto $\operatorname{cone}_+(A+\operatorname{int} P)$ es convexo.

c) implica d) es inmediato.

Mostremos d) implica b). Como $\operatorname{cone}(A+\operatorname{int} P)$ y por propiedades de conjuntos convexos implica $\overline{\operatorname{cone}}(A+\operatorname{int}(P))$ es convexo. Luego, veamos que $\overline{\operatorname{cone}}(A+\operatorname{int}(P)) = \overline{\operatorname{cone}}(A+P)$. Sabemos que $A+\operatorname{int}(P) \subset A+P$ implica que $\operatorname{cone}(A+\operatorname{int}(P)) \subset \operatorname{cone}(A+P)$, por lo tanto $\overline{\operatorname{cone}}(A+\operatorname{int}(P)) \subset \overline{\operatorname{cone}}(A+P)$. Recíprocamente, por la demostración en c) implica b) tenemos que $\operatorname{cone}(A+P) \subset \overline{\operatorname{cone}_+}(A+\operatorname{int}(P)) \subset \overline{\operatorname{cone}}(A+\operatorname{int}(P))$, por tanto $\overline{\operatorname{cone}}(A+P) \subset \overline{\operatorname{cone}}(A+\operatorname{int}(P))$. Así $\overline{\operatorname{cone}}(A+P)$ es convexo.

c) implica e). En efecto, sean $y, -y \in \operatorname{cone}(A+\operatorname{int} P)$. Razonamos por contradicción, supongamos $y \neq 0$. Por hipótesis $\operatorname{cone}_+(A+\operatorname{int} P)$ es convexo. Luego $0 = \frac{1}{2}y + \frac{1}{2}(-y) \in \operatorname{cone}_+(A+\operatorname{int} P)$ lo cual implica que existen $t>0$, $a \in A$ y $p \in \operatorname{int} P$ tales que $0 = t(a+p)$. Entonces $a = -p$ e implica que $A \cap (-\operatorname{int} P) \neq \emptyset$ lo cual es una contradicción.

Veamos e) implica f). Supongamos por contradicción que existe $x \in \operatorname{co}(A) \bigcap (\operatorname{int} -P)$, luego $x \in \operatorname{co}(A)$ y $-x \in \operatorname{int} P$. Así, $0 = x + (-x) \in \operatorname{co}(A) + \operatorname{int} P$. Además, $\operatorname{co}(A)+\operatorname{int} P$ es un abierto que contiene al cero. Concluimos diciendo que $\operatorname{cone}(\operatorname{co}(A)+\operatorname{int} P) = \mathbb{R}^2$ lo cual es una contradicción, pues $\operatorname{cone}(A+\operatorname{int} P)$ es pointed.

Finalmente, mostremos f) implica a). Por teorema de separación entre convexos, existe $p \in \mathbb{R}^2 \setminus \{0\}$ y $\alpha \in \mathbb{R}$, tal que $\langle p, z\rangle \geq \alpha$, para todo $z \in \operatorname{co}(A)$ y $\langle p, w\rangle \leq \alpha$, para toda $w \in -\overline{\operatorname{int} P} = -P$. De la segunda desigualdad tendremos que $\langle p, -w\rangle \geq -\alpha$, para toda $-w$ en P. Sea $-w = nq$, para todo $q \in P$ y tendremos que $\langle p, nq\rangle \geq -\alpha$, es decir; $\langle p, q\rangle \geq -\frac{\alpha}{n}$.

Tomando el límite a la última desigualdad cuando $n \to +\infty$ obtenemos $\langle p, q\rangle \geq 0$, para todo $q \in P$. Y por la definición de P^* tendremos que $p \in P^*$. Ahora, pongamos $p = \lambda^*$ y en la primera desigualdad tendremos que $\langle \lambda^*, z\rangle \geq 0$, para todo $z \in \operatorname{co}(A)$, en particular $\langle \lambda^*, a\rangle \geq 0$, para todo $a \in A$. $\square$

En virtud del resultado anterior y del razonamiento empleado en A complete characterization of strong duality in nonconvex optimization with a single constraint, la idea es particionar el cono $\operatorname{cone}(F(C)-\mu(1,0)+\mathbb{R}^2_{++})$, para determinar las condiciones que permitan asegurar la condición de Pointed con respecto a las particiones, se denota $\mathbb{R}^2_{++} = \operatorname{int} \mathbb{R}^2_+$.

Definición 14. $K = \{x \in C : g(x) \le 0\} = S_g^-(0) \cup S_g^=(0)$, en donde,

$$\begin{aligned}
S_g^-(0) &= \{x \in C : g(x) < 0\}\,,\\
S_g^=(0) &= \{x \in C : g(x) = 0\}\,,\\
S_g^+(0) &= \{x \in C : g(x) > 0\}\,.
\end{aligned}$$

Definición 15.

$$\begin{aligned}
S_f^-(\mu) &= \{x \in C : f(x) < \mu\}\,,\\
S_f^=(\mu) &= \{x \in C : f(x) = \mu\}\,,\\
S_f^+(\mu) &= \{x \in C : f(x) > \mu\}\,.
\end{aligned}$$

Al particionar $F(C) - \mu(1,0) + \mathbb{R}^2_{++} = \Omega_1 \cup \Omega_2 \cup \Omega_3$ se obtiene $\operatorname{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \operatorname{cone}(\Omega_1) \cup \operatorname{cone}(\Omega_2) \cup \operatorname{cone}(\Omega_3)$, donde:

- $\Omega_1 = \displaystyle\bigcup_{x \in \operatorname{argmin}_K f \cap S_g^=(0)} \left[(0,0) + \mathbb{R}^2_{++}\right] \cup \bigcup_{x \in \operatorname{argmin}_K f \cap S_g^-(0)} \left[(0, g(x)) + \mathbb{R}^2_{++}\right].$
- $\Omega_2 = \displaystyle\bigcup_{x \in S_f^+(\mu) \cap S_g^-(0)} \left[(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}\right] \cup \bigcup_{x \in S_f^+(\mu) \cap S_g^=(0)} \left[(f(x) - \mu, 0) + \mathbb{R}^2_{++}\right].$
- $\Omega_3 = \displaystyle\bigcup_{x \in S_f^-(\mu) \cap S_g^+(0)} \left[(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}\right] \cup \bigcup_{x \in S_f^=(\mu) \cap S_g^+(0)} \left[(0, g(x)) + \mathbb{R}^2_{++}\right] \cup$
  $\displaystyle\bigcup_{x \in S_f^+(\mu) \cap S_g^+(0)} \left[(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}\right].$

Observación 5. Por definición de μ tenemos que $f(x) < \mu$ implica $g(x) > 0$, luego $S_f^-(\mu) \cap S_g^+(0) = S_f^-(\mu)$.

Por otro lado, mientras $S_f^+(\mu) \cap S_g^-(0)$, $S_f^-(\mu)$ sean no vacíos, se definen respectivamente:

$$r = \inf_{x \in S_f^+(\mu) \cap S_g^-(0)} \frac{g(x)}{f(x) - \mu}, \qquad s = \sup_{x \in S_f^-(\mu)} \frac{g(x)}{f(x) - \mu}.$$

Es claro que, $-\infty \le r < 0$, $-\infty < s \le 0$.

La siguiente proposición recoge propiedades básicas de los conjuntos previamente definidos.

