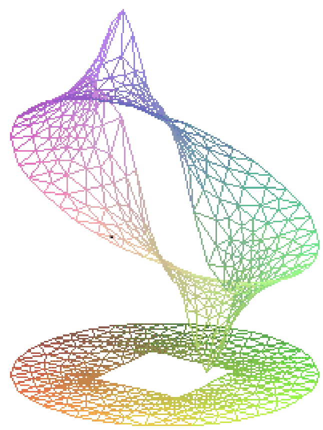
Universidad de Concepción
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática



Dualidad Fuerte en Optimización No Convexa

Tesis para optar al título de
Ingeniero Matemático

Gabriel Osvaldo Cárcamo Aravena.

Concepción, 21 de Marzo 2012

A mi Familia ...

Agradecimientos

Quisiera expresar mis más sinceros agradecimientos a las personas que me apoyaron para que este trabajo se concretara. En primer lugar a mi familia, sostén incondicional en mi vida, la cual a través de su crianza me brindaron el apoyo y amor suficiente para poder estar hoy en esta etapa de mi vida.

Agradecer a CONYCYT por haber ayudado a financiar este trabajo a través del proyecto FONDECYT 110-0667, cuyo investigador responsable es mi profesor guía Fabián Flores Bazán.

Quisiera agradecer también al Departamento de Ingeniería Matemática, en particular a los profesores Mauricio Sepúlveda, Walter Gomez, Carlos Mora, Juan Aracena, Anahí Gajardo, Gabriel Gatica, Marcelo Sobottka, Juan Molina, quienes me han formado y aconsejado en los diferentes pasajes de mi carrera.

Párrafo aparte para mi profesor guía, profesor Fabián Flores Bazán, quien me ha apoyado, formado y sobre todo a mantenido una enorme paciencia, siempre estaré agradecido por sus consejos y comentarios en mi formación académica.

Índice general

1. Sumario **5**

2. Introducción **6**
2.1. Introducción al problema. 6
2.2. Objetivos 8

3. Definiciones y resultados previos. **9**
3.1. Análisis Convexo: Conjuntos convexos, conos pointed y otros. 9
3.2. Propiedades 11
3.3. Funciones y conos asintóticos. 12

4. Strong Duality **16**
4.1. El caso general 16
4.2. Caso general con una restricción: Formulación del problema 20

5. Strong Duality v/s condiciones de optimalidad: Caso Cuadrático con m restricciones. **30**
5.1. Formas cuadráticas. 30
5.2. Condiciones de optimalidad. 32
5.3. Lagrangiano regularizado para problemas cuadráticos con restricciones lineales . . . 40

6. Strong Duality v/s condiciones de optimalidad: Caso cuadrático con un restricción cuadrática y una lineal **47**
6.1. Convexidad de imagen de formas cuadráticas y teoremas alternativos 47
6.2. Caso cuadrático no-homogéneo con restricción lineal y cuadrática 54
6.3. Existencia de mínimos para un problema cuadrático. 68

7. Conclusiones y perspectivas **73**

Capítulo 1

Sumario

La presente memoria de Ingeniero Matemático se adentra en la problemática de hallar condiciones que permitan caracterizar, en algún sentido, la propiedad de Strong Duality, para así establecer condiciones de optimalidad en soluciones de un problema de optimización con una restricción cuadrática y otra lineal.

Siempre que se está planteando un problema de optimización cabe preguntarse si existe otro problema asociado al anterior que permita, entre otras cosas, resolver el primero en forma más sencilla, aprovechando las propiedades que el segundo tiene, como por ejemplo, pasar de un problema con restricciones a uno sin restricciones, estos, son conocidos como los problemas primal y dual. La propiedad de Strong Duality o Dualidad fuerte es la condición que afirma que el dual de un problema tiene solución y que este coincide con el valor optimal del problema primal.

Estructuramos el trabajo en la siguiente forma: en los capítulos 2 y 3 marco teórico, introducimos definiciones y propiedades a la formulación del problema primal y dual en un contexto general del punto de vista matemático. Se dan algunas referencias a investigaciones anteriores para contextualizar los distintos enfoques en los cuales a sido abordado este problema. En el capítulo 4 nos enfocamos a estudiar la propiedad de Strong Duality de acuerdo al enfoque dado en [5], en donde interesa el resultado, con respecto a la descripción del conjunto $\text{cone}(F(C) - (\mu, 0) + \mathbb{R}^2_{++})$ y a la caracterización de Strong Duality para el caso general con una restricción. En el capítulo 5, estudiamos la propiedad Strong Duality y su relación con las condiciones de optimalidad, para esto, se introduce el Lagrangiano regularizado. El capítulo 6 es en donde se presentan las novedades para un problema cuadrático con una restricción cuadrática y otra lineal. Por una parte aseguramos que siempre el conjunto $\text{cone}(F(C) - (\mu, 0) + \mathbb{R}^2_{++})$ es convexo. Entregamos por separado condiciones necesarias y suficiente para que se cumpla Strong Duality, lo cual a su vez, genera nuevas condiciones de optimalidad para el caso no-convexo. Para finalizar damos un resultado que caracteriza la existencia de un mínimo para una función cuadrática, sobre un subespacio afín.

Capítulo 2

Introducción

2.1. Introducción al problema.

Motivado por la teoría de dualidad bién desarrollada en optimización lineal, planteamos el problema de encontrar condiciones (ojalá optimales en algun sentido) a fin que el valor óptimo del problema de minimización primal (**P**) coincida con el valor óptimo del problema (Lagrangiano) dual (**D**) y que tal problema admita solución.

Una de las ventajas de la propiedad anterior, es que el problema (**P**), generalmente con restricciones, se reduce a un problema sin restricciones, el cual es mucho más tratable desde el punto de vista algorítmico.

Para ser más precisos, formulamos el problema matemáticamente. Sea X un espacio vectorial topológico localmente convexo; Y espacio normado; $P \subseteq Y$ cono, convexo, cerrado y C subconjunto de X. Dado $f : C \to \mathbb{R}$ y $g : C \to Y$, se considera el problema de minimización siguiente:

$$\mu \doteq \inf_{\substack{g(x) \in -P \\ x \in C}} f(x), \tag{P}$$

El problema dual Lagrangiano asociado a P se define como:

$$\nu \doteq \sup_{\lambda^* \in P^*} \inf_{x \in C} \left[f(x) + \langle \lambda^*, g(x) \rangle \right], \tag{D}$$

Donde P^* es el cono polar no-negativo. Se dice que el problema P tiene la propiedad Z.D.G (*zero duality gap*) si los valores óptimos de P y D coinciden, esto es, $\mu = \nu$. El problema P se dice que tiene la propiedad de dualidad fuerte (*strong duality*) si P tiene la propiedad Z.D.G y además el problema D admite solución.

El problema de encontrar condiciones bajo las cuales se tiene la dualidad fuerte no es nuevo; caracterizar dicha propiedad sí es un tema contingente, más aun si no se impone la hipótesis de convexidad estándar.

Con el propósito de contextualizar el proyecto de tesis, referenciamos algunos trabajos previos de cierta trascendencia para nuestro propósito. La mayoría de ellos requieren alguna restricción de cualificación (CQ) o del tipo Slater.

1. En el trabajo [9, 2009] se considera: $C = X = \mathbb{R}^n$, $P = [0, +\infty[$, $g : \mathbb{R}^n \to \mathbb{R}$ una función cuadrática no nula, los autores demuestran que, P tiene la propiedad (*strong duality*) para cada función cuadrática $f : \mathbb{R}^n \to \mathbb{R}$ si, y sólo si se cumple la condición de Slater, ésto es, existe $x_0 \in \mathbb{R}^n$, tal que $g(x_0) < 0$.

2. En [2, 2006], se cosidera $C = X$, $g : X \to Y$ un función P-convexa y continua, se demuestra que P tiene la propiedad (*strong duality*) para cada $f \in X^*$ si, y sólo si cierta condición (CQ) se cumple, la cual involucra el epigrafo de una función soporte en C y el epigrafo del congujado de la función $x \to \langle \lambda^*, g(x) \rangle$. Esta (CQ) también es equivalente a que P tenga la propiedad (*strong duality*) para cada función continua y convexa f, tal como se cita en [7, 2008].

3. En el trabajo [10, 2009], $C = X = \mathbb{R}^n$, $P = \mathbb{R}_+ \times \{0\}$ $f : \mathbb{R}^n \to \mathbb{R}$, $g = (g_1, g_2)$, con g_2 lineal y f, g_1, cuadráticas, los autores dan condiciones de optimalidad para el problema P, el cual depende de la condición de Slater aplicado en g_1.

4. En el trabajo [14, 1997], los autores dieron condiciones que aseguren la existencia de un mínimo global, de los cuales derivan en condiciones de optimalidad del problema P, para el caso en que $C = X = \mathbb{R}^n$, $P = \mathbb{R}^2_+$ y $f, g = (g_1, g_2)$ funciones cuadráticas.

5. En [5, 2011] Se considera $C \subseteq \mathbb{R}^n$; $f : C \to \mathbb{R}$ y $g : C \to \mathbb{R}$ funciones. Los autores caracterizaron de manera completa la propiedad (*strong duality*) asociado al problema con una restricción:

$$\operatorname*{ínf}_{\substack{g(x) \leq 0 \\ x \in C}} f(x),$$

   donde las funciones tanto objetivo como de restricción no son necesariamente convexas, es más, se observa que la propiedad de (*strong duality*) se cumple aún cuando Slater no se satisfaga.

6. En [16, 2011] Se considera $C = X = \mathbb{R}^n$, $P = \mathbb{R}^m_+$, $f : \mathbb{R}^n \to \mathbb{R}$ y $g = (g_1, \ldots, g_2)$, $g_i : \mathbb{R}^n \to \mathbb{R}$ funciones cuadráticas no convexas, en este trabajo, los autores dan condiciones suficientes para que se cumpla la propiedad Z.D.G (*zero duality gap*) usando técnicas de programación semidefinda el cual requiere como hipótesis la condición de Slater.

7. En [11, 2011] Se considera $X = \mathbb{R}^n$, $C = H^{-}(d) \doteq \{x : \ Hx = d\}$, $P = \mathbb{R}^m_+$, $g = (g_1, \ldots, g_m)$, $g_i : \mathbb{R}^n :\to \mathbb{R}$ funciones cuadráticas no convexas, en este trabajo, los autores dan condiciones necesarias y suficientes para que se cumpla (*strong duality*) para cada función cuadrática $f : \mathbb{R}^n \to \mathbb{R}$ si, y sólo si se cumple la condición de Slater.

8. En [12, 2010] se considera $X = C = \mathbb{R}^n$, $P = \mathbb{R}_+$, $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$ funciones cuadráticas no convexas, en este trabajo, dan condiciones de optimalidad al problema con una restricción, en donde, se impone la hipótesis que la matriz asociada a la forma cuadrática g sea no nula, además requiere que se satisfaga Slater.

2.2. Objetivos

Contrariamente a lo que se persigue en donde dado g y C se caracteriza la validez de la propiedad de dualidad fuerte para una clase de funciones f (ver 1, 2, 3 y 4), esta tesis estará dedicada a caracterizar dicha propiedad en términos de f, g y C conjuntamente.

Debido a la importancia que reviste el caso, prueba de ello es que en la actualidad la caracterización de la propiedad de Strong duality para el caso cuadrático no homogéneo sigue siendo tema de investigación ([16, 11]), el caso particular que consideraremos en esta tesis, un problema de optimización cuadrático compuesta por dos restricciones, en el cual, la función objetivo es una forma cuadrática, una restricción cuadrática y otra lineal

$$\min_{\substack{g(x)\leq 0 \\ x\in C}} f(x) \tag{P}$$

en donde f, g, son formas cuadráticas no necesariamente convexas y $C = H^{-1}(d)$ con $H \in \mathbb{R}^{m\times n}$, $d \in \mathbb{R}^m$.

Como consecuencia de lo anterior, se pretende presentar una caracterización de optimalidad global que involucra los multiplicadores de Lagrange en el caso no convexo y cuadrático no homogéneo, ya que, apunta a la relevancia que tienen las cuadráticas en las aplicaciones, tales como se muestra en [2, 13].

De esta manera el presente trabajo de tesis se desgloza en dos objetivos:

1. Caracterizar la propiedad de (*strong duality*) asociado a (P).

2. Establecer condiciones de optimalidad en base a los resultados obtenidos en el objetivo 1. para el problema (P), particularmente el caso cuadrático no homogéneo.

Capítulo 3

Definiciones y resultados previos.