Proposición 11. $K \neq \emptyset$ y μ finito. Entonces:

a) $C = K$ si y solo si $S_g^+(0) = \emptyset$;

b) $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$ y $S_f^+(\mu) \cap S_g^-(0) = \emptyset$ si y solo si $S_g^-(0) = \emptyset$;

c) $S_f^+(\mu) \cap S_g^-(0) = \emptyset$ si y solo si $S_g^-(0) \subseteq \operatorname{argmin}_K f$;

d) $S_f^-(\mu) = \emptyset$ si y solo si $\mu = \inf\limits_{x \in C} f(x)$.

Demostración. Comencemos mostrando a). Supongamos que $S_g^+(0) \neq \emptyset$ entonces existe $x \in C$ tal que $g(x) > 0$ luego $x \notin K$. Por lo tanto $C \neq K$.

Recíprocamente. Sea $x \in K$ entonces $x \in S_g^-(0)$ o $x \in S_g^=(0)$. Tomamos en cuenta los siguientes casos:

i) Si $x \in S_g^-(0)$ entonces $g(x) < 0$; $x \in C$, por tanto $K \subseteq C$.

ii) Si $x \in S_g^=(0)$ entonces $g(x) = 0$; $x \in C$, por tanto $K \subseteq C$.

Por otro lado, sea $x \in C$. Entonces $x \in S_g^-(0) \cup S_g^=(0) \cup S_g^+(0)$ y como $S_g^+(0) = \emptyset$ se tendrá que $x \in S_g^-(0) \cup S_g^=(0)$ y por definición $x \in K$, por tanto $C \subseteq K$.

Mostremos b). Por hipótesis $S_g^-(0) = \emptyset$ luego $\operatorname{argmin}_K f \cap \emptyset = \operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$ y $S_f^+(\mu) \cap \emptyset = S_f^+(\mu) \cap S_g^-(0) = \emptyset$.

Recíprocamente, supongamos $S_g^-(0) \neq \emptyset$ entonces existe $x_0 \in S_g^-(0)$ y por definición de $S_g^-(0)$ se tendrá $x_0 \in C$ tal que $g(x_0) < 0$ y por la observación 4 tenemos $f(x_0) \geq \mu$ lo cual genera dos casos:

i) Si $f(x_0) = \mu$ y como $\mu \in \mathbb{R}$ entonces $x_0 \in \operatorname{argmin}_K f$. Luego $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$, lo que es una contradicción.

ii) Si $f(x_0) > \mu$ entonces $x_0 \in S_f^+(\mu)$. Por tanto $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, también es una contradicción.

En consecuencia $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$ y $S_f^+(\mu) \cap S_g^-(0) = \emptyset$.

Ahora procedemos a demostrar c). De la hipótesis $S_f^+(\mu) \cap S_g^-(0) = \emptyset$ decimos que $S_g^-(0) \subseteq \left[S_f^+(\mu)\right]^c$ luego $S_g^-(0) \subseteq S_f^-(\mu) \cup S_f^=(\mu)$. Tomando un $x_0 \in S_g^-(0)$ se tendrá $g(x_0) < 0$. Así $f(x_0) \geq \mu$ luego $x_0 \in S_f^=(\mu)$, además como $x_0 \in K$ concluimos que $x_0 \in \operatorname{argmin}_K f$.

Recíprocamente, supongamos que $f(x) > \mu$ y $g(x) < 0$ entonces por hipótesis tenemos que $f(x) = \mu$ para algún $x \in K$ pues $f(x) > \mu$, que es una contradicción.

Veamos d). Como $C = S_f^+(\mu) \cup S_f^=(\mu)$ y si $x \in C$ entonces $f(x) \geq \mu$ y $\inf\limits_{x \in C} f(x) \geq \mu$. Luego $\mu = \inf\limits_{x \in K} f(x) \geq \inf\limits_{x \in C} f(x) \geq \mu$ por lo tanto $\mu = \inf\limits_{x \in C} f(x)$. Recíprocamente, supongamos que $S_f^-(\mu) \neq \emptyset$ entonces existe un $x \in C$ tal que $f(x) < \mu$, que es una contradicción. $\square$

Observación 6. Asumiendo que μ es finito y que el conjunto factible K es no vacío, por tanto

$$[F(C) - \mu(1,0)] \cap \left(-\mathbb{R}^2_{++}\right) = \emptyset.$$

Y por teorema afirmamos que la convexidad de cone $\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$ es equivalente a que sea pointed.

Antes de mostrar el siguiente teorema veamos un corolario.

Corolario 7. Sea $\Omega \in \mathbb{R}^n$. Si $\mathbb{L}_a \cap \Omega \neq \emptyset$ entonces $a \in \text{cone}(\Omega)$.

Demostración. Por hipótesis existe un $x \in \mathbb{L}_a \cap \Omega$, donde $\mathbb{L}_a = \{sa/s \geq 0\}$. Ahora como $x = sa$, $s \geq 0$ y $tx \in \text{cone}(\Omega)$ para todo $t \geq 0$, tendremos los siguientes casos:

Si $s = 0$ entonces $t0a \in \text{cone}(\Omega)$.

Y si $s > 0$ entonces $a = \frac{1}{s}sa = \frac{1}{s}x \in \text{cone}(\Omega)$. $\square$

Teorema 15. Consideremos $\mu = \inf\limits_{x \in K} f(x)$ tal que $K \neq \emptyset$ y μ finito.

a) Asumimos que $\text{argmin}_K f \neq \emptyset$. Entonces, $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ es pointed si se cumple cualquiera de las siguientes condiciones:

- a1) $\text{argmin}_K f \cap S^-_g(0) \neq \emptyset$ y, $S^+_g(0) = \emptyset$ ó $\left[S^+_g(0) \cap S^-_f(\mu) = \emptyset, S^+_g(0) \neq \emptyset\right]$;
- a2) $\text{argmin}_K f \cap S^-_g(0) = \emptyset$, $\text{argmin}_K f \cap S^=_g(0) \neq \emptyset$ y $K = \text{argmin}_K f$;
- a3) $\text{argmin}_K f \cap S^-_g(0) = \emptyset$, $\text{argmin}_K f \cap S^=_g(0) \neq \emptyset$, $S^-_g(0) \cap S^+_f(\mu) \neq \emptyset$, $-\infty < r < 0$ y cualquiera $S^+_g(0) = \emptyset$ ó $\left[S^+_g(0) \cap S^-_f(\mu) \neq \emptyset, s \leq r\right]$, ó $\left[S^+_g(0) \cap S^-_f(\mu) = \emptyset, S^+_g(0) \neq \emptyset\right]$;
- a4) $\text{argmin}_K f \cap S^-_g(0) = \emptyset$, $\text{argmin}_K f \cap S^=_g(0) \neq \emptyset$, $S^-_g(0) \cap S^+_f(\mu) \neq \emptyset$, $r = -\infty$ y, cualquiera $S^+_g(0) = \emptyset$ ó $\left[S^+_g(0) \cap S^-_f(\mu) = \emptyset, S^+_g(0) \neq \emptyset\right]$;
- a5) $\text{argmin}_K f \cap S^-_g(0) = \emptyset$, $\text{argmin}_K \cap S^=_g(0) \neq \emptyset$, $S^-_g(0) \cap S^+_f(\mu) = \emptyset$ y $S^=_g(0) \cap S^+_f(\mu) \neq \emptyset$.

b) Asumiendo que $\text{argmin}_K f = \emptyset$. Entonces, $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ es pointed si se cumple cualquiera de las siguientes condiciones:

- b1) $S^-_g(0) \cap S^+_f(\mu) \neq \emptyset$, $-\infty < r < 0$ y, cualquiera $S^+_g(0) = \emptyset$ ó $\left[S^+_g(0) \cap S^-_f(\mu) \neq \emptyset,\ s \leq r\right]$ ó $\left[S^+_g(0) \cap S^-_f(\mu) = \emptyset, S^+_g(0) \neq \emptyset\right]$;
- b2) $S^-_g(0) \cap S^+_f(\mu) \neq \emptyset$, $r = -\infty$ y, $S^+_g(0) = \emptyset$ o $\left[S^+_g(0) \cap S^-_f(\mu) = \emptyset, S^+_g(0) \neq \emptyset\right]$;
- b3) $S^-_g(0) \cap S^+_f(\mu) = \emptyset$, $S^=_g(0) \cap S^+_f(\mu) \neq \emptyset$.

Demostración. Veamos a1) implica a). Si $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$ y $S_g^+(0) = \emptyset$ entonces existe un $x_1 \in K$, $f(x_1) = \mu$, $g(x_1) < 0$ luego $\Omega = (f(x_1) - \mu, g(x_1)) + \mathbb{R}^2_{++} \subset \Omega_1$.

Afirmamos que $\operatorname{cone}(\Omega_1) = \mathbb{R}_+ \times \mathbb{R} \cup (0,0)$. En efecto, para $t = 0$ y $a \in \Omega_1$ implica que $ta = (0,0) \in \{(0,0)\}$. Por otro lado sabemos que el $\operatorname{cone}(\Omega_1) \subset \mathbb{R}_+ \times \mathbb{R}$ implica que $\operatorname{cone}(\operatorname{cone}(\Omega_1)) \subset \operatorname{cone}(\mathbb{R}_+ \times \mathbb{R}) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$. Recíprocamente, si $(a,b) \in \mathbb{R}_+ \times -\mathbb{R}_+$ tendremos $\mathbb{L}_{ab} \cap \Omega \neq \emptyset$ como consecuencia $\mathbb{L}_{ab} \cap \Omega_1$ luego por corolario $(a,b) \in \operatorname{cone}(\Omega_1)$. Además $\operatorname{cone}(\Omega_2) \subset \operatorname{cone}(\Omega_1)$ pues $\Omega_2 \subset \mathbb{R}_+ \times \mathbb{R}$ y $\operatorname{cone}(\Omega_3) = \emptyset$ pues $S_g^+(0) = \emptyset$. Finalmente concluimos que

$$\operatorname{cone}(\Omega_1) \cup \operatorname{cone}(\Omega_2) \cup \operatorname{cone}(\Omega_3) = \operatorname{cone}(\Omega_1) = \operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$$

es convexo y por teorema $\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$ es pointed.

Ahora, sea la hipótesis con $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$ y $\left[S_g^+(0) \cap S_f^-(\mu) = \emptyset, S_g^+(0) \neq \emptyset\right]$ entonces $\operatorname{cone}(\Omega_1) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$, $\operatorname{cone}(\Omega_2) \subset \operatorname{cone}(\Omega_1)$, pues $\Omega_2 \subset \mathbb{R}_+ \times \mathbb{R}$ y $\operatorname{cone}(\Omega_3) \in \operatorname{cone}(\Omega_1)$, pues $\Omega_3 \subset \mathbb{R}^2_{++}$. Por lo tanto

$$\bigcup_{i=1,2,3.} \operatorname{cone}(\Omega i) = \operatorname{cone}(\Omega_1)$$

es convexo y también decimos que $\operatorname{cone}(\Omega_1) = \operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$ es pointed.

Mostremos que a2) implica a). Si $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, $\operatorname{argmin}_K f \cap S_g^=(0) \neq \emptyset$ y $K = \operatorname{argmin}_K f$ entonces existe $x_1 \in K$ tal que $f(x_1) = \mu$, $g(x_1) = 0$ y se tendrá $\Omega = (f(x_1) - \mu, g(x_1)) + \mathbb{R}^2_{++} \subset \Omega_1$. Luego, tenemos los siguientes casos:

i) $\Omega_1 = \mathbb{R}^2_{++}$ y $\operatorname{cone}(\Omega_1) = \mathbb{R}^2_{++} \cup \{(0,0)\}$. En efecto, sea $(a,b) \in \operatorname{cone}(\Omega_1)$, entonces $(a,b) = t\{(0,0) + (r,p)\} = t(r,p)$ está en $\mathbb{R}^2_{++}$.Recíprocamente, sea $(a,b) \in \mathbb{R}^2_{++}$ y $\mathbb{L}_{(a,b)} \cap \Omega \neq \emptyset$, entonces por corolario $(a,b) \in \operatorname{cone}(\Omega_1)$

ii) Si $x \in S_f^+(\mu) \cap S_g^-(0)$ ó $x \in S_f^+(\mu) \cap S_g^=(0)$ entonces $g(x) \leq 0$ y en consecuencia $x \in K$ lo que es una contradicción. Por lo tanto $\Omega_2 = \emptyset$ y $\operatorname{cone}(\Omega_2) = \emptyset$.

iii) Y si $S_f^-(\mu) \cap S_g^+(0) = \emptyset$ entonces $\Omega_3 \subset \mathbb{R}^2_{++}$ y así $\operatorname{cone}(\Omega_3) \subset \operatorname{cone}(\Omega_1)$.

Concluimos de los anteriores casos que $\bigcup \operatorname{cone}(\Omega i) = \operatorname{cone}(\Omega_1)$, $i = 1,2,3$ y que el $\operatorname{cone}(\Omega_1) = \operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$ es pointed.

Por otro lado, asumiendo la hipótesis juntamente con $S_f^-(\mu) \cap S_g^+(0) \neq \emptyset$ entonces

$$0 < \frac{g(x)}{-(f(x) - \mu)} < \infty, \text{ donde } \tan\theta = \frac{g(x)}{-(f(x) - \mu)},$$

luego

$$0 \leq \inf\left\{\left(\frac{g(x)}{-(f(x) + \mu)}\right) / x \in S_f^-(\mu) \cap S_g^+(0)\right\} < \infty,$$

donde $\alpha = \inf\left\{\left(\frac{g(x)}{-(f(x)+\mu)}\right)/x \in S_f^-(\mu)\cap S_g^+(0)\right\}$.

Sea $A = \{(u,v)\in\mathbb{R}\times\mathbb{R}_+/u\geq 0\}\cup\left\{(u,v)\in\mathbb{R}\times\mathbb{R}_+/\frac{v}{-u}>\alpha, u<0\right\}$. Entonces afirmamos que $\text{cone}(\Omega_3) = A = A_1\cup A_2$, donde $A_1 = \{(u,v)\in\mathbb{R}\times\mathbb{R}_+/u\geq 0\}$ y $A_2 = \left\{(u,v)\in\mathbb{R}\times\mathbb{R}_+/\frac{v}{-u}>\alpha, u<0\right\}$. En efecto, como A_1 es un cono y por la definición de Ω_3 se tiene $\text{cone}(\Omega_3)\subset A_1$. Por otro lado sea $(a,b)\in\text{cone}(\Omega_3)$, luego existe un $t_0\geq 0$ tal que $t_0(a,b)\in\Omega_3$, es decir $t_0(a,b)\in(f(x)-\mu, g(x))+\mathbb{R}^2_{++}$ para algún $x\in S_f^-(\mu)\cap S_g^+(0)$, así

$$\alpha\leq\tan\frac{tb}{-ta}=\tan\frac{b}{-a},\ a<0,$$

entonces $(a,b)\in A_2$ por lo tanto $\text{cone}(\Omega_3)\subset A_2$. De ésta forma hemos concluido que $\text{cone}(\Omega_3)\subset A_1\cup A_2 = A$. Recíprocamente, podemos ver que $A_1\subset\text{cone}(\Omega_3)$. Ahora analicemos el caso en el que $\alpha = 0$. Sea $(a,b)\in\mathbb{R}\times\mathbb{R}_+,\ a<0$. Luego, $\tan\theta = \frac{b}{-a}>0$ entonces existe $x\in S_f^-(\mu)\cap S_g^+(0)$ tal que $\frac{g(x)}{-(f(x)-\mu)}<\frac{b}{-a},\ \frac{b}{-a}>0$ luego