3.1. Análisis Convexo: Conjuntos convexos, conos pointed y otros.

Definición 3.1. *Sea $A \subseteq \mathbb{R}^n$.*

1. *A es un conjunto convexo $\Leftrightarrow \forall\ x,\ y\ \in A,\ \forall\ \lambda\ \in\ [0,1]:\ \lambda x + (1-\lambda)y\ \in\ A$.*

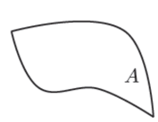
2. *Se define como $\mathrm{co}(A)$ la envoltura convexa de A, al menor conjunto convexo que contiene a A, donde:*

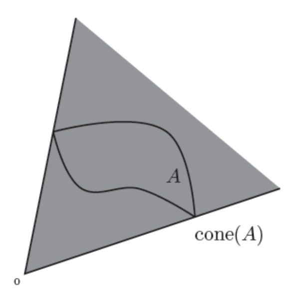
$$\mathrm{co}(A) \doteq \left\{ \sum_{i=1}^{n} \lambda_i x_i :\ \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \geq 0, \forall i = 1, \ldots, n;\ x_i \in A;\ n \in \mathbb{N} \right\}$$

3. *P es un cono $\Leftrightarrow tP \subseteq P,\ \forall t \geq 0$.*

4. *Se define $\mathrm{cone}(A)$, al menor cono que contiene a A, esto es:*

$$\mathrm{cone}(A) \doteq \bigcup_{t \geq 0} tA$$





5. $\text{cone}_+(A) = \bigcup_{t>0} tA$. *Observemos que:*

$$\text{cone}(A) = \text{cone}_+(A) \cup \{0\}$$

6. *La clausura topológica de A la denotaremos como $\overline{A}$.*

7. *El interior topológico de A lo denotaremos como* int A.

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

$$P^* \doteq \{p^* \in \mathbb{R}^n : \langle p^*, p \rangle \geq 0 \ \forall \ p \in P\},$$

   en donde $\langle \cdot, \cdot \rangle$ es el producto interno en $\mathbb{R}^n$.

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

$$M^\perp \doteq \{x \in \mathbb{R}^m : \langle x, m \rangle = 0, \ \forall \ m \in M\}.$$

Una propiedad útil aplicada posteriormente en los conos, corresponde a la definición de Pointed.

Definición 3.2. *Un cono $P \subset \mathbb{R}^n$ es pointed si y sólo 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 este es a su vez convexo, es decir, si $P \subset Y$, cono, convexo.

Lema 3.1. *Sea $K \subseteq \mathbb{R}^n$ un cono, convexo. Entonces:*

$$K \text{ es pointed} \iff K \cap (-K) = \{0\}$$

Demostración. $\Rightarrow \rfloor$ Razonando por contradicción, suponemos que existe $y \neq 0$, $y \in K \cap -K$, sigue que $y, \ -y \in K$. Como K es pointed entonces $y = 0$ lo cual contradice el supuesto. Por lo tanto $K \cap (-K) = \{0\}$.

$\Leftarrow\rfloor$ Supongamos que $K \cap (-K) = \{0\}$ y que existen $y_i \in K,\ i = 1, \ldots, m,\ m \in \mathbb{N}$, de modo que $\sum_{i=1}^{m} y_i = 0$, supongamos además que $\exists i_0 \in \{1, \ldots, n\} :\ y_{i_0} \neq 0$. Como K es cono y convexo se tiene,

$$y \doteq \frac{1}{m-1} y_{i_0} = -\frac{1}{m-1} \sum_{\substack{j=1 \\ j \neq i_0}}^{m} y_j,$$

lo cual demuestra que $y \in K \cap (-K)$, hecho por el cual se genera una contradicción pues $y \neq 0$. Por lo tanto $y_i = 0$, para todo $i = 1, \ldots, m$, es decir, K es pointed.

$\square$

A continuación algunas propiedades importantes, que seran usadas en los capitulos posteriores.

3.2. Propiedades

Proposición 3.1. [4, Proposición 1] *Sea $A,\ K \subseteq \mathbb{R}^n$ un conjunto no vacío.*

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

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

c) $\text{cone}(\text{co}(A)) = \text{co}(\text{cone(A)})$; $\text{cone}_+(\text{co}(A)) = \text{co}(\text{cone}_+(\text{A}))$*;*

d) $\text{co}(A + K) = \text{co}(A) + K$, *cuando K es convexo;*

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

f) $\overline{A + K} = \overline{A + \overline{K}}$;

g) $K \subseteq \overline{\text{cone}}(A + K)$, *cuando K es un cono;*

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

$$\text{cone}(A + K) = \text{cone}(A) + K;$$

Demostración. a), b), c), d) y e) son directas.

f): $\subseteq \rfloor$ Dado que $K \subseteq \overline{K}$, entonces $\overline{A+K} \subseteq \overline{A+\overline{K}}$.

$\supseteq \rfloor$ Sea $a+k \in A+\overline{K}$, esto es, $a \in A$ y existe $k_n \in K$ para cada $n \in \mathbb{N}$ de modo que $a+k_n \longrightarrow a+k$ cuando $n \to +\infty$. Lo cual demuestra que $a+k$ tambien pertenece a $\overline{A+K}$.

g): Para todo $x \in K$ y para cada elemento $a \in A$, $\frac{1}{n}(a+nx) \longrightarrow x$ cuando $n \to +\infty$. Como K es un cono, entonces para cada $n \in \mathbb{N}$ se tiene que $\frac{1}{n}(a+nx) \in \operatorname{cone}(A+K)$. Lo cual demuestra la inclusión.

h): La primera inclusión es obvia.

Para la segunda inclusión, sea $x = t_1 a + t_2 k \in \operatorname{cone}(A) + K$, en donde $a \in A$, $k \in K$, $t_i \geq 0, \ i = 1, 2$.

i) $t_1 = t_2 = 0$ entonces la inclusión se cumple de manera trivial.

ii) $t_1 > 0$ y $t_2 = 0$, entonces $x = t_1(a+0) \in \overline{\operatorname{cone}}(A+K)$.

iii) $t_1 = 0$ y $t_2 > 0$, entonces existe $x_n \doteq \frac{1}{n}(a+nt_2k) \in \operatorname{cone}(A+K)$, tal que $x_n \longrightarrow x$ cuando $n \to +\infty$. Por consiguiente se cumple la inclusión.

iv) $t_1, \ t_2 > 0$, entonces $x \in \operatorname{cone}_+(A)+K$ y de acuerdo con e) $\operatorname{cone}_+(A)+K = \operatorname{cone}_+(A+K) \subseteq \overline{\operatorname{cone}}(A+K)$.

Si $0 \in A$, entonces $\operatorname{cone}(A) = \bigcup_{t>0} tA$. Por lo tanto dado un elemento $ta+k \in \operatorname{cone}(A)+K$ con $t > 0$ se tiene que $ta+k = t(a+\frac{1}{t}k) \in \operatorname{cone}(A+K)$ y así la inclusión necesaria para la igualdad.

□

3.3. Funciones y conos asintóticos.

Definición 3.3. *Sea* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$.

1. *Se dice que* f *es convexa si y sólo si* $\operatorname{dom} f$ *es convexo y si* $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \ \forall \ x, \ y \in \operatorname{dom} f, \ \forall \ \lambda \in [0,1]$.

2. *Se define el conjunto epígrafo de* f*, denotado como* epi f*, al conjunto:*

$$\operatorname{epi} f \doteq \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \ f(x) \leq t\}.$$

3. *Se dice que f es semicontinua inferior (s.c.i) en un punto x, si:*

$$f(x) \leq \liminf_{k \to +\infty} f(x_k), \ [\forall \ (x_k) \ \ \textit{que converge a } x].$$

4. *Se dice que f semi-estricta casi convexa (s.c.c), si para todo par $u, \ v \ \in \ \mathbb{R}^n, \ f(u) \neq f(v)$, se cumple $f(z) < \max\{f(u), f(v)\}$ para todo $z \ \in \ ]u, v[$.*

5. *Se dice que f es cuasiconvexa (c.c), si cada subnivel de f es convexo, o equivalentemente, si $f(tx + (1-t)y) \leq \max\{f(x), f(y)\}$ para todo $x, \ , y \ \in \ \mathbb{R}^n$ y para todo $t \ \in \ [0,1]$, o equivalentemente, $\{x \ \in \ \mathbb{R}^n : \ f(x) < \lambda\}$ es convexa para todo $\lambda \ \in \ \mathbb{R}$.*

Observación 3.1. [3]

1. *Cuando f es (s.c.i) y (s.c.c) se puede probar que tambien es (c.c).*

2. *f es (s.c.i) es equivalente a decir que todo subnivel de f es cerrado.*

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

$$C^\infty \doteq \{v : \ \exists \ t_k \ \downarrow \ 0, \ \exists \ x_k \ \in \ C, \ t_k x_k \to v\}.$$

Proposición 3.2. [3, Proposición 3.2] *Propiedades de los conos asintóticos.*

a) *C^∞ es un cono convexo.*

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

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

d) *$C^\infty = (C + x_0)^\infty$ para todo $x_0 \ \in \ \mathbb{R}^n$.*

e) *$C^\infty = (\overline{C})^\infty$.*

f) *$C_1 \subseteq C_2$ implica $(C_1)^\infty \subseteq (C_2)^\infty$.*

g) *Si C es no vacío, cerrado y convexo, $x_0 \in C$, entonces,*

$$C^\infty = \{u \in \mathbb{R}^n : \ x_0 + tu \in C \ \forall \ t > 0\}.$$

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

$$\left(\bigcup_{i=1}^{m} C_i\right)^\infty = \bigcup_{i=1}^{m} (C_i)^\infty.$$

i) *Sea $\{C_i\}$, $i \in I$, un 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

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

Dada cualquier función $f : C \subseteq \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, la función asintótica de f, se denota como f^∞ y verifica:

$$\text{epi } f^\infty = (\text{epi } f)^\infty.$$

Algunas propiedades.

Proposición 3.3. [3, Proposición 3.3] *Para todo $x \in \mathbb{R}^n$,*

$$f^\infty(x) = \text{lím}_{t\to 0^+,} \, \text{ínf}_{y\to x} \, t f(\frac{y}{t});$$

$$f^\infty(x) = \text{ínf}\left[\text{lím inf}_{k\to+\infty} \, t_k f(\frac{x_k}{t_k}) : \ t_k \downarrow 0, \ x_k \to x\right].$$

Observación 3.2. *En el caso en que f es convexa se cumple para todo x,*

$$f^{\infty}(x) = \sup_{\lambda>0} \frac{f(x_0 + \lambda x) - f(x_0)}{\lambda}, \ x_0 \ \in \ \text{dom } f.$$

Proposición 3.4. [3, Proposición 3.5] *Sea $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, entonces:*

a) *Sí x es un mínimo para f entonces f es semicontinua inferior en x.*

b) *Sí $\min\limits_{\mathbb{R}^n} f(x)$ existe, entonces $f^{\infty}(x) \geq 0$ para todo $x \ \in \ \mathbb{R}^n$.*

c) *Sí $f^{\infty}(x) > 0$ para todo $x \neq 0$, entonces f es coerciva. Adicionalmente se cumple la implicación contraria si f es convexa.*

d) *Sí $C \subseteq \mathbb{R}^n$, es no vacío, convexo con* dom $f \cap C \neq \emptyset$ y $\widetilde{f}(x) = f(x) + I_C(x)$, $x \ \in \ \mathbb{R}^n$ *(I_C denota la función indicadora del conjunto C), entonces:*

$$\{x \ \in \ \mathbb{R}^n : \ \widetilde{f}(x) \leq \lambda\} = \{x \ \in \ \mathbb{R}^n : \ f(x) \leq \lambda\}, \ \textit{para todo} \ \lambda \ \in \ \mathbb{R}^n$$

y

$$\widetilde{f}^{\infty}(x) \geq f^{\infty}(x) + I_{C^{\infty}}, \ x \ \in \ \mathbb{R}^n.$$

e) *Sí $f^{\infty}(x) > 0$ para todo $x \ \in \ C^{\infty}$, $x \neq 0$, entonces $\widetilde{f}$ es coerciva.*

Capítulo 4

Strong Duality

En este capítulo se profundiza el concepto de Dualidad fuerte desde la temática general hasta llegar al caso particular cuadrático, en donde se relaciona esta propiedad con las soluciones del problema (P) lo cual se conoce como condiciones de optimalidad.

4.1. El caso general

Sea $C \subseteq \mathbb{R}^n$ no vacío, la función $f : C \to \mathbb{R}$. Consideremos el siguiente problema de minimización,

$$\mu \doteq \inf_{x \in K} f(x), \tag{P}$$

donde $g : C \to \mathbb{R}^m$, $K \doteq \{x \in C : g(x) \in -P\}$ con $P \subseteq \mathbb{R}^m$ cono convexo. Cabe notar que como el interior topológico de P puede ser vacío, en este caso, el conjunto de restricciones es descrito por desigualdades e igualdades.

Se introduce de el Lagrangiano:

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

De la definición de K, notemos que, para todo $x \in K$, $\gamma^* f(x) \geq \gamma^* f(x) + \langle \lambda^*, g(x) \rangle$. Por lo tanto se obtiene de manera directa la desigualdad,

$$\gamma^* \inf_{x \in K} f(x) \geq \inf_{x \in K} L(\gamma^*, \lambda^*, x) \geq \inf_{x \in C} L(\gamma^*, \lambda^*, x), \ \forall \ \lambda^* \in P^*, \ \forall \ \gamma^* \geq 0 \tag{4.1}$$

la propiedad de Strong Duality 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 tal que

$$\gamma^* \inf_{x \in K} f(x) = \inf_{x \in C} L(\gamma^*, \lambda_0^*, x) \tag{4.2}$$

En el trabajo [5] se obtuvieron condiciones que permitan obtener la igualdad en (4.2), en donde la condición de pointed juega un rol crucial en la caracterización de Strong Duality. A continuación algunos resultados formulados en [5] y presentados en detalle.

Teorema 4.1. [5, Teorema 3.1] *Consideremos el problema* (P)*, en donde* $F(C) \doteq \{(f(x), g(x)) : x \in C\}$. *Suponiendo que* $\mu \in \mathbb{R}$ *y*

$$\mathrm{int}(\mathrm{co}(F(C)) + (\mathbb{R}_+ \times P)) \neq \emptyset.$$

Las siguientes afirmaciones son equivalentes:

a) *Existe multipiclador de Lagrange* $(\gamma_0^*, \lambda_0^*) \in \mathbb{R}_+ \times P^*,\ (\gamma_0^*, \lambda_0^*) \neq (0,0)$ *tal que*

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

b) $\mathrm{cone}\,(\mathrm{int}(\mathrm{co}(F(C)) - \mu(1,0) + (\mathbb{R}_+ \times P)))$ *es pointed;*

c) $(0,0) \notin \mathrm{int}(\mathrm{co}(F(C)) - \mu(1,0) + (\mathbb{R}_+ \times P))$;

d) $\mathrm{cone}\,(\mathrm{co}(F(C)) - \mu(1,0) + \mathrm{int}(\mathbb{R}_+ \times P))$ *es pointed, en el caso* $\mathrm{int}\ P \neq \emptyset$.

Demostración. $b) \Rightarrow c)\rfloor$ Es directo, ya que, en caso contrario,

$$(0,0) \in \mathrm{int}(\mathrm{co}(F(C)) - \mu(1,0) + (\mathbb{R}_+ \times P)) \Rightarrow \mathrm{Cone}\,(\mathrm{int}(\mathrm{co}(F(C)) - \mu(1,0) + (\mathbb{R}_+ \times P))) = \mathbb{R} \times \mathbb{R}^m,$$

lo cual contradice el hecho que $\mathrm{cone}\,(\mathrm{int}(\mathrm{co}(F(C)) - \mu(1,0) + (\mathbb{R}_+ \times P)))$ sea pointed.

$c) \Rightarrow a)\rfloor$ Usando argumentos de separación entre convexos se garantiza la existencia $(\gamma_0^*, \lambda_0^*) \in \mathbb{R} \times P^* \setminus \{(0,0)\}$ de modo tal que

$$\gamma_0^* f(x) + \langle \lambda_0^*, g(x) \rangle \geq \gamma_0^* \mu, \ \forall \ x \ \in \ C.$$

Lo anterior implica que $\inf_{x \in C} L(\gamma_0^*, \lambda_0^*, x) \geq \gamma_0^* \mu$, en definitiva y teniendo en cuenta que la otra desigualdad se cumple por definición, que

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

a) $\Rightarrow$ *b)*$\rfloor$ Por hipótesis, (4.2) se cumple, reescribiéndolo como:

$$\langle (\gamma_0^*, \lambda_0^*), (f(x) - \mu, g(x)) \rangle \geq 0, \ \forall \ x \ \in \ C. \tag{4.4}$$

Sea $A \doteq F(C) - \mu(1,0)$. Dado que $\text{cone}(\text{int}(\text{co}(A) + (\mathbb{R}_+ \times P)))$ es convexo, queda por mostrar que cualquiera sea $x, \ -x \ \in \ \text{cone}(\text{int}(\text{co}(A) + (\mathbb{R}_+ \times P)))$, se tiene que $x = 0$. Razonando por contradicción, se asume $x \neq 0$, entonces podemos escribir $x = t_1\xi_1, \ -x = t_2\xi_2, \ t_i > 0$, $\xi_i \ \in \ \text{int}(\text{co}(A) + (\mathbb{R}_+ \times P))$, $i = 1, 2$. De la ecuación (4.4) se extiende por convexidad, para obtener que $\langle (\gamma_0^*, \lambda_0^*), \xi \rangle \geq 0 \ \forall \ \xi \ \in \ \text{co}(A) + (\mathbb{R}_+ \times P)$. Por otro lado dado cualquier $y \ \in \ \mathbb{R} \times \mathbb{R}^m$, existe $\delta > 0$ tal que:

$$\xi_i + \lambda y \ \in \ \text{co}(A) + \mathbb{R}_+ \times P, \ \forall |\lambda| < \delta, \ \forall i = 1, 2.$$

Poniendo $p^* \doteq (\gamma_0^*, \lambda_0^*)$, se tiene

$$\langle p^*, \xi_i + \lambda y \rangle \geq 0, \ \forall \ |\lambda| < \delta, \ \forall i = 1, 2.$$

Se sigue de lo anterior que $\langle p^*, \lambda y(t_1 + t_2) \rangle \geq 0$, lo cual implica que $\langle p^*, y \rangle = 0$ para todo $y \in \mathbb{R} \times \mathbb{R}^m$ y por ende $(\gamma_0^*, \lambda_0^*) = p^* \equiv 0$ lo cual es una contradicción.

b) $\Rightarrow$ *d)*$\rfloor$ Primero, dado un cono convexo Q con interior no vacío y un conjunto K. Se cumple $K + \text{int } Q = \text{int}(K + Q)$ y $\text{cone}(\text{co}(K)) = \text{co}(\text{cone}(K))$, así $\text{cone}(\text{int}(\text{co}(A) + P)) = \text{co}(\text{cone}(\text{cone}(A + \text{int } P))$. Teniendo en cuenta este hecho, la demostración es directo de la definición de pointed.

□

Para obtener la propiedad de Strong Duality se requiere asegurar que el parámetro γ_0^* sea estrictamente positivo, la cual necesita una condición del tipo Slater. El siguiente corolario caracteriza este hecho.

Corolario 4.1. [5, Corolario 3.2] *Consideremos el problema* (P). *Suponiendo que* $\mu \in \mathbb{R}$,

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

y se cumple la condición de Slater generalizado $\overline{\operatorname{cone}}(g(C)+P) = \mathbb{R}^m$. *Entonces son equivalentes*

a) *Existe multipiclador 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) $\operatorname{cone}\left(\operatorname{int}(\operatorname{co}(F(C)) - \mu(1,0) + \operatorname{int}(\mathbb{R}_+ \times P))\right)$ *es pointed.*

Demostración. $a) \Leftrightarrow b)]$ Una implicación es obvia. Por otro lado suponiendo a) se tiene:

$$\mu \leq \max_{\lambda^* \in P^*} \inf_{x \in C} L(\lambda^*, x),$$

esto, junto a (4.1) implica b).

$c) \Leftrightarrow a)]$ Una implicación proviene del teorema4.1. Si $\gamma_0^* = 0$, entonces $0 \neq \lambda_0^* \in P^*$ y $\langle \lambda_0^*, g(x) \rangle \geq 0$ para todo $x \in C$. Esto implica que $\lambda_0^* = 0$, por la condición de Slater generalizado, lo cual es una contradicción. Por lo tanto $\gamma_0^* > 0$ y en base a este parámetro construir uno que sea unitario.

$\square$

4.2. Caso general con una restricción: Formulación del problema

En esta sección, abarcamos Strong Duality 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 $\mathrm{cone}(F(C) - \mu(1,0) + \mathrm{int}\ \mathbb{R}^2)$ mencionado en el teorema 4.1. A raiz de esto, una nueva caracterización de Strong Duality, la cual entre otras, evidencia la no dependencia de la condición de Slater.

En este caso al tratarse de una restricción, el problema (P) queda,

$$\mu = \inf_{\substack{g(x)\leq 0 \\ x\in C}} f(x),$$

por otro lado el Lagrangiano usual asociado a (P) es,

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

donde $\gamma^* \geq 0$ y $\lambda^* \geq 0$ se conocen como multiplicadores de Lagrange. Notemos además que $K \doteq \{x \ \in \ C : \ g(x) \leq 0\}$ y de esta manera la desigualdad que permite obtener la igualdad (4.2) para el caso con una restricción es,

$$\gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall \ x \ \in \ C \tag{4.3}$$

La cual implica Strong Duality, cuando $\gamma^* > 0$.

Dados $\gamma^* \geq 0, \ \lambda^* \geq 0$, considerando la anterior desigualdad (4.3) y haciendo $F(C) \doteq \{(f(x), g(x)) : \ x \ \in \ C\}$, con $\rho \doteq (\gamma^*, \lambda^*)$, esta se puede escribir como

$$\langle \rho, a \rangle \geq 0 \ \forall \ a \ \in \ F(C) - \mu(1,0),$$

la que a su vez, es equivalente a

$$\langle \rho, a \rangle \geq 0 \ \forall \ a \ \in \ F(C) - \mu(1,0) + \mathbb{R}^2_+.$$

El siguiente resultado el cual es dado en [4], entrega un completa caracterización de la relación anterior.

Teorema 4.2. [4, Teorema 5] *Sea $P \subseteq \mathbb{R}^2$ un cono, convexo, cerrado de modo tal que* $\operatorname{int} P \neq \emptyset$ *y* $A \subseteq \mathbb{R}^2$ *un conjunto no vacío. Entonces las siguiente proposiciones son equivalentes.*

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

b) $A \cap -\operatorname{int} P = \emptyset$ *y* $\overline{\operatorname{cone}}(A + P)$ *es convexo;*

c) $A \cap -\operatorname{int} P = \emptyset$ *y* $\operatorname{cone}_+(A + \operatorname{int} P)$ *es convexo;*

d) $A \cap -\operatorname{int} P = \emptyset$ *y* $\operatorname{cone}(A + \operatorname{int} P)$ *es convexo;*

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

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

Demostración. $c) \Rightarrow d) \Rightarrow b)$ y $f) \Rightarrow a)\rfloor$ Es directo.

$b) \Rightarrow a)\rfloor$ Es una consecuencia de argumentos de separación de convexos debido a que

$$A \cap (-\operatorname{int} P) = \emptyset \Leftrightarrow \overline{\operatorname{cone}(A + P)} \cap (-\operatorname{int} P) = \emptyset.$$

$a) \Rightarrow b)\rfloor$ Claramente $\langle \lambda^*, a\rangle \geq 0$, para todo $x \in \operatorname{cone}(A + P)$. Sean $u \in \operatorname{int} P$, $y,\ z \in A$. Entonces obviamente

$$\operatorname{cone}(\{y\}) + \operatorname{cone}(\{u\}) = \{sy + tu : \ s,\ t \geq 0\}$$

es un cono convexo cerrado que contiene a u e y y esta contenido en $\operatorname{cone}(A + P)$. Lo mismo es cierto para el cono $\operatorname{cone}(\{z\}) + \operatorname{cone}(\{u\})$. Ambos conos tienen en común a $\operatorname{cone}(\{u\})$ y la unión esta contenida en $\operatorname{cone}(\{A + P\})$, el cual a su vez esta contenido en $\{x \in \mathbb{R}^2 : \ \langle \lambda^*, x\rangle \geq 0\}$. Definiendo $B \doteq \operatorname{cone}(\{y\}) + \operatorname{cone}(\{u\}) \cup \operatorname{cone}(\{z\}) + \operatorname{cone}(\{u\})$, se tiene que B es un cono convexo. Dados $x,\ y \in B$ se deduce que $[y, z] \subseteq B \subseteq \operatorname{cone}(A + P)$. Sigue que $\operatorname{co}(A) \subseteq \overline{\operatorname{cone}}(A + P)$, con lo que se infiere que $\overline{\operatorname{cone}}(A + P)$ es convexo, pues $P \subseteq \overline{\operatorname{cone}}(A + P)$.

$b) \Leftrightarrow c)\rfloor$ Obiamente $c) \Rightarrow b)$. Si $\overline{\operatorname{cone}}(A + P)$ es convexo, entonces

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

es tambien convexo.

$c) \Rightarrow e)\rfloor$ Sean $y,\ -y\ \in\ \mathrm{cone}(A+\mathrm{int}\ P)$. Razonando por contradicción supongamos que $y \neq 0$. Por hipótesis, $\mathrm{cone}_+(A+\mathrm{int}\ P)$ es convexo, luego $0 = \frac{1}{2}y + \frac{1}{2}(-y)\ \in\ \mathrm{cone}_+(A+\mathrm{int}\ P)$, lo cual implica que existen $t > 0$, $a\ \in\ A$ y $p\ \in\ \mathrm{int}\ P$ tales que $0 = t(a+p)$. Luego $a = -p$ implica que $A \cap -\mathrm{int}\ P \neq \emptyset$ lo cual es una contradicción. Así, del lema 3.1, el conjunto $\mathrm{cone}(A+\mathrm{int}\ P)$ es pointed.