$$\Omega = (f(x)-\mu, g(x))+\mathbb{R}^2_{++}\subset\Omega_3.$$

Sea $\mathbb{L}_{ab}$ la recta que une (a,b) con $(0,0)$, luego $\mathbb{L}_{ab}\cap\Omega\neq\emptyset$. Así, existe $(c,d)\in\mathbb{R}_{ab}\cap\Omega\subset\Omega_3$. Por tanto $(a,b)\in\text{cone}(\Omega_3)$. Ahora, si $\alpha>0$ entonces $0<\alpha<\frac{\pi}{4}$ luego $\frac{b}{-a}<\alpha$ para $(a,b)\in\mathbb{R}\times\mathbb{R}_+$ y en consecuencia $(a,b)\in\text{cone}(\Omega_3)$ y por lo tanto se tiene que $A = A_1\cup A_2\subset\text{cone}(\Omega_3)$.

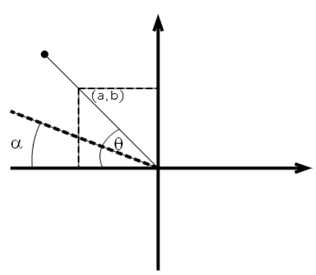


Figura 3.1: a2) implica a)

Veamos a3) implica a). Asumiendo la hipótesis y que $S_g^+(0)=\emptyset$, entonces $\Omega_1 = \mathbb{R}^2_{++}$ luego $\text{cone}(\Omega_1) = \mathbb{R}^2_{++}\cup\{(0,0)\}$. Y por hipótesis $\Omega_3=\emptyset$ lo cual implica que $\text{cone}\,\Omega_3=\emptyset$. Ahora bien, por como está definido $r = \inf\left\{\frac{g(x)}{f(x)-\mu}/x\in S_g^+(0)\cap S_f^+(\mu)\right\}$ se tendrá

$$-r = -\inf\left\{\frac{g(x)}{f(x)-\mu}/x\in S_g^-(0)\cap S_f^+(\mu)\right\} = \sup\left\{\frac{-g(x)}{f(x)-\mu}/x\in S_g^+(0)\cap S_f^+(\mu)\right\},$$

donde $0<-r<\infty$.

Sea $A = \left\{(u,v)\in\mathbb{R}_+\times\mathbb{R}/\frac{-v}{u}>r\right\}$, por demostrar que $\text{cone}(\Omega_2) = A\cup\{(0,0)\}$. En efecto, existe $x_1\in S_g^-(0)\cap S_f^+(\mu)$ tal que $(f(x_1)-\mu, g(x_1))+\mathbb{R}^2_{++}\subset\Omega_2$ y sea $(a,b)\in$

cone(Ω_2). Entonces existe $t_0 \geq 0$ tal que $t_0(a,b) \in \Omega_2$ es decir, $t_0(a,b) \in (f(x_1)-\mu, g(x_1)) + \mathbb{R}^2_{++}$ para algún $x_1 \in S^-_g(0) \cap S^+_f(\mu)$ e implica que $r < \tan \frac{-tb}{ta} = \tan \frac{-b}{a}$, $b < 0$, por lo tanto $(a,b) \in A$. Por otro lado, sea $(a,b) \in$ cone(Ω_2) entonces existe $X \in (\Omega_2)$ tal que $tX = (a,b)$ para $t \geq 0$. Entonces si $X \in (f(x_1) - \mu, g(x)) + \mathbb{R}^2_{++}$ para algún $x \in S^{=}_g(0) \cap S^+_f(\mu)$ del cual implica que $(a,b) \in A$.

Recíprocamente, si $(a,b) \in A$, basta ver cuándo $(a,b) \notin A$, pues $(f(x_1) - \mu, g(x_1)) + \mathbb{R}^2_{++} \subset (\Omega_2)$ y cone$\big((f(x_1) - \mu, g(x_1)) + \mathbb{R}^2_{++}\big) \subset \Omega_2 \subset$ cone$\,\Omega_2 = A$. Como $-\frac{b}{a} < -r$ entonces existe $x \in S^-_g(0) \cap S^+_f(\mu)$ tal que $-\frac{b}{a} < \frac{-g(x)}{f(x)-\mu}$, entonces $(f(x_1) - \mu, g(x_1)) + \mathbb{R}^2_{++} \cap \mathbb{L}_{ab}$ es no vacío, luego $(a,b) \in$ cone(Ω_2). Concluimos que cone$\,\Omega_1 \subset$ cone$\,\Omega_2$ y que $\bigcup$cone$\,\Omega_i =$ cone$\,\Omega_2$ para $i = 1,2,3$ y por teorema decimos que el cone$\big(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\big)$ es pointed.

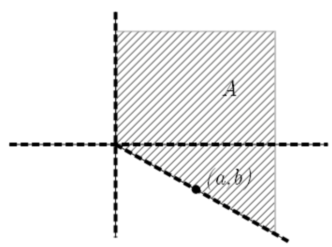


Figura 3.2: a3) implica a).

Por otro lado, además de la hipótesis asumamos que $S^+_g(0) \cap S^-_f(\mu) \neq \emptyset$ con $s \leq r$, entonces como $\Omega_1 = \mathbb{R}^2_{++}$ implica que el cone$(\Omega_1) = \mathbb{R}^2_{++} \cup \{(0,0)\}$. Considerando $A = \big\{(u,v) \in \mathbb{R}_+ \times \mathbb{R} / \frac{-v}{u} > r\big\}$ la región del anterior apartado, entonces cone$(\Omega_2) = A \cup \{(0,0)\}$.

Ahora, sea $B = \big\{(u,v) \in \mathbb{R} \times \mathbb{R}_+ / \frac{v}{-u} > s\big\}$ por demostrar que cone$(\Omega_3) = B \cup \{(0,0)\}$. En efecto, sea $(a,b) \in$ cone(Ω_3) entonces existe $t_0 \geq 0$ tal que $t_0(a,b) \in (f(x_1)-\mu, g(x_1)) \in \Omega_3$ para algún $x \in S^+_g(0) \cap S^-_f(\mu)$. Luego $s \leq \tan \frac{t_0 b}{-t_0 a} = \tan \frac{b}{-a}$, por lo tanto $(a,b) \in B$ y así cone$(\Omega_3) \subset B$.

Recíprocamente, sea $(a,b) \in \mathbb{R} \times \mathbb{R}_+$, $a < 0$. Luego $\tan\theta = \frac{b}{a} > 0$, además existe $x \in S^+_g(0) \cap S^-_f(\mu)$ tal que $s \leq \frac{g(x)}{-(f(x)-\mu)} < \frac{b}{-a}$ y luego $\Omega = (f(x) - \mu, g(x)) + \mathbb{R}^2_{++} \subset \Omega_3$. Sea $\mathbb{L}_{ab}$ la recta que une (a,b) con $(0,0)$ luego $\mathbb{L}_{ab} \cap \Omega$ es distinto del vacío. Así, existe

$(c,d) \in \mathbb{L}_{ab} \cap \Omega_3$. Por lo tanto $(a,b) \in \text{cone}(\Omega_3)$.