$e) \Rightarrow f)\rfloor$ Asumiendo por el contrario que $\mathrm{co}(A) \cap -\mathrm{int}\ P \neq \emptyset$. Entonces, existe $a_i\ \in\ A$, $p_0\ \in\ \mathrm{int}\ P$, $\alpha_i \geq 0$, $\displaystyle\sum_{i=1}^{m} \alpha_i = 1$, de tal manera que $0 = \displaystyle\sum_{i=1}^{m} \alpha_i a_i + p_0 = \sum_{i=1}^{m} \alpha_i(a_i + p_0)$. Como $\mathrm{cone}(A+\mathrm{int}\ P)$ es pointed, se sigue que $\alpha_i(a_i+p_0) = 0$ para todo $i = 1,\ldots,m$. Por lo tanto $a_j + p_0 = 0$ para algún j. Sin embargo esto implica que $\mathrm{cone}(A+\mathrm{int}\ P) = \mathbb{R}^2$ lo cual contradice el hecho que sea pointed.

$\square$

En virtud del resultado anterior y del razonamiento empleado en [5], la idea es particionar el cono $\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$, para determinar las condiciones que permitan asegurar la condición de Pointed con respecto a las particiones, se denota $\mathbb{R}^2_{++} \doteq \mathrm{int}\ \mathbb{R}^2_+$.

Por otro lado hacemos $K = \{x\ \in\ C:\ g(x) \leq 0\} = S^-_g(0) \cup S^=_g(0)$, en donde, $S^-_g(0) \doteq \{x\ \in\ C:\ g(x) < 0\}$, $S^=_g(0) \doteq \{x\ \in\ C:\ g(x) = 0\}$, $S^+_g(0) \doteq \{x\ \in\ C:\ g(x) > 0\}$. Similarmente se definen

$$S^-_f(\mu) \doteq \{x\ \in\ C:\ f(x) < \mu\},\ S^=_f(\mu) \doteq \{x\ \in\ C:\ f(x)\mu\},\ S^=_f(\mu) \doteq \{x\ \in\ C:\ f(x) = \mu\}.$$

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

$$\Omega_1 \doteq \bigcup_{x \in \mathrm{argmin}_K f \cap S^=_g(0)} [(0,0) + \mathbb{R}^2_{++}] \cup \bigcup_{x \in \mathrm{argmin}_K f \cap S^-_g(0)} [(0, g(x)) + \mathbb{R}^2_{++}]$$

$$\Omega_2 \doteq \bigcup_{x \in S^+_f(\mu) \cap S^-_g(0)} [(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}] \cup \bigcup_{x \in S^+_f(\mu) \cap S^=_g(0)} [(f(x) - \mu, 0) + \mathbb{R}^2_{++}]$$

$$\Omega_3 \doteq \bigcup_{x \in S^-_f(\mu) \cap S^+_g(0)} [(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}] \cup \bigcup_{x \in S^=_f(\mu) \cap S^+_g(0)} [(0, g(x)) + \mathbb{R}^2_{++}] \cup \bigcup_{x \in S^+_f(\mu) \cap S^+_g(0)} [(f(x) - \mu, g(x)) + \mathbb{R}^2_{++}]$$

Observación 4.1. *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

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

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

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

Proposición 4.1. *$K \neq \emptyset$ y μ finito. Entonces:*

a) $C = K \iff S_g^+(0) = \emptyset$;

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

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

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

Demostración. a) Evidente de la definición de K.

b) $\Leftarrow\rfloor$ Es directo.

$\Rightarrow\rfloor$ Por contrarecíproco, existe un elemento $x_0 \in S_g^-(0)$ tal que $f(x_0) \geq \mu$. Con lo cual se generan dos posibilidades:

Caso 1) $f(x_0) = \mu$. Como $\mu \in \mathbb{R}$ entonces $x_0 \in \operatorname{argamin}_K f$ y por lo tanto $\operatorname{argamin}_K f \cap S_g^-(0) \neq \emptyset$.

Caso 2) $f(x_0) > \mu$. Entonces $x_0 \in S_f^+(\mu)$ y con lo cual $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$.

c) Primero notemos que de la definición de μ $S_f^-(\mu) \cap S_g^-(0) = \emptyset$. Por lo tanto,

$$S_f^+(\mu) \cap S_g^-(0) = \emptyset \Leftrightarrow S_g^-(0) \subseteq [S_f^+(\mu)]^c = S_f^-(\mu) \cup \operatorname{argmin}_K f \Leftrightarrow S_g^-(0) \subseteq \operatorname{argmin}_K f.$$

d) Es directo de la definición de $S_f^-(\mu)$.

$\square$

Observación 4.2. *Notemos que para efectos prácticos, en lo que respecta a la obtención de resultados, se asume que μ es finito y que el conjunto factible K es no vacío. Notar que a raiz de esta hipótesis se verifica que*

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

Luego teniendo en cuenta el teorema 4.2, podemos afirmar que la convexidad del conjunto $\operatorname{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ *es equivalente a que sea Pointed.*

En [5], los autores mostraron bajo que condiciones el conjunto $\operatorname{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ es convexo (teorema 4.1). Además en el mismo trabajo, del corolario 4.2 se da un conjunto de solución para los multiplicadores de Lagrange, para lo cual se verifica Strong Duality. El siguiente resultado agrupa ambos resultados.

Teorema 4.3. [5, Teorema 4.1, Corolario 4.2] *Consideremos el problema* (P), $K \neq \emptyset$ *y* μ *finito.*

a) *En el caso en que* $\operatorname{argmin}_K f \neq \emptyset$ *se tiene que* $\operatorname{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ *es convexo si y sólo si, se cumple cualquiera de las siguientes condiciones:*

a.1) $[\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$ *ó* $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ r = -\infty]$ *y* $[S_g^+(0) = \emptyset$ *ó* $S_f^-(\mu) = \emptyset, S_g^+(0) \neq \emptyset]$. *Entonces*

$$\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v):\ u > 0\}$$

Por lo tanto

$$\exists\ (\gamma^*, \lambda^*)\ \in\ \mathbb{R}^2_+ \setminus \{(0,0)\}:\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\ x\ \in\ C\ \iff\ \gamma^* > 0,\ \lambda^* = 0.$$

a.2) $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, $\operatorname{argmin}_K f \cap S_g^=(0) \neq \emptyset$ *y* $K = \operatorname{argmin}_K f$. *Entonces*

$$\mathrm{cone}_+(F(C)-\mu(1,0)+\mathbb{R}^2_{++}) = \begin{cases} \mathbb{R}^2_{++} & , S_g^+(0)=\emptyset \\ \{(u,v):\ v>su,\ v>0\} & , S_f^-(\mu)\neq\emptyset,\ s<0 \\ \{(u,v):\ v>0\} & , S_f^-(\mu)\neq\emptyset,\ s=0 \end{cases}$$

Por lo tanto

a.2.1) *Si* $S_g^+(0)=\emptyset$*,*

$$\exists\ (\gamma^*,\lambda^*)\ \in\ \mathbb{R}^2_+\setminus\{(0,0)\}:\ \gamma^*(f(x)-\mu)+\lambda^* g(x)\geq 0,\ \forall\ x\ \in\ C\iff \gamma^*\geq 0,\ \lambda^*\geq 0.$$

a.2.2) *Si* $S_f^-(\mu)\neq\emptyset,\ s<0$*. Entonces*

$$\exists\ (\gamma^*,\lambda^*)\ \in\ \mathbb{R}^2_+\setminus\{(0,0)\}:\ \gamma^*(f(x)-\mu)+\lambda^* g(x)\geq 0,\ \forall\ x\ \in\ C\iff \gamma^*>0,\ \lambda^*\geq -\frac{1}{s}\gamma^*,\ \lambda^*\neq 0.$$

a.2.3) *Si* $S_f^-(\mu)\neq\emptyset,\ s=0$*. Entonces*

$$\exists\ (\gamma^*,\lambda^*)\ \in\ \mathbb{R}^2_+\setminus\{(0,0)\}:\ \gamma^*(f(x)-\mu)+\lambda^* g(x)\geq 0,\ \forall\ x\ \in\ C\iff \gamma^*=0,\ \lambda^*>0.$$

a.3) $\mathrm{argmin}_K f\cap S_g^-(0)=\emptyset$*,* $\mathrm{argmin}_K f\cap S_g^=(0)\neq\emptyset$*,* $S_f^+(\mu)\cap S_g^-(0)\neq\emptyset$*,* $-\infty<r<0$ *y* $[S_g^+(0)=\emptyset$ *ó* $S_f^-(\mu)=\emptyset,\ S_g^+(0)\neq\emptyset]$*. Entonces*

$$\begin{aligned} \mathrm{cone}_+(F(C)-\mu(1,0)+\mathbb{R}^2_{++}) &= \mathrm{cone}_+(F(S_f^+(\mu)\cap S_g^-(0))-\mu(1,0)+\mathbb{R}^2_{++}) \\ &= \{(u,v):\ v>ru,\ u>0\} \end{aligned}$$

Por lo tanto

$$\exists\ (\gamma^*,\lambda^*)\ \in\ \mathbb{R}^2_+\setminus\{(0,0)\}:\ \gamma^*(f(x)-\mu)+\lambda^* g(x)\geq 0,\ \forall\ x\ \in\ C\iff \gamma^*>0,\ 0\leq\lambda^*\leq -\frac{1}{r}\gamma^*.$$

a.4) $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, $\operatorname{argmin}_K f \cap S_g^=(0) \neq \emptyset$, $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, $-\infty < r < 0$ *y* $S_f^-(\mu) \cap S_g^+(\mu) \neq \emptyset$, $s \leq r$.

Definiendo $A \doteq (S_f^+(\mu) \cap S_g^-(\mu)) \cup S_f^-(\mu)$, *se tiene*

$$\begin{aligned}\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \operatorname{cone}_+(F(A) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) :\ v > ru,\ v > su\}\end{aligned}$$

Por lo tanto

$$\exists\, (\gamma^*, \lambda^*) \in \mathbb{R}^2_+ \setminus \{(0,0)\} :\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\, x \in C \iff \gamma^* > 0,\ -\frac{1}{s}\gamma^* \leq \lambda^* \leq -\frac{1}{r}\gamma^*.$$

a.5) [$\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, $\operatorname{argmin}_K f \cap S_g^=(0) \neq \emptyset$ *ó* $S_f^+(\mu) \cap S_g^=(0) \neq \emptyset$.] *y* $S_f^+(\mu) \cap S_g^-(0) = S_f^-(\mu) = \emptyset$.

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

Por lo tanto

$$\exists\, (\gamma^*, \lambda^*) \in \mathbb{R}^2_+ \setminus \{(0,0)\} :\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\, x \in C \iff \gamma^* \geq 0,\ \lambda^* \geq 0.$$

b) *En el caso en que* $\operatorname{argmin}_K f = \emptyset$ *se tiene que* $\operatorname{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ *es convexo si y sólo si, se cumple cualquiera de las siguientes condiciones:*

b.1) $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, $-\infty < r < 0$ *y* $S_g^+(0) = \emptyset$ *ó* [$S_f^-(0) \neq \emptyset$, $s \leq r$], *ó* [$S_f^-(\mu) = \emptyset$, $S_g^+(0) \neq \emptyset$].

Como consecuencia,

$$\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \begin{cases} \{(u,v) :\ v > ru,\ u > 0\} & , S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ -\infty < r < 0,\ [S_g^+(0) = \emptyset \\ & \text{ó } S_f^-(\mu) = \emptyset,\ S_g^+(0) \neq \emptyset] \\ \{(u,v) :\ v > ru,\ v > su\} & , S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ S_f^-(\mu) \neq \emptyset,\ s \leq r \end{cases}$$

Por lo tanto

b.1.1) *Si* $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ -\infty < r < 0,\ [S_g^+(0) = \emptyset$ *ó* $S_f^-(\mu) = \emptyset,\ S_g^+(0) \neq \emptyset]$. *Entonces*

$$\exists\ (\gamma^*, \lambda^*)\ \in\ \mathbb{R}_+^2 \setminus \{(0,0)\} :\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\ x\ \in\ C \iff \gamma^* > 0,\ 0 \leq \lambda^* \leq -\frac{1}{r}.$$

b.1.2) *Si* $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ S_f^-(\mu) \neq \emptyset,\ s \leq r$. *Entonces*

$$\exists\ (\gamma^*, \lambda^*)\ \in\ \mathbb{R}_+^2 \setminus \{(0,0)\} :\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\ x\ \in\ C \iff \gamma^* > 0,\ -\frac{1}{s} \leq \lambda^* \leq -\frac{1}{r}.$$

b.2) $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, $r = -\infty$ *y* $S_g^+(0) = \emptyset$ *ó* $[S_f^-(\mu) = \emptyset,\ S_g^+(0) \neq \emptyset]$.

Como consecuencia,

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}_{++}^2) = \{(u,v) :\ u > 0\}$$

Por lo tanto

$$\exists\ (\gamma^*, \lambda^*)\ \in\ \mathbb{R}_+^2 \setminus \{(0,0)\} :\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\ x\ \in\ C \iff \gamma^* > 0,\ \lambda^* = 0.$$

b.3) $S_f^+(\mu) \cap S_g^-(0) = S_f^-(\mu) = \emptyset$, $S_f^+(\mu) \cap S_g^=(0) \neq \emptyset$.

$$\begin{aligned} \text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}_{++}^2) &= \text{cone}_+(F(S_f^+(\mu) \cap S_g^=(0)) - \mu(1,0) + \mathbb{R}_{++}^2) \\ &= \mathbb{R}_{++}^2 \end{aligned}$$

Por lo tanto

$$\exists\ (\gamma^*, \lambda^*)\ \in\ \mathbb{R}_+^2 \setminus \{(0,0)\} :\ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\ x\ \in\ C \iff \gamma^* \geq 0,\ \lambda^* \geq 0.$$

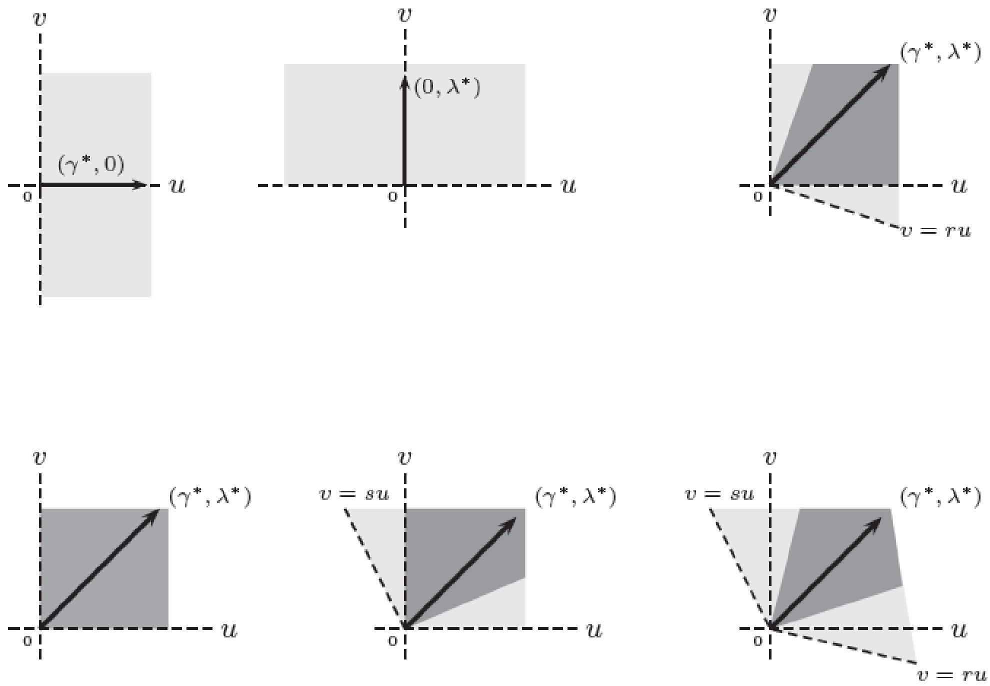


Figura 4.1: Teorema 4.3

Demostración. La demostración se remite a la visualización de la figura.

□

El siguiente resultado provisto en [5], caracteriza de manera completa la condición de Strong Duality asociado al problema (P).

Teorema 4.4. [5, Teorema 4.4] *Sea $K \neq \emptyset$ y $\mu \in \mathbb{R}$. Son equivalentes:*

a) *Se cumple Strong Duality (S.D)*

b) $\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ *es convexo y* $\left[S_f^-(\mu) = \emptyset \text{ ó } S_f^-(\mu) \neq \emptyset, s < 0\right]$

Demostración. $a) \Rightarrow b)\rfloor$ De acuerdo al teorema 4.1 y a la observación 4.2, Strong Duality implica que $\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ es convexo. Si $S_f^-(\mu) = \emptyset$ entonces no hay nada más que probar, por lo cual suponemos que $S_f^-(\mu) \neq \emptyset$. Por hipótesis existe $\lambda^* \geq 0$ de modo que $f(x) + \lambda^* g(x) \geq \mu$, $\forall\ x \in C$, además necesariamente $\lambda^* > 0$ pues en caso contrario $f(x) \geq \mu$, $\forall\ x \in C$ lo cual de acuerdo a la proposición 4.1 d), $S_f^-(\mu) = \emptyset$ lo cual daría una contradicción en este caso.

Ahora supongamos que $s = 0$. Entonces, existe $\overline{x} \in S_f^-(\mu)$ tal que

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

ya que en caso contrario se concluye que $-\dfrac{1}{\lambda^*}$ es una cota superior estrictamente negativa para s lo cual no puede ser.

De la anterior relación se concluye que $f(\overline{x}) + \lambda^* g(\overline{x}) < \mu$, lo cual es una contradicción pues estamos bajo el supuesto de Strong Duality. Por consiguiente, $s < 0$.

$b) \Rightarrow a)\rfloor$ Se sigue del teorema 4.3

$\square$

Capítulo 5

Strong Duality v/s condiciones de optimalidad: Caso Cuadrático con m restricciones.

En este capítulo se estudia la relación que existe entre la propiedad de Dualidad fuerte y las condiciones de optimalidad para problemas cuadráticos, con el objetivo de estudiar el caso propuesto por los objetivos de esta tesis así como tambien presentar nuevos resultados. También se hace un repaso a resultados de trabajos afínes.

5.1. Formas cuadráticas.

Sea M_n el espacio de las matrices reales cuadradas de tamaño $n \times n$, S^n al espacio de las matrices, reales, simétricas de orden n. Una forma cuadrática se define como una función:

$$f: \begin{array}{c} \mathbb{R}^n \longrightarrow \mathbb{R} \\ x \longmapsto f(x) \end{array} \quad , \; f(x) = \frac{1}{2} x^\top A x + a^\top x + \alpha,$$

en donde $A \in M_n$, $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. Dado que ,

$$x^\top A x = x^\top \left(\frac{A + A^\top}{2} \right) x, \; \frac{A + A^\top}{2} \in S^n,$$

es que sin perdida de generalidad podemos considerar a lo largo del trabajo que $A \in S^n$.

Definición 5.1. *Sea $A = (a_{ij}) \in S^n$ y $\lambda(A)$ los valores propios de A. Entonces diremos que A es:*

1. *Ortogonal si y sólo si:*

$$A^\top A = AA^\top = I,\ I \text{ matriz identidad};$$

2. *Simétrica si y sólo si:*

$$A = A^\top;$$

Como consecuencia, para todo $i = 1, \dots, n$ $\lambda_i(A) \in \mathbb{R}$ *y* A *es diagonalizable, es decir, existe matriz* $O \in M_n$ *ortogonal de modo tal que:*

$$O^T A O = \begin{pmatrix} \lambda_1(A) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n(A) \end{pmatrix} \doteq \operatorname{diag}(\lambda_1(A), \dots, \lambda_n(A))$$

3. *Semi-definida positiva si y sólo si:*

$$\forall x \in \mathbb{R}^n : x^\top A x \geq 0,\ (A \succeq 0),$$

equivalentemente, para todo $i = 1, \dots, n$ $\lambda_i(A) \geq 0$;

4. *Definida positiva si y sólo si:*

$$\forall x \in \mathbb{R}^n \setminus \{0\} : x^\top A x > 0,\ (A \succ 0),$$

equivalentemente, para todo $i = 1, \dots, n$ $\lambda_i(A) > 0$;

5. *Dado un cono convexo* P *diremos que* A *es copositiva en* P *si*

$$x^\top A x \geq 0,\ \forall\, x \in P.$$

Definición 5.2. *Dada una matriz* $A \in M_n$*, se define* $\operatorname{Ker} A \doteq \{x \in \mathbb{R}^n : Ax = 0\}$

Teorema 5.1. *Sea $A \in S^n$. Si $A \succeq 0$ entonces,*

$$x^\top Ax = 0 \Longleftrightarrow x \in \text{Ker } A$$

Demostración. Como $A \in S^n$, existe matriz ortogonal $O \in S^n$ de modo tal que $A = O\text{diag}(\lambda_1(A), \ldots, \lambda_n(A))O^\top$. Definiendo $y = Ox$, se sigue que

$$\begin{aligned}
&x^\top Ax = 0 && \Leftrightarrow \\
&y^\top \text{diag}(\lambda_1(A), \ldots, \lambda_n(A))y = 0 && \Leftrightarrow \\
&\textstyle\sum_{i=1}^n \lambda_i(A)y_i^2 = 0,
\end{aligned}$$

como $A \succeq 0$, entonces $y_i = 0,\ i = 1, \ldots, n$ y así $O^\top x = 0$, con lo cual $Ax = 0$.

□

Ejemplo 5.1. *La hipótesis $A \succeq 0$ es necesaria. En efecto consideremos*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad x = (1, 1)$$

Se tiene que $x^\top Ax = 0$, sin embargo, $Ax \neq 0$.

5.2. Condiciones de optimalidad.

Uno de los primeros resultados [12] cronológicamente hablando, fue presentado en 1982 por Jorge J. Moré, en este, se dan condiciones de optimalidad para un problema con una restricción, bajo hipótesis de restricción.

Teorema 5.2. [12, Teorema 3.2] *Sean $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$ funciones cuadráticas, esto es $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$, $B \neq 0$, $A,\ B \in S^n$, $a,\ b \in \mathbb{R}^n$, $\alpha,\ \beta \in \mathbb{R}$. Consideremos el problema de minimización:*

$$\min_{\substack{g(x)=0 \\ x \in \mathbb{R}^n}} f(x). \tag{5.1}$$

Además supongamos que se satisface:

$$\inf_{x\in\mathbb{R}^n} g(x) < 0 < \sup_{x\in\mathbb{R}^n} g(x).$$

Entonces:

$$\overline{x} \quad \textit{es solución de } (5.1) \Longleftrightarrow \begin{cases} \exists\ \lambda^* \geq 0 : & \nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0 \\ & A + \lambda^* B \succeq 0 \end{cases}$$

Demostración. $\Rightarrow\rfloor$ Supongamos que $\overline{x}$ es mínimo de (5.1). La demostración requiere considerar dos casos.

Caso 1) Si $\nabla g(\overline{x}) \neq 0$ se cumplen las condiciones de K.K.T., es decir:

$$\exists\ \lambda^* \in \mathbb{R} :\ \nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0.$$

Falta probar que $A + \lambda^* B \succeq 0$. En términos del Lagrangiano

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

se tiene $\nabla_x L(\lambda, \overline{x}) = 0$, además haciendo el desarrollo de Taylor en torno a $\overline{x}$ tenemos:

$$\begin{aligned} L(\lambda^*, x) = &\quad L(\lambda^*, \overline{x}) + \underbrace{\langle \nabla_x L(\lambda^*, \overline{x}), x - \overline{x} \rangle}_{=0} + \tfrac{1}{2}(x - \overline{x})^\top \nabla_x^2 L(\lambda^*, \overline{x})(x - \overline{x}) \\ = &\quad f(\overline{x}) + \lambda^* \underbrace{g(\overline{x})}_{=0} + \tfrac{1}{2}(x - \overline{x})^\top \nabla_x^2 L(\lambda^*, \overline{x})(x - \overline{x}), \end{aligned}$$

definiendo $w \doteq x - \overline{x}$ se tiene:

$$f(w + \overline{x}) + \lambda^* g(w + \overline{x}) = L(\lambda^*, x) = f(\overline{x}) + \frac{1}{2} w^\top \nabla_x^2 L(\lambda^*, \overline{x}) w,$$

es decir:

$$\frac{1}{2} w^\top \nabla_x^2 L(\lambda^*, \overline{x}) w = f(w + \overline{x}) - f(\overline{x}) + \lambda^* g(w + \overline{x}).$$

Así tenemos que si $w + \overline{x}$ es factible ($g(w + \overline{x}) = 0$) entonces $w^\top \nabla_x^2 L(\lambda^*, \overline{x}) w \geq 0$, afirmación por la cual requiere particionar de manera conveniente a $\mathbb{R}^n$, esto es:

$$\mathbb{R}^n = \bigcup_{i=1}^{4} S_i,$$

en donde

$$S_1 \doteq \{w \in \mathbb{R}^n : \ \nabla g(\overline{x})^\top w \neq 0, w^\top Bw \neq 0\}$$

$$S_2 \doteq \{w \in \mathbb{R}^n : \ \nabla g(\overline{x})^\top w = 0, w^\top Bw = 0\}$$

$$S_3 \doteq \{w \in \mathbb{R}^n : \ \nabla g(\overline{x})^\top w = 0, w^\top Bw \neq 0\}$$

$$S_4 \doteq \{w \in \mathbb{R}^n : \ \nabla g(\overline{x})^\top w \neq 0, w^\top Bw = 0\}$$

Notemos que

$$g(tw+\overline{x}) = \frac{1}{2}(tw+\overline{x})^\top B(tw+\overline{x})+b^\top(tw+\overline{x})+\beta = \frac{1}{2}t^2w^\top Bw+tb^\top w+t\overline{x}^\top Bw+\frac{1}{2}\overline{x}^\top B\overline{x}+b^\top\overline{x}+\beta$$

$$= \frac{1}{2}t^2w^\top Bw+t(\overline{x}^\top B+b^\top)w+g(\overline{x}) = \frac{1}{2}t^2w^\top Bw+t(B\overline{x}+B)^\top w+g(\overline{x}) = \frac{1}{2}t^2w^\top Bw+t\nabla g(\overline{x})^\top w+g(\overline{x}).$$

Como $\overline{x}$ es factible, $g(\overline{x}) = 0$, luego

$$g(tw + \overline{x}) = t^2w^\top Bw + t\nabla g(\overline{x})^\top w \tag{5.2}$$

Para concluir que $A + \lambda^* B \succeq 0$, se procede de la siguiente manera. Dado un elemento $w \in S_i$ para cada $i = 1, \ldots, 4$, construir una dirección factible $\overline{x} + tw$ de modo que $g(\overline{x} + tw) = 0$.