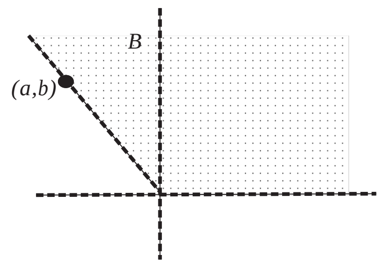


Figura 3.3: a3) implica a).

Mostremos que $A \cup B$ es convexo. En efecto, sean $A = \text{int}\,[\text{cone}\,\{e_2\} + \text{cone}\,\{(1,-r)\}]$ y $B = \text{int}\,[\text{cone}\,\{e_1\} + \text{cone}\,\{(-1,s)\}]$. Dado que $\text{cone}\,\{e_2\} + \text{cone}\,\{(1,-r)\}$ y $\text{cone}\,\{e_1\} + \text{cone}\,\{(-1,s)\}$ son convexos, lo que a su vez implica que A y B son convexos respectivamente, entonces $\overline{A}$ y $\overline{B}$ son convexos, esto por propiedades de convexos. Luego $\overline{A} \cup \overline{B}$ es convexo, además $\overline{A} \cup \overline{B} = \text{cone}\,\{(-1,s)\} + \text{cone}\,\{(1,-r)\}$. Por tanto, $\text{int}\left(\overline{A} \cup \overline{B}\right) = A \cup B$ es convexo.

De ésta manera, $\text{cone}(\Omega_1)$ está contenido en $A \cup B$ y

$$\bigcup_{i=1,2,3} \text{cone}(\Omega_i) = A \cup B \cup \{(0,0)\}\,.$$

Luego concluimos diciendo que

$$A \cup B \cup \{(0,0)\} = \text{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$$

es pointed.

En el caso en el que además de la hipótesis se tiene $S_g^+(0) \cap S_f^-(\mu) = \emptyset$ y $S_g^+(0) \neq \emptyset$ y considerando $A = \left\{(u,v) \in \mathbb{R}_+ \times \mathbb{R} / \frac{-v}{u} > r\right\}$ la región de los anteriores apartados se tendrá que $\text{cone}(\Omega_1) = \mathbb{R}^2_{++} \cup \{(0,0)\}$, $\text{cone}(\Omega_2) = A \cup \{(0,0)\}$ y $\text{cone}(\Omega_3) = \mathbb{R}^2_{++} \cup \{(0,0)\}$. Luego $\text{cone}(\Omega_1) \subset A \cup \{(0,0)\}$, $\text{cone}(\Omega_3) \subset A \cup \{(0,0)\}$. Por lo tanto, $\text{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = A \cup \{(0,0)\} = \text{cone}(\Omega_2)$ el cual es pointed.

Veamos a4) implica a). Con los datos de la hipótesis y asumiendo que $S_g^+(0) = \emptyset$ obtendremos que $\text{cone}(\Omega_1) = \mathbb{R}^2_{++} \cup \{(0,0)\}$, $\text{cone}(\Omega_2) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$ y $\text{cone}\,\Omega_3 = \emptyset$, observemos que $\operatorname{cone}(\Omega_1) \subset \operatorname{cone}(\Omega_2)$ y que

$$\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_2)$$

es pointed.

Por otro lado dada la hipótesis, y además si $S_g^+(0) \cap S_f^-(\mu) = \emptyset$ y $S_g^+(0) \neq \emptyset$ se tendrá que $\operatorname{cone}(\Omega_1) = \mathbb{R}^2_{++} \cup \{(0,0)\}$, $\operatorname{cone}(\Omega_2) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = \mathbb{R}^2_{++} \cup \{(0,0)\}$. Entonces $\operatorname{cone}(\Omega_1) \subset \operatorname{cone}(\Omega_2)$ y $\operatorname{cone}(\Omega_3) \subset \operatorname{cone}(\Omega_2)$. Por lo tanto

$$\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_2)$$

el cual es pointed.

Ahora veamos que a5) implica a). De manera directa afirmamos que $\operatorname{cone}(\Omega_1) = \mathbb{R}^2_{++} \cup \{(0,0)\}$, $\operatorname{cone}(\Omega_2) = \mathbb{R}^2_{++} \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$, entonces $\operatorname{cone}(\Omega_1) \subset \operatorname{cone}(\Omega_3)$ y $\operatorname{cone}(\Omega_2) \subset \operatorname{cone}(\Omega_3)$. Por lo tanto $\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_3)$.

Mostremos que b1) implica b). Los detalles de la demostración son similares a los realizados para demostrar a). Ahora, asumiendo la hipótesis juntamente con $S_g^+(0) = \emptyset$ se tiene que $\operatorname{cone}(\Omega_1) = \emptyset$, $\operatorname{cone} \Omega_2 = A \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = \emptyset$ donde $A = \left\{(u,v) \in \mathbb{R}_+ \times \mathbb{R} / \frac{v}{u} > r\right\}$. Por lo tanto $\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_2)$ que es pointed.

Dada la hipótesis b1) y $S_g^+(0) \cap S_f^-(\mu) \neq \emptyset$ con $s \leq r$ obtendremos que $\operatorname{cone}(\Omega_1) = \emptyset$, $\operatorname{cone}(\Omega_2) = A \cup B \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = \mathbb{R} \times \mathbb{R}_+ \cup \{(0,0)\}$ donde
$A = \left\{(u,v) \in \mathbb{R}_+ \times \mathbb{R} / \frac{-v}{u} > r\right\}$ y $B = \left\{(u,v) \in \mathbb{R} \times \mathbb{R}_+ / \frac{v}{-u} > s\right\}$. Por lo tanto $\operatorname{cone}(\Omega_3) \subset \operatorname{cone}(\Omega_2)$ y

$$\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_2).$$

Ahora asumiendo la hipótesis con $S_g^+(0) \cap S_f^-(\mu) = \emptyset$ y $S_g^+(0) \neq \emptyset$. Entonces el $\operatorname{cone}(\Omega_1) = \emptyset$, $\operatorname{cone}(\Omega_2) = A \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = A \cup \{(0,0)\}$. Luego $\operatorname{cone}(\Omega_3) \subset \operatorname{cone}(\Omega_2)$ $\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_2) = \operatorname{cone}(\Omega_3)$.

Veamos b2) implica b). Dada la hipótesis con $S_g^+(0) = \emptyset$ entonces $\operatorname{cone}(\Omega_1) = \emptyset$, $\operatorname{cone}(\Omega_2) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = \emptyset$. Por lo tanto $\operatorname{cone}\left((f(x_1) - \mu, g(x_1)) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_2)$

Por otro lado, dada la hipótesis con $S_g^+(0) \cap S_f^-(\mu) = \emptyset$ y $S_g^+(0) \neq \emptyset$, se tendrá que $\operatorname{cone}(\Omega_1) = \emptyset$, $\operatorname{cone}(\Omega_3) = \operatorname{cone}(\Omega_2) = \mathbb{R}_+ \times \mathbb{R} \cup \{(0,0)\}$, notemos que $\operatorname{cone}(\Omega_2) \subset \operatorname{cone}(\Omega_3)$, concluimos ésta parte diciendo que $\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_3)$.