Si $w \in S_1$ entonces $\nabla g(\overline{x})^\top w \neq 0, w^\top Bw \neq 0$, por lo que escogiendo

$$t = -\frac{\nabla g(\overline{x})^\top w}{w^\top Bw},$$

tenemos de (5.2) que $g(tw + \overline{x}) = 0$ por lo tanto $w^\top \nabla_x^2 L(\lambda^*, \overline{x})w \geq 0$.

Si $w \in S_2$ entonces $\nabla g(\overline{x})^\top w = w^\top Bw = 0$ tenemos de (5.2) que $g(tw + \overline{x}) = 0$, $\forall\, t \in \mathbb{R}$ por lo tanto $w^\top \nabla_x^2 L(\lambda^*, \overline{x})w \geq 0$.

Sea $w \in S_3$. Como $\nabla g(\overline{x}) \neq 0$ entonces existe $v \neq 0 :\ \nabla g(\overline{x})^\top v \neq 0$ ya que en caso contrario $\nabla g(\overline{x})^\top v = 0,\ \forall\ v \in \mathbb{R}^n$ lo que concluye de manera contradictoria $\nabla g(\overline{x}) = 0$.

De esta manera se define $w_t \doteq w + tv,\ t \in \mathbb{R}$, el cual, satisface $\nabla g(\overline{x})^\top w_t \neq 0, w_t^\top B w_t \neq 0$ para t suficientemente pequeo, por lo tanto $w_t \in S_1$ es decir

$$w_t^\top \nabla_x^2 L(\lambda^*, \overline{x}) w_t \geq 0,$$

en el límite $w^\top \nabla_x^2 L(\lambda^*, \overline{x}) w \geq 0$.

Sea $w \in S_4$. Como $B \in S^n$ es no nula, existe al menos un valor propio no nulo. De esta manera definiendo $w_t \doteq w + tv,\ v^\top B v \neq 0$ y razonando de manera análoga al caso iii) se concluye que $w^\top \nabla_x^2 L(\lambda^*, \overline{x}) w \geq 0$.

De esta manera se a probado que bajo el supuesto $\nabla g(\overline{x}) \neq 0$, que $w^\top \nabla_x^2 L(\lambda^*, \overline{x}) w \geq 0$, para todo $w \in \mathbb{R}^n$.

Si $\nabla g(\overline{x}) = 0$, falta por demostrar que $\nabla f(\overline{x}) = 0$ y que $\nabla_x^2 L(\lambda^*, \overline{x}) \succeq 0$ para concluir la prueba se sigue el siguiente orden:

Probar que para todo $w \in \mathbb{R}^n :\ w^\top B w = 0$ se cumple que $f(\overline{x} + \alpha w) \geq f(\overline{x}),\ \forall\ \alpha \in \mathbb{R}^n$.

Primero notemos que la existencia de w tal que $w^\top B w = 0$ esta garantizada por hipótesis en la restricción g. En efecto, supongamos en caso contrario que:

$$w^\top B w \neq 0,\ \forall\ w \in \mathbb{R}^n \setminus \{0\} \Leftrightarrow w^\top B w < 0 \text{ ó } w^\top B w > 0,\ \forall\ w \in \mathbb{R}^n \setminus \{0\}.$$

Por otro lado, para todo $x \in \mathbb{R}^n$, $x \neq \overline{x}$ del desarrollo de Taylor de g alrededor de $\overline{x}$ se tiene,

$$g(x) = g(\overline{x}) + \langle \nabla g(\overline{x}), x - \overline{x} \rangle + \frac{1}{2}(x - \overline{x})^\top B (x - \overline{x}) = \frac{1}{2}(x - \overline{x})^\top B (x - \overline{x}),$$

ya que, $\nabla g(\overline{x}) = 0$ y $g(\overline{x}) = 0$. En el caso que $\frac{1}{2}(x - \overline{x})^\top B (x - \overline{x}) > 0$ se obtiene que $g(x) > 0$ lo cual implica que $\inf_{x \in \mathbb{R}^n} g(x) \geq 0$ resultado análogo, es decir, $\sup_{x \in \mathbb{R}^n} g(x) \leq 0$ en el caso $\frac{1}{2}(x - \overline{x})^\top B (x - \overline{x}) < 0$, contradiciendo la hipótesis en la restricción g.

De esta forma para todo w tal que $w^\top B w = 0$, $\nabla g(\overline{x}) = 0$ y $g(\overline{x}) = 0$ de la ecuación (5.2) se tiene:

$$g(\overline{x} + tw) = t^2 w^\top B w = 0,\ \forall\ t \in \mathbb{R},$$

esto implica que $f(\overline{x}+tw) \geq f(\overline{x}),\ \forall\ t\ \in\ \mathbb{R}$ ya que $\overline{x}+tw$ es factible y $\overline{x}$ es mínimo.

Probar que para todo $w\ \in\ \mathbb{R}^n:\ w^\top Bw = 0$, entonces $\nabla f(\overline{x})^\top w = 0$.

Como sabemos que $w \in \mathbb{R}^n:\ w^\top Bw = 0$ implica $f(\overline{x}+tw) \geq f(\overline{x}), \forall\ t \in \mathbb{R}^n$. Desarrollando $f(\overline{x}+tw)$ de acuerdo a (5.2) se tiene que $t^2 w^\top Aw + t\nabla f(\overline{x})^\top w \geq 0,\ \forall\ t \in \mathbb{R}$.

$$\text{Entonces}\begin{cases} tw^\top Aw + \nabla f(\overline{x})^\top w \geq 0,\ \forall\ t > 0 \\ \text{y} \\ tw^\top Aw + \nabla f(\overline{x})^\top w \leq 0,\ \forall\ t < 0 \end{cases}$$

Por lo tanto $\nabla f(\overline{x})^\top w = 0$.

Probar que $\nabla f(\overline{x})^\top v = 0$, para todo vector propio v de B.

De la condición $g(\overline{x}) = 0$, $\nabla g(\overline{x}) = 0$ y

$$\inf_{x\in\mathbb{R}^n} g(x) < 0 < \sup_{x\in\mathbb{R}^n} g(x),$$

necesariamente B es indefinida. Sean v_1, v_2 en $\mathbb{R}^n \setminus \{0\}:\ \ v_1^\top Bv_1 < 0$ y $v_1^\top Bv_1 > 0$. Sin pérdida de generalidad se asume que $v_i,\ \ i = 1, 2.$ son vectores propios pues,

$$\nu_{\text{mín}} \leq \frac{x_1^\top Bx}{x^\top x} \leq \nu_{\text{máx}},\ \forall\ x \neq 0,\ \nu_{\text{mín}},\ \nu_{\text{máx}},$$

mínimo y máximo valor propio de B respectivamente y además se escogen de modo que $\nabla f(\overline{x})^\top v_i \geq 0,\ i = 1, 2.$

Como $(v_1^\top Bv_1)(v_2^\top Bv_2) < 0$ entonces el discriminante $\Delta = (2v_1^\top Bv_2)^2 - 4(v_1^\top Bv_1)(v_2^\top Bv_2)$ de

$$(v_2^\top Bv_2)t^2 + 2(v_1^\top Bv_2)t + v_1^\top Bv_1 = (v_1 + tv_2)^\top B(v_1 + tv_2),$$

es positivo, por lo tanto existe $t \in \mathbb{R}$ tal que $(v_1 + tv_2)^\top B(v_1 + tv_2) = 0$ lo cual implica que $\langle \nabla f(\overline{x}), v_1 + tv_2 \rangle = 0$, suponiendo $t > 0$ se obtiene que $\nabla f(\overline{x})^\top v_i = 0,\ i = 1, 2.$ Así queda demostrado que $\nabla f(\overline{x})^\top v = 0$, para todo vector propio v de B, como B es símetrica $\nabla f(\overline{x})$ es ortogonal a $\mathbb{R}^n$ por lo que $\nabla f(\overline{x}) = 0$.

Para finalizar queda probar que $L(\lambda^*, \overline{x}) \succeq 0$, para lo cual se usa el siguiente resultado.

Lema 5.1. [12, Teorema 2.3] *Sean A, C matrices simétricas en $\mathbb{R}^{n\times n}$ y C indefinida. Entonces*

$$w^\top Cw = 0 \Longrightarrow w^\top Aw \geq 0$$

si y sólo si $A + \lambda C \succeq 0$ algún $\lambda \in \mathbb{R}$.

De acuerdo al lema anterior, esto es, suponiendo $w^\top Bw = 0$ y el hecho que $g(\overline{x}) = 0$, $\nabla g(\overline{x}) = 0$, se tiene:

$$g(\overline{x} + w) = w^\top Bw + \nabla g(\overline{x})^\top w + g(\overline{x}) = 0,$$

por lo que $f(\overline{x} + w) \geq f(\overline{x})$. Como $\nabla f(\overline{x}) = 0$ se obtiene:

$$w^\top Aw = f(\overline{x} + w) - f(\overline{x}) \geq 0.$$

Lo cual implica de acuerdo al lema 5.1 que $A + \lambda^* B \succeq 0$ para algún $\lambda^* \in \mathbb{R}$.

$\Leftarrow\rfloor$ Supongamos que $g(\overline{x}) = 0$ y que existe $\lambda^* \in \mathbb{R}$ tal que

$$\nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0$$

$$A + \lambda^* B \succeq 0.$$

Estas condiciones aseguran que el Lagrangiano $L(\lambda^*, \cdot)$ alcanza el mínimo en $\overline{x}$ sobre $\mathbb{R}^n$. De esta manera

$$f(x) + \lambda^* g(x) \geq f(\overline{x}) + \lambda^* g(\overline{x}) = f(\overline{x}),\ \forall\ x\ \in\ \mathbb{R}^n,\ \text{por lo que}$$

$$f(x) \geq f(\overline{x}),\ \forall\ x :\ g(x) = 0,$$

por lo tanto $\overline{x}$ es un mínimo de (5.1).

$\square$

Ejemplo 5.2. [12] *La hipótesis $B \neq 0$ en el teorema anterior es necesaria. Por ejemplo, consideremos $f(x) = x_1^2 - x_2^2$, $g(x) = x_2$. Luego el problema*

$$\min_{\substack{g(x)=0\\ x\in\mathbb{R}^n}} f(x),$$

tiene a $\overline{x} = (0,0)$ *como mínimo y a* $\lambda^* = 0$ *multiplicador. Sin embargo,*

$$\nabla_x^2 L(\overline{x}, \lambda^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix},$$

el cual claramente no es semidefinida positiva.

Por otro lado la hipótesis sobre la restricción tambien es necesaria como debería esperarse. En efecto, al considerar $f(x) = x_1^2 + x_2$, $g(x) = x_2^2$, *el problema*

$$\min_{\substack{g(x)=0\\ x\in\mathbb{R}^n}} f(x),$$

tiene a $\overline{x} = (0,0)$ *como mínimo, sin embargo,*

$$\nabla f(\overline{x}) + \lambda \nabla g(\overline{x}) = (0,1), \ \forall \ \lambda \ \in \ \mathbb{R}.$$

Corolario 5.1. [12, Teorema 3.4] *Sean* $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, esto es,* $f(x) = \frac{1}{2}x^T Ax + a^T x + \alpha$, $g(x) = \frac{1}{2}x^T Bx + b^T x + \beta$, $B \neq 0$, *donde* $A,\ B \ \in \ S^n$, $a,\ b \ \in \ \mathbb{R}^n$, $\alpha,\ \beta \ \in \ \mathbb{R}$. *Se Considera el problema de minimización:*

$$\min_{\substack{g(x)\leq 0\\ x\in\mathbb{R}^n}} f(x). \tag{P}$$

Además supongamos que se satisface:

$$\inf_{x\in\mathbb{R}^n} g(x) < 0.$$

Entonces:

$$\overline{x} \text{ es mínimo global de (P)} \iff \begin{cases} g(\overline{x}) \leq 0 \ y \ \exists \ \lambda^* \ \in \ \mathbb{R} : & \nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0 \\ & A + \lambda^* B \succeq 0 \end{cases}$$

Más aún, $\lambda^* = 0$, si $g(\overline{x}) < 0$.

Demostración. $\Rightarrow\rfloor$ Supongamos que $\overline{x}$ es solución de (P) entonces $\overline{x}$ es factible, es decir, $g(\overline{x}) \leq 0$, lo que genera dos casos:

Caso 1) Si $g(\overline{x}) < 0$, entonces f alcanza el mínimo $\overline{x}$ sobre un abierto, por lo cual $\nabla f(\overline{x}) = 0$ y $A \succeq 0$. De modo que, escogiendo $\lambda^* = 0$ se tiene:

$$\nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0$$

$$A + \lambda^* B \succeq 0$$

Caso 2) Si $g(\overline{x}) = 0$, entonces

$$f(\overline{x}) = \min_{\substack{g(x)\leq 0\\ x\in\mathbb{R}^n}} f(x) = \min_{\substack{g(x)=0\\ x\in\mathbb{R}^n}} f(x).$$

Por otro lado notar que si $\sup_{x\in\mathbb{R}^n} g(x) \leq 0$, entonces $g(x) \leq 0,\ \forall\ x\ \in\ \mathbb{R}^n$, así se obtendría:

$$f(\overline{x}) = \min_{\substack{g(x)\leq 0\\ x\in\mathbb{R}^n}} f(x) = \min_{x\in\mathbb{R}^n} f(x).$$

Por lo tanto nuavemente se alcanza el mínimo sobre un abierto y se concluye de la misma manera que en el caso anterior. Por lo tanto, suponiendo que $\sup_{x\in\mathbb{R}^n} g(x) > 0$, se cuenta con las mismas hipótesis que en el teorema anterior y así entonces existe $\lambda^* \in \mathbb{R}$ tal que:

$$\nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0$$

$$A + \lambda^* B \succeq 0.$$

Estas condiciones aseguran que el Lagrangiano $L(\lambda^*, \cdot)$ alcanza el mínimo en $\overline{x}$ sobre $\mathbb{R}^n$. Esto es:

$$f(x) + \lambda^* g(x) \geq f(\overline{x}) + \lambda^* g(\overline{x}) = f(\overline{x}),\ \forall\ x \in \mathbb{R}^n.$$

De esta manera

$$f(x)+\lambda^* g(x) \geq f(\overline{x}),\ \forall\ x \in \mathbb{R}^n \tag{5.3}$$

Se quiere concluir que $\lambda^* \geq 0$. En efecto supongomas que $\lambda^* < 0$. Entonces se tiene que:

$$S_f^-(\mu) \cap S_g^+(0) = \emptyset \tag{5.4}$$

ya que en caso contrario, es decir, tomando un elemento $x_0 \in S_f^-(\mu) \cap S_g^+(0)$ se cumple que $f(x_0) < \mu = f(\overline{x})$ y $g(x_0) > 0$ por lo que:

$$f(x_0)+\lambda^* g(x_0) < f(\overline{x}),$$

contradiciendo (5.3). Más aún de (5.4) se obtiene que $S_f^-(\mu) = \emptyset$. En efecto, puesto que en caso contrario $f(x) < f(\overline{x})$ implica $g(x) \leq 0$ lo cual contradice que $\overline{x}$ es mínimo. De esta manera $f(x) \geq f(\overline{x}),\ \forall\ x \in \mathbb{R}^n$ lo cual implica que $\nabla f(\overline{x}) = 0$ y $A \succeq 0$.

Lo que se acaba de mostrar es que si $\lambda^* < 0$, siempre es posible redefinir un nuevo multiplicador $\overline{\lambda} = 0$ de manera que se cumplan las condiciones de optimalidad.

$$\begin{aligned} &\nabla f(\overline{x}) + \overline{\lambda}\nabla g(\overline{x}) = 0 \\ &A + \overline{\lambda} B \succeq 0. \end{aligned}$$

finalizando de esta manera la primera parte de la demostración.

$\Leftarrow\rfloor$ Es claro usando el mismo argumento dado por el Lagrangiano.

$\square$

5.3. Lagrangiano regularizado para problemas cuadráticos con restricciones lineales

En esta sección, introducimos el concepto de dualidad fuerte asociado al Lagrangiano regularizado (para más detalle ver [11]), el cual justificamos su uso en los capítulos posteriores. Consideremos el problema de optimización cuadrática con m restricciones cuadráticas y r restricciones lineales,

$$\inf_{\substack{g(x)\in -P \\ x\in C}} f(x), \tag{P$_\mathrm{m}$}$$

en donde, $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g_i(x) = \frac{1}{2}x^\top B_i x + b_i^\top x + \beta_i$, $A,\ B_i \in S^n$, $a,\ b_i \in \mathbb{R}^n$, $\alpha,\ \beta_i \in \mathbb{R}$, para cada $i = 1, \ldots, m$. $C = \{x : Hx = d\}$, $H \in \mathbb{R}^{r\times n}$, $d \in \mathbb{R}^r$. Para efectos prácticos escribimos la matriz H de la forma

$$H = \begin{pmatrix} h_1^\top \\ \vdots \\ h_r^\top \end{pmatrix},\ h_j \in \mathbb{R}^n.$$

Es común ver en la literatura asociada al tema, definir al Lagrangiano asociado al problema $(\mathrm{P_m})$ como

$$L_S(x, \lambda, \nu) \doteq f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{r} \nu_j (h_j^T x - d_j),\ x \in \mathbb{R}^n,\ \lambda \in \mathbb{R}_+^m,\ \nu \in \mathbb{R}^r.$$

Lamentablemente, para este Lagrangiano mostraremos que Strong Duality falla. Veamos el siguiente ejemplo que evidencia este hecho.

Ejemplo 5.3. [11] *Sean $f(x, y) = -y^2$, $g(x, y) = x^2+y^2-1$ y $H(x, y) = x+y$ con $d = 0$. Considere el siguiente problema cuadrático con restricción lineal,*

$$\min_{\substack{g(x,y)\leq 0 \\ H(x,y)=0}} f(x, y)$$

El valor óptimo es $-\frac{1}{2}$ y es alcanzado en los puntos $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$, $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. El problema dual asociado al Lagrangiano usual es

$$\max_{\substack{\lambda\geq 0 \\ \nu\in\mathbb{R}}} \inf_{(x,y)\in\mathbb{R}^2} \{-y^2 + \lambda(x^2 + y^2 - 1) + \nu(x + y)\}.$$

el cual tiene como valor óptimo -1. Por lo tanto Strong Duality no se cumple.

Sin embargo considerando el problema dual

$$\max_{\lambda\geq 0} \inf_{(x,y)\in H^{-1}(0)} \{-y^2 + \lambda(x^2 + y^2 - 1)\},$$

se tiene que el óptimo es $-\frac{1}{2}$.

Definición 5.3. *Para el problema* ($\mathrm{P_m}$) *se define el Lagrangiano Regularizado como*

$$L(x,\lambda) \doteq f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),\ x \ \in \ H^{-1}(d),\ \lambda \ \in \ \mathbb{R}^m_+.$$

Notemos que para todo $\lambda \ \in \ \mathbb{R}^m_+,\ \nu \ \in \ \mathbb{R}^r$, tenemos $f(x) + \langle \lambda, g(x)\rangle = f(x) + \langle \lambda, g(x)\rangle + \langle \nu, Hx - d\rangle,\ \forall\ x \ \in \ C$. Luego se cumple

$$\operatorname*{ínf}_{x\in\mathbb{R}^n} \{f(x) + \langle \lambda, g(x)\rangle + \langle \nu, Hx-d\rangle\} \leq \operatorname*{ínf}_{x\in C}\{f(x) + \langle \lambda, g(x)\rangle\},$$

entonces

$$\sup_{\substack{\lambda\in\mathbb{R}^m_+\\ \nu\in\mathbb{R}^n}} \operatorname*{ínf}_{x\in\mathbb{R}^n} \{f(x) + \langle \lambda, g(x)\rangle + \langle \nu, Hx-d\rangle\} \leq \sup_{\lambda\in\mathbb{R}^m_+} \operatorname*{ínf}_{x\in C}\{f(x) + \langle \lambda, g(x)\rangle\}.$$

Por lo tanto, si se cumple Strong Duality con respecto al Lagrangiano usual entonces también se cumple Strong Duality con respecto al Lagrangiano regularizado. Teniendo en cuenta este hecho y el ejemplo 5.3, se sigue que la propiedad de Strong Duality con respecto al Lagrangiano regularizado es una condición más débil por lo cual etrega más información. De esta manera los resultados obtenidos en este trabajo se obtendrán considerando el Lagrangiano regularizado.

El siguiente resultado dado en [11], los autores caracterizan la propiedad de dualidad fuerte, por medio de el Lagrangiano Regularizado, bajo el supuesto que el problema primal tiene solución ($\mathrm{argmin}_K f \neq \emptyset$). Este resultado, generaliza en parte, los resultados de problemas cuadráticos con una restricción.

Teorema 5.3. [11, Teorema 2.1] *Sean* $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, esto es* $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g_i(x) = \frac{1}{2}x^\top B_i x + b_i^\top x + \beta_i$, $A,\ B_i \ \in \ S^n$, $a,\ b_i \ \in \ \mathbb{R}^n$, $\alpha,\ \beta_i \ \in \ \mathbb{R}$, *para cada* $i = 1, \ldots, m$. *Consideremos* $P = \mathbb{R}^m_+$, $C = H^{-1}(d) \doteq \{x :\ Hx = d\}$, *donde* $H \in \mathbb{R}^{r\times n},\ d \in \mathbb{R}^r$. *Entonces,*

$$\overline{x} \in \mathrm{argmin}_K f(x),\ \mu = \max_{\lambda\in\mathbb{R}^m_+} \ \inf_{x\in H^{-1}(d)} L(\lambda, x) \Longleftrightarrow \begin{cases} \exists \lambda^* = (\lambda_1^*, \ldots, \lambda_m^*) \in \mathbb{R}^m_+, y \in \mathbb{R}^r : \\ \\ \nabla f(\overline{x}) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\overline{x}) + H^\top y = 0 \\ \\ A + \sum_{i=1}^m \lambda_i^* B_i,\ \text{ es copositiva en } \mathrm{Ker}(H). \\ \\ \lambda_i^* g_i(\overline{x}) = 0,\ i = 1, \ldots, m. \end{cases}$$

Demostración. En este caso

$$L(x,\lambda^*) = f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x),\ x \in H^{-1}(d).$$

$\Rightarrow\rfloor$ Supongamos que $\overline{x} \in \operatorname{argmin}_K f,\ \mu = \max\limits_{\lambda^* \in \mathbb{R}_+^m} \inf\limits_{x \in H^{-1}(d)} L(\lambda^*, x),$

$$\exists\ \lambda^* \in \mathbb{R}_+^m :\ f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geq f(\overline{x}),\ \forall\ x \in H^{-1}(d)$$

$$\Leftrightarrow \exists\ \lambda^* \in \mathbb{R}_+^m :\ f(x+\overline{x}) + \sum_{i=1}^{m} \lambda_i^* g_i(x+\overline{x}) \geq f(\overline{x}),\ \forall\ x \in\ \operatorname{Ker}(H),$$

en particular para $x = 0 \in\ \operatorname{Ker}(H),\ \ \sum\limits_{i=1}^{m} \lambda_i^* g_i(\overline{x}) \geq 0$ y $g_i(\overline{x}) \leq 0,\ \forall i = 1, \ldots, m$, con lo cual

$$\lambda_i^* g_i(\overline{x}) = 0,\ \forall\ i = 1, \ldots, m.$$

Así

$f(x+\overline{x}) + \sum\limits_{i=1}^{m} \lambda_i^* g_i(x+\overline{x}) \geq f(\overline{x}) = f(\overline{x}) + \sum\limits_{i=1}^{m} \lambda_i^* g_i(\overline{x}),\ \forall\ x \in \operatorname{Ker}(H) \Rightarrow 0$ es un mínimo global de

$$L(\cdot + \overline{x}, \lambda^*) \text{ sobre } \operatorname{Ker}(H) \Rightarrow \langle \nabla L(\overline{x}, \lambda^*), x \rangle \leq 0,\ \forall\ x\ \in\ \operatorname{Ker}(H),$$

Como $\operatorname{Ker}(H)$ es subespacio vectorial se concluye que

$$\langle \nabla L(\overline{x}, \lambda^*), x \rangle = 0,\ \forall\ x\ \in\ \operatorname{Ker}(H) \Rightarrow \nabla L(\overline{x}, \lambda^*)\ \in\ (\operatorname{Ker}(H))^{\perp} = \operatorname{Im}(H^{\top})$$

$$\Leftrightarrow \exists\ y \in \mathbb{R}^m :\ \nabla f(\overline{x}) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\overline{x}) + H^{\top} y = 0.$$

Por otro lado

$$L(x+\overline{x},\lambda^*) = L(\overline{x},\lambda) + \underbrace{\langle \nabla L(\overline{x},\lambda^*), x\rangle}_{=0} + \frac{1}{2}x^\top \sum_{i=1}^m \lambda_i^* B_i x, \ \forall\ x\ \in\ \mathrm{Ker}(H)$$

$$\Rightarrow \frac{1}{2}x^\top \sum_{i=1}^m \lambda_i^* B_i x = L(x+\overline{x},\lambda^*) - L(\overline{x},\lambda^*) \geq 0 \Rightarrow \sum_{i=1}^m \lambda_i^* B_i \text{ copositiva en } \ \mathrm{Ker}(H)$$

$\Leftarrow\rfloor$ Supongamos que se cumplen las condiciones de optimalidad,

$\exists\ \lambda^* \in \mathbb{R}^m_+, y \in \mathbb{R}^r :$

$\nabla f(\overline{x}) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\overline{x}) + H^\top y = 0$

$\nabla_x^2 f(\overline{x}) + \sum_{i=1}^m \lambda_i^* B_i$ copositiva en $\mathrm{Ker}(H)$.