Finalmente veamos b3) implica b). Directamente notemos que $\operatorname{cone}(\Omega_1) = \emptyset$, $\operatorname{cone}(\Omega_2) = \mathbb{R}^2_{++} \cup \{(0,0)\}$ y $\operatorname{cone}(\Omega_3) = \mathbb{R} \times \mathbb{R}_+ \cup \{(0,0)\}$. Por lo tanto $\operatorname{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \operatorname{cone}(\Omega_3)$. $\square$

Corolario 8. Sea $K \neq \emptyset$, μ finito.

a) Si $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$ ó $[S_g^-(0) \cap S_f^+(\mu) \neq \emptyset$ con $r = -\infty]$. Entonces

$$\lambda^* \geq 0,\ f(x) + \lambda^* g(x) \geq \mu \text{ para todo } x \in C \text{ implica que } \lambda^* = 0.$$

En consecuencia $\inf_{x \in K} f(x) = \inf_{x \in C} f(x)$.

b) Asuma que $S_g^+(0) = \emptyset$ ó $\left[S_g^+(0) \cap S_f^-(\mu) = \emptyset, S_g^+(0) \neq \emptyset\right]$ son satisfechas; si a3) ó a5) con $\operatorname{argmin}_K f \neq \emptyset$ o [b3) con $\operatorname{argmin}_K f = \emptyset$], entonces, cualquier $(\gamma^*, \lambda^*) \in \mathbb{R}^2_+ \setminus \{(0,0)\}$, verifica

$$\gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \text{ para todo } x \in C.$$

En consecuencia $\gamma^* \inf_{x \in K} f(x) = \inf_{x \in C} L(\gamma^*, \lambda^*, x)$.

c) Asumiendo que $s \leq r$; si [a3) con $\operatorname{argmin}_K f \neq \emptyset$] o [b1) con $\operatorname{argmin}_K f = \emptyset$], entonces cualquier λ^* tal que $-\frac{1}{s} \leq \lambda^* \leq -\frac{1}{r}$ satisface

$$f(x) + \lambda^* g(x) \geq \mu, \text{ para todo } x \in C.$$

d) Asumiendo que $-\infty < s < 0$ y $S_g^+(0) \cap S_f^-(\mu) \neq \emptyset$ si cualquiera, a2) o a5) con $\operatorname{argmin}_K f \neq \emptyset$, o [b3) con $\operatorname{argmin}_K f = \emptyset$], entonces cualquier λ^* tal que $-\frac{1}{s} \leq \lambda^*$ verifica

$$f(x) + \lambda^* g(x) \geq \mu, \text{ para todo } x \in C.$$

e) Asumiendo que $S_g^+(0) = \emptyset$ ó $\left[S_g^+(0) \cap S_f^-(\mu) = \emptyset, S_g^+(0) \neq \emptyset\right]$ son satisfechas; si [a3) con $\operatorname{argmin}_K f \neq \emptyset$] ó [b1) con $\operatorname{argmin}_K f = \emptyset$], entonces cualquier λ^* tal que $-\frac{1}{r} \geq \lambda^* > 0$ verifica

$$f(x) + \lambda^* g(x) \geq \mu, \text{ para todo } x \in C.$$

f) Si $S_g^+(0) \cap S_f^-(\mu) \neq \emptyset$ con $s = 0$, entonces

$$\gamma^* \geq 0,\ \gamma^*(f(x) - \mu) + g(x) \geq 0, \text{ para todo } x \in C \text{ implica } \gamma^* = 0.$$

Demostración. Mostremos a). Asumiendo la hipótesis $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$ supongamos por contradicción que $\lambda^* > 0$ entonces existe $y \in \operatorname{argmin}_K f \cap S_g^-(0)$. Tomemos $\overline{x} \in C$ tal que $f(\overline{x}) + \lambda^* g(\overline{x}) \geq \mu$ luego $g(\overline{x}) \geq 0$ en particular para $y \in \operatorname{argmin}_K f \cap S_g^-(0)$ lo cual es una contradicción.

Por otro lado, asumiendo que $S_g^-(0) \cap S_f^+(\mu) \neq \emptyset$ y $r = -\infty$ existe $y \in S_g^-(0) \cap S_f^+(\mu)$ y supongamos por contradicción que $\lambda^* > 0$. Tomemos $\overline{x} \in C$ tal que $f(\overline{x}) + \lambda^* g(\overline{x}) \geq \mu$ luego

$$f(\overline{x}) - \mu \geq \lambda^*(-g(\overline{x})) > 0$$
$$\frac{1}{\lambda^*} \geq \frac{-g(\overline{x})}{f(\overline{x}) - \mu}$$

entonces

$$\begin{aligned} \inf_{x \in S_g^-(0) \cap S_g^+(\mu)} \left\{ \frac{g(\overline{x})}{f(\overline{x} - \mu)} \right\} &\geq -\frac{1}{\lambda^*} \\ r &\geq -\frac{1}{\lambda^*} \\ -\infty &\geq -\frac{1}{\lambda^*} \end{aligned}$$

lo cuál es una contradicción.

En consecuencia, si $\lambda^* = 0$ entonces $f(x) \geq \mu$ para todo $x \in C$, luego $\inf_{x \in C} f(x) \geq \inf_{x \in K} f(x)$ y por la definición de K concluimos que

$$\inf_{x \in C} f(x) = \inf_{x \in K} f(x).$$

Veamos b). Suponiendo $S_g^+(0) = \emptyset$ con a3), tomemos $(\gamma_0^*, \lambda_0^*) \in \mathbb{R}_+^2 \setminus \{(0,0)\}$, luego por

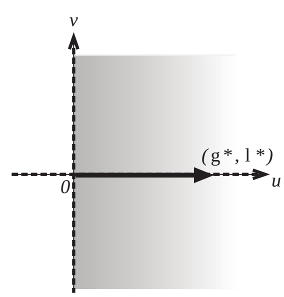


Figura 3.4: Corolario 8.a)

definición de cono polar positivo se tendrá $\langle (\gamma_0^*, \lambda_0^*), (f(x) - \mu, g(x)) \rangle \geq 0$ para todo $x \in C$. Por tanto $\gamma_0^*(f(x) - \mu) + \lambda_0^* g(x) \geq 0$ para todo $x \in C$. Para las otras hipótesis tendremos el mismo razonamiento.

Mostremos c). Por la primera hipótesis de a3) existe $x_1 \in S_g^+(0) \cap S_f^-(\mu)$ tal que $g(x) < 0$, $f(x) > \mu$. Luego por la definición de r y para $x \in C$ tendremos que $r \leq \frac{g(x)}{f(x) - \mu}$ lo cual implica que $-r \geq \frac{-g(x)}{f(x) - \mu}$. Por la condición de λ^* tendremos que

$$\frac{f(x) - \mu}{-g(x)} \geq \frac{1}{-r} \geq \lambda^*.$$

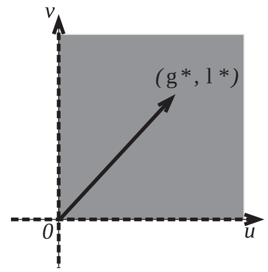


Figura 3.5: Corolario 8.b)

Por tanto $f(x) - \mu \geq \lambda^*(-g(x))$, es decir $f(x) + \lambda^* g(x) \geq \mu$.

Por otro lado, de la segunda hipótesis de a3) existe $x \in C$ tal que $g(x) > 0$ y $f(x) < \mu$. Luego por la definición de s y para $x \in C$ tendremos que $\frac{g(x)}{f(x)-\mu} \leq s$ lo cual implica que $\frac{g(x)}{-(f(x)-\mu)} \geq -s$. Por la condición de λ^* obtendremos que

$$\lambda^* \geq \frac{1}{-s} \geq \frac{-(f(x) - \mu)}{g(x)},$$

por tanto $f(x) + \lambda^* g(x) \geq \mu$.

Ahora, de la tercera hipótesis de a3) se tendrá por teorema anterior que

$$\operatorname{cone}\left(F(C) - (\mu, 0) + \mathbb{R}^2_{++}\right)$$

es pointed. Sean $(1, \lambda_0^*)$ y $x \in C$ tal que $\langle (1, \lambda_0^*), (f(x) - \mu, g(x)) \rangle \geq 0$, luego $f(x) + \lambda_0^* g(x) \geq 0$.

Las demostraciones tomando en cuenta las hipótesis de b1) son similares a la primera parte.

Verifiquemos d). Supongamos a2) y por la definición de s para $x \in C$ tendremos que $-s \leq \frac{g(x)}{-(f(x)-\mu)}$. Luego, por la condición de λ^* tendremos que

$$\frac{-(f(x) - \mu)}{g(x)} \leq \frac{1}{-s} \leq \lambda^*.$$

Por tanto $\mu \leq \lambda^* g(x) + f(x)$.

Suponiendo a5) ó b3) las demostraciones serán similares.

Ahora mostremos el inciso e). Suponiendo que se cumplen a3) con $S_g^+(0) = \emptyset$ y por la definición de r para $x \in C$ tendremos que $r \leq \frac{g(x)}{f(x)-\mu}$ lo cual implica que $-r \geq \frac{-g(x)}{f(x)-\mu}$.

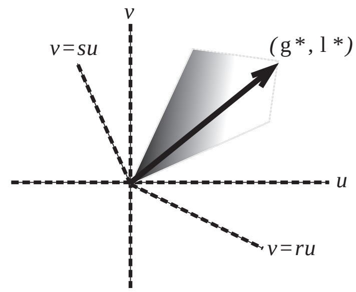


Figura 3.6: Corolario 8.c)

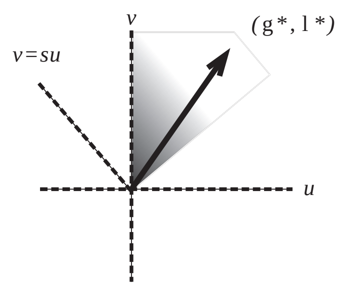


Figura 3.7: Corolario 8.d)

Luego por la condición de λ^* tendremos

$$\frac{f(x)-\mu}{g(x)} \geq \frac{1}{-r} \geq \lambda^* > 0,$$

por tanto $f(x)+\lambda^* g(x) \geq -\mu$. Ahora consideremos la segunda hipótesis de a3). Sea $x \in C$ y por la definición de r tendremos $r \leq \frac{g(x)}{f(x)-\mu}$ luego por la hipótesis concluimos diciendo que $f(x)+\lambda^* g(x) \geq \mu$. Para la tercera hipótesis el tratamiento es similar a las demostraciones anteriores.

Mostremos f). Sea $x \in C$ tal que $\gamma^* \left(f(x)-\mu\right)+g(x) \geq 0$ entonces

$$g(x) \geq \gamma^* \left(-(f(x)-\mu)\right) \geq 0,$$

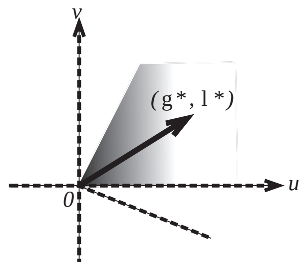


Figura 3.8: Corolario 8.e)

luego $\frac{g(x)}{-(f(x)-\mu)} \geq \gamma^* \geq 0$ del cual decimos que

$$s = \sup_{x \in S_f^-(\mu)} \frac{g(x)}{f(x)-\mu} \leq \frac{g(x)}{f(x)-\mu} \leq -\gamma^* \leq 0.$$

Luego tenemos de la definición de s, en la anterior relación tendremos que $0 \leq -\gamma^* \leq 0$. Por tanto $\gamma^* = 0$, como queríamos.

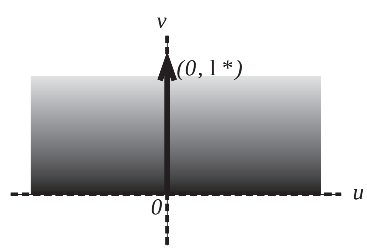


Figura 3.9: Corolario 8. f)

□

Teorema 16. Sea $K \neq \emptyset$ y $\mu \in \mathbb{R}$. Son equivalentes:

a) se cumple Dualidad Fuerte (D.F.), esto es, existen $\lambda_0^* \in P^*$ y $\gamma^* > 0$ de modo que

$$\gamma^* \inf_{x \in K} f(x) = \inf_{x \in C} L\left(\gamma^*, \lambda_0^*, x\right),$$

b) $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ es pointed y $[S_g^+(0) \cap S_f^-(\mu) = \emptyset$ ó $S_g^+(0) \cap S_f^-(\mu) \neq \emptyset, s < 0]$.

Demostración. Mostremos primero que a) implica b). El pointed del cone $\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right)$ la obtenemos del teorema 13. Y si $S_g^+(0) \cap S_f^-(\mu) = \emptyset$ el pointed de cone $\left(F(C) - (\mu, 0) + \mathbb{R}^2_{++}\right)$ es como se muestra en 3.10.

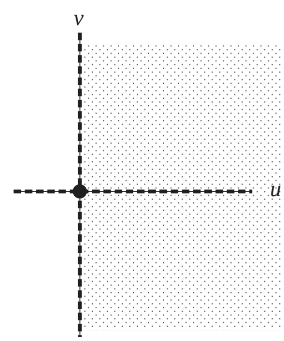


Figura 3.10: Teorema 16.

Supongamos que existe un $y \in S_g^+(0) \cap S_f^-(\mu)$. De la hipótesis a) existen $\lambda^* \in P^*$ y $\gamma^* > 0$ tal que

$$\begin{aligned}\mu &= \inf_{x \in K} f(x) \\ &= \inf_{x \in C} \frac{1}{\gamma^*} L\left(\gamma^*, \lambda^*, x\right) \\ &= \inf_{x \in C} \left\{ f(x) + \frac{\lambda^*}{\gamma^*} g(x) \right\} \\ &\leq f(x) + \frac{\lambda^*}{\gamma^*} g(x).\end{aligned}$$

Luego, de la anterior relación decimos que existe $\lambda_0^* \in P^*$ con $\lambda_0^* = \frac{\lambda^*}{\gamma^*} \geq 0$ tal que $f(x) + \lambda_0^* g(x) \geq \mu$ para todo $x \in C$. Implica que $\lambda_0^* > 0$. En efecto, si $\lambda_0^* = 0$ entonces $f(x) \geq \mu$ para todo $x \in C$, en particular para $y \in S_g^+(0) \cap S_f^-(\mu)$ lo cual es una contradicción.

Ahora, si $S_g^+(0) \cap S_f^-(\mu) \neq \emptyset$ analizamos los siguientes casos:

i) Supongamos que $s > 0$ lo cual es imposible por definición de s.

ii) Supongamos que $s = 0$ entonces existen $\overline{x} \in S_g^+(0) \cap S_f^-(\mu) \neq \emptyset$ y $\lambda_0^* > 0$ tal que $\frac{g(\overline{x})}{f(\overline{x}) - \mu} > -\frac{1}{\lambda_0^*}$, luego $\frac{g(\overline{x})}{-(f(\overline{x}) - \mu)} < \frac{1}{\lambda_0^*}$. Y de esta ultima desigualdad tenemos que $\lambda^* g(\overline{x}) < -(f(\overline{x}) - \mu)$; es decir, $f(\overline{x}) + \lambda_0^* g(\overline{x}) < \mu$ lo cual es una contradicción, pues $f(x) + \lambda_0^* g(x) \geq \mu$ para todo $x \in C$.

iii) Supongamos que $s<0$, entonces por la definición de s tendremos que $s \geq \frac{g(x)}{f(x)-\mu}$ lo cual es cierto por que $x \in S_g^+(0) \cap S_f^-(\mu)$.

Ahora mostremos *b*) implica *a*). Dado la hipótesis que cone $\left(F(C) - (\mu, 0) + \mathbb{R}^2_{++}\right)$ es pointed y $S_g^+(0) \cap S_f^-(\mu) = \emptyset$, por el corolario 8.b) y teorema 15 a3) implica a), tomemos $(\gamma^*, \lambda^*) \in \mathbb{R}^2_+ \setminus \{(0,0)\}$ y $x \in C$ tal que $\gamma^*(f(x)-\mu) + \lambda^* g(x) \geq 0$. Luego por la definición de K concluimos diciendo que

$$\inf_{x\in C} L\left(\gamma^*, \lambda^*, x\right) = \gamma^* \inf_{x\in K} f(x).$$

Ahora de la segunda hipótesis, además del pointed tenemos que $S_g^+(0) \cap S_f^-(\mu) = \emptyset$ con $s<0$. Y por el corolario 8.d), sea λ^* tal que $-\frac{1}{s} \leq \lambda^*$ verifica $f(x) + \lambda^* g(x) \geq \mu$ para todo $x \in C$. Luego existe $\gamma^* > 0$ tal que

$$\frac{\gamma^*\left(f(x)-\mu\right) + \lambda^*\gamma^* g(x)}{\gamma^*} \geq 0,$$

es decir $\gamma^*\left(f(x)-\mu\right) + \lambda_0^* g(x) \geq 0$, donde $\lambda_0^* = \lambda^*\gamma^* \geq 0$, para todo $x \in C$. Entonces por la definición de K concluimos que

$$\gamma^* \inf_{x\in K} f(x) = \inf_{x\in C} L\left(\gamma^*, \lambda_0^*, x\right).$$

$\square$

Ejemplo 7. Tomemos $C = \{(x_1,x_2) : x_2 \leq x_1, x_1 \geq 0, x_2 \geq -1\}$ $P = \mathbb{R}_+$, $f(x_1,x_2) = x_1^2 + x_1x_2 - 2x_2^2$, $g(x_1,x_2) = x_1 - x_2$. Así, $K = \{(x_1,x_2) : x_1 = x_2, x_2 \geq 0\}$, notemos además que no hay x en C tal que $g(x) < 0$ es decir, $S_g^-(0) = \emptyset$.

En éste caso cone $\left(F(C) + \mathbb{R}^2_{++}\right) = \left\{(u,v) \in \mathbb{R}^2 : v > -\frac{1}{2}u, v > 0\right\} \cup \{(0,0)\}$ es pointed; $\mu = 0$, $\operatorname{argmin}_K f = K$, $S_g^+(0) = \{(x_1,x_2) \in C : x_1 > x_2\}$ y

$$S_f^-(\mu) = \left\{(x_1,x_2) : x_2 < -\frac{1}{2}x_1, x_2 \geq -1, x_1 \geq 0\right\}.$$

Así, $S_g^+(0) \cap S_f^-(\mu) = S_f^-(\mu)$, y entonces $s = -\frac{1}{2} < 0$. De acuerdo al corolario 4,2d) cualquier λ^* tal que $-\frac{1}{s} \leq \lambda^*$ lo cual implica que $\lambda^* \geq 2$ y satisface que $f(x) + \lambda^* g(x) \geq 0$ para todo $x \in C$ es decir,

$$\min_{x\in C}\left(f(x) + \lambda^* g(x)\right) = \min_{g(x)\leq 0} f(x).$$

Ejemplo 8. Ahora veamos un contraejemplo del teorema.

Sean $C = \mathbb{R}$, $P = \mathbb{R}_+$, $f(x) = -e^{-x^2}$. Consideremos el siguiente problema primal

$$\begin{aligned} \min \quad & -e^{-x^2} \\ \text{s.a.} \quad & x \leq -1 \end{aligned} \tag{Primal}$$

cuya solución óptima está en $\overline{x} = -1$ y $\text{mín}(f(-1)) = -\frac{1}{e}$.

Ahora, el dual lagrangeano asociado a λ^* de éste problema primal y en acuerdo a nuestro contexto es

$$\begin{array}{ll} \min & [F(x) = -f(x) + \lambda^*(x+1)] \\ x \in \mathbb{R} & \end{array} \tag{Dual}$$

Luego $F'(x) = e^{-x^2}(-2x) + \lambda^* = 0$ entonces $\lambda^* = 2xe^{-x^2}$ del cual decimos que $x > 0$ pues $\lambda^* \geq 0$. Luego, reemplazando λ^* en $F(x)$ tendremos que $F(x) = e^{-x^2}\left[2x^2 + 2x - 1\right]$ y su derivada es $F'(x) = e^{-x^2}\left[-4x^3 - 4x^2 + 6x + 2\right] = 0$ y sus puntos críticos son $-1 + \frac{\sqrt{2}}{2}$, $-1 - \frac{\sqrt{2}}{2}$ y 1, del cuál tendremos que $F(x)$ no tiene mínimo, es decir

$$\text{mín}_{x \leq -1} f(x) \neq \text{mín}_{x \in \mathbb{R}} L\left(1, \lambda^*, x\right)$$

Lo cual implica que para $x \in \mathbb{R}$ se tendrá que el

$$\text{cone}\left(F(C) - \mu(1,0) + \mathbb{R}^2_{++}\right) = \mathbb{R}^2$$

no es pointed.

Bibliografía

[1] Alfred Auslender, Marc Teboulle, Asymptotic Cones and Functions in Optimization and Variational Inequalities, Ed. Springer, 2003.

[2] Bazaraa, M.S. y C.M. Shetty. Non-linear Programming Theory and algorithms, Ed. John Wiley, 1979.

[3] Dimitri P. Bertsekas, Convex analysis and optimization, Athena Scientific, 2003.

[4] F. Flores-Bazán, Fernando Flores-Bazán, Cristian Vera, A complete characterization of strong duality in onconvex optimization with a single constraint, J. Global Optim., DOI 10.1007/s10898-011-9673-6, 2011.

[5] F. Flores-Bazán, Fernando Flores-Bazán, Cristián Vera, Gordan-type alternative theorems and vector optimization revisited in Recent Developments in Vector Optimization, Q. H. Ansari and J. C. Yao (Eds), Springer-Verlag, Berlin, Vol 1, 2959, 2012.

[6] Fabián Flores-Bazán, Optimización y cálculo de variaciones sin convexidad: Una introducción, IMCA, 1998.

[7] Gabriel Osvaldo Carcamo Aravena, Dualidad Fuerte en Optimización No Convexa, Universidad de Concepción, Tesis, 2012.

[8] Javier Márquez Diez-Canedo, Fundamentos de teoria de optimización, Ed. Limusa, 1987.

[9] Ji-Ming Peng and Ya-Xiang Yuan, Optimality Conditions for the Minimization of a quadratic with two quadratic constraints, SIAM, J. Optim., 579594, 1997.

[10] José Miguel Manzano Prego, Geometría de convexos.

[11] R.I. Bot, G. Wanka, An alternative formulation for a new closed cone constraint qualification, Nonlinear Analysis, 13671381, 2006.

[12] Rockafellar R.T., Convex Análysis, Princeton University Press, 1970.

[13] V. Jeyakumar, N.Q. Huy, G, Y. LI, Necessary and suficient conditions for S-lemma and nonconvex quadratic optimization, Optim. Eng., 491503, 2009.

[14] X. J. Zheng, X. L. Sun, D. Li, Y. F. Xu, On zero duality gap in nonconvex quadratic programming problems, J. Global Optim., (2011), 229242, 52(2)2012.