$\lambda_i^* g_i(\overline{x}) = 0.$

Del desarrollo de Taylor de $L(\cdot,\lambda^*)$ en torno a $\overline{x}$ se tiene que $\forall\ x\ \in\ \mathrm{Ker}(H)$

$$L(x+\overline{x},\lambda^*) = L(\overline{x},\lambda) + \langle \nabla L(\overline{x},\lambda^*), x\rangle + \frac{1}{2}\underbrace{x^\top \sum_{i=1}^m \lambda_i^* B_i x}_{\geq 0}$$

$$\geq L(\overline{x},\lambda^*) + \langle -H^\top y, x\rangle = L(\overline{x},\lambda^*) + \underbrace{\langle -y, Hx\rangle}_{=0} = L(\overline{x},\lambda^*)$$

$$\Rightarrow f(x+\overline{x}) + \sum_{i=1}^m \lambda_i^* g_i(x+\overline{x}) \geq f(\overline{x}) + \underbrace{\sum_{i=1}^m \lambda_i^* g_i(\overline{x})}_{=0} = f(\overline{x}), \ \forall\ x\ \in\ \mathrm{Ker}(H)$$

$$\Leftrightarrow f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) \geq f(\overline{x}), \ \forall\ x\ \in\ H^{-1}(d) \Rightarrow \overline{x}\ \in\ \mathrm{argmin}_K f(x), \ \mu = \max_{\lambda\in\mathbb{R}^m_+} \inf_{x\in H^{-1}(d)} L(\lambda,x)$$

$\square$

El siguiente resultado muestra que con respecto al Lagrangiano usual también se pueden formular condiciones de optimalidad.

Teorema 5.4. *Sean* $f:\mathbb{R}^n\to\mathbb{R}$ $g:\mathbb{R}^n\to\mathbb{R}$ *funciones cuadráticas, esto es* $f(x)=\frac{1}{2}x^\top Ax+a^\top x+\alpha$, $g_i(x)=\frac{1}{2}x^\top B_i x+b_i^\top x+\beta_i$, $A,\ B_i\ \in\ S^n$, $a,\ b_i\ \in\ \mathbb{R}^n$, $\alpha,\ \beta_i\ \in\ \mathbb{R}$, *para cada* $i=1,\ldots,m$. *Consideremos* $P=\mathbb{R}^m_+\times\Theta^r$, $C=\mathbb{R}^n$, *donde* $H\in\mathbb{R}^{r\times n}$, $d\ \in\ \mathbb{R}^r$. *Entonces*

$$\overline{x}\in\ \operatorname{argmin}_K f(x),\ \mu=\max_{\substack{\lambda\in\mathbb{R}^m_+\\ \nu\in\mathbb{R}^r}}\ \inf_{x\in\mathbb{R}^n} L(\lambda,\nu,x)\Longleftrightarrow\begin{cases}\exists\ \lambda^*\in\mathbb{R}^m_+,\nu^*\in\mathbb{R}^r:\\[2mm] \nabla f(\overline{x})+\sum_{i=1}^m\lambda_i^*\nabla g_i(\overline{x})+H^\top\nu^*=0\\[2mm] A+\sum_{i=1}^m\lambda_i^*B_i\succeq 0\\[2mm] \lambda_i^*g_i(\overline{x})=0\end{cases}$$

Demostración. $\Rightarrow\rfloor$ Supongamos que $\overline{x}\in\operatorname{argmin}_K f,\ \mu=\max\limits_{\substack{\lambda^*\in\mathbb{R}^m_+\\ \nu\in\mathbb{R}^r}}\ \inf\limits_{x\in\mathbb{R}^n} L(\lambda,\nu,x)$, entonces

$$\exists\ \lambda^*\in\mathbb{R}^m_+,\nu^*\in\mathbb{R}^r:\ f(x)+\sum_{i=1}^m\lambda_i^*g_i(x)+\sum_{i=1}^m\nu_i^*(h_j^\top x-d_j)\geq f(\overline{x}),\ \forall x\in\mathbb{R}^n$$

en particular para $x=\overline{x}\in K$, $\sum_{i=1}^m\lambda_i^*g_i(\overline{x})\geq 0$, $g_i(\overline{x})\leq 0,\forall\ i=1,...,m$ y $\nu_j^*(h_j^\top\overline{x}-d_j)=0,\forall\ j=1,\ldots,r$. Como cada $\lambda_i^*\geq 0$ se tiene que $\lambda_i^*g_i(\overline{x})\leq 0$, con lo cual $\sum_{i=1}^m\lambda_i^*g_i(\overline{x})\leq 0$. De esta manera

$$\sum_{i=1}^m\lambda_i^*g_i(\overline{x})=0,\ \sum_{j=1}^r\nu_j^*(h_j^\top\overline{x}-d_j)=0,$$

así $\lambda_i^*g_i(\overline{x})=0,\ \forall\ i=1,\ldots,m$. Por otro lado

$$f(x)+\sum_{i=1}^m\lambda_i^*g_i(x)+\sum_{j=1}^r\nu_j^*(h_j^\top x-d_j)\geq f(\overline{x})=f(\overline{x})+\sum_{i=1}^m\lambda_i^*g_i(\overline{x})+\sum_{j=1}^r\nu_j^*(h_j^\top\overline{x}-d_j),\ \forall\ x\in\mathbb{R}^n$$

$$\Rightarrow\overline{x}\text{ es un mínimo global de }L(\lambda^*,\nu^*,\cdot)\Rightarrow\nabla L(\lambda^*,\nu^*,\overline{x})=0,\ A+\sum_{i=1}^m\lambda_i^*B_i\succeq 0$$

$$\Leftrightarrow\exists\ \lambda^*\in\mathbb{R}^m_+,\nu^*\in\mathbb{R}^r:\ \nabla f(\overline{x})+\sum_{i=1}^m\lambda_i^*\nabla g_i(\overline{x})+\nabla(\sum_{j=1}^r\nu_j^*h_j^\top\overline{x}-d_j)$$

$$= \nabla f(\overline{x}) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\overline{x}) + \sum_{j=1}^{r} \nu_j^* h_j^\top = \nabla f(\overline{x}) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(\overline{x}) + H^T \nu^* = 0,$$

$$A + \sum_{i=1}^{m} \lambda_i^* B_i \succeq 0.$$

$\Leftarrow\rfloor$ Supongamos que:

$$\exists\ \lambda^* \in \mathbb{R}_+^m, \nu^* \ \in \mathbb{R}^r :$$

$$\nabla f(\overline{x}) + \textstyle\sum_{i=1}^{m} \lambda_i^* \nabla g_i(\overline{x}) + H^\top \nu^* = 0$$

$$A + \textstyle\sum_{i=1}^{m} \lambda_i^* B_i \succeq 0$$

$$\lambda_i^* g_i(\overline{x}) = 0$$

Estas hipótesis permiten asegurar que $L(\lambda^*, \nu^*, \cdot)$ es convexa con $\overline{x}$ punto crítico, entonces

$$L(\lambda^*, \nu^*, x) \geq L(\lambda^*, \nu^*, \overline{x}),\ \forall\ x \in \mathbb{R}^n$$

$$\Leftrightarrow f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) + \sum_{j=1}^{r} \nu_j^* (h_j^\top x - d_j) \geq f(\overline{x}) + \underbrace{\sum_{i=1}^{m} \lambda_i^* g_i(\overline{x})}_{=0} + \underbrace{\sum_{j=1}^{r} \nu_j^* (h_j^\top \overline{x} - d_j)}_{=0} = f(\overline{x}),\ \forall\ x \in \mathbb{R}^n.$$

$\square$

Observación 5.1. *En ausencia de las restricciones lineales ($H = 0,\ d = 0$) los teoremas 5.3 y 5.4 son equivalentes.*

Capítulo 6

Strong Duality v/s condiciones de optimalidad: Caso cuadrático con un restricción cuadrática y una lineal

6.1. Convexidad de imagen de formas cuadráticas y teoremas alternativos

Esta sección esta enfocada a estudiar algunos resultados clásicos asociado a las formas cuadráticas. Uno de estos corresponde al teorema de Dines de 1942 y el cual asegura que la imagen de dos formas cuadráticas homogéneas es siempre un conjunto convexo de $\mathbb{R}^2$. A raíz de lo anterior los teoremas del tipo alternativo para dos desigualdades cuadráticas el cual estudiaremos, nos permiten obtener una versión relajada del teorema de Dines y de acuerdo a nosotros nuevo en la literatura.

Teorema 6.1. [15, Teorema 4.4] (Dines) *Sean $f(x) = x^\top Ax$, $g(x) = x^\top Bx$ con $A,\ B \in\ \mathcal{S}^n$. Entonces,*

$$\{(f(x), g(x)) :\ x\ \in\ \mathbb{R}^n\}$$

es un conjunto convexo en $\mathbb{R}^2$.

Demostración. Sea $F(\mathbb{R}^2) \doteq \{(f(x), g(x)) :\ x\ \in\ \mathbb{R}^n\}$. La prueba se hace usando la definición de convexidad, para esto, dado dos elementos $u,\ v\ \in\ F(\mathbb{R}^2)$, se presentan 2 casos.

Caso 1) Si u, v son colineales con respecto al origen entonces existe $\alpha\ \in\ \mathbb{R}$, tal que $|\alpha| < 1$ y $u = \alpha v$, luego para cualquier $\lambda\ \in\ (0, 1)$ se tiene que $\lambda u + (1 - \lambda)v = [\underbrace{(\alpha - 1)\lambda + 1}_{>0}]v\ \in\ F(\mathbb{R}^n)$, pues f, g son homogéneas.

Caso 2) $u,\ v$ no son colineales, luego $u = (u_f, u_g)$ y $v = (v_f, v_g)$, entonces existen $x_u,\ x_v\ \in\ \mathbb{R}^n$

tal que, $u = (u_f, u_g) = (f(x_u), g(x_u))$ y $v = (v_f, v_g) = (f(x_v), g(x_v))$. Puesto que u, v son no colineales entonces

$$d \doteq \begin{vmatrix} u_f & v_f \\ u_g & v_g \end{vmatrix} = u_f v_g - u_g v_f \neq 0.$$

Sea $\lambda \in (0,1)$ arbitrario. Buscamos $x_\lambda \in \mathbb{R}^n$ de modo que

$$(f(x_\lambda), g(x_\lambda)) = \lambda u + (1-\lambda)v. \tag{6.1}$$

Para esto, consideremos el plano $x_\lambda = \rho(x_u \cos\theta + x_v \sin\theta)$ en donde $\rho,\ \theta$ son variables. Resolviendo a través de sus componentes el sistema (6.1) con respecto a las variables ρ, θ, λ se obtiene,

$$\begin{cases} \rho^2 f(x_u \cos\theta + x_v \sin\theta) = (1-\lambda)u_f + \lambda v_f \\ \rho^2 g(x_u \cos\theta + x_v \sin\theta) = (1-\lambda)u_g + \lambda v_g \end{cases}$$

Al igualar ambas ecuaciones con respecto a ρ^2 se obtiene,

$$\frac{(1-\lambda)u_f + \lambda v_f}{f(x_u \cos\theta + x_v \sin\theta)} = \frac{(1-\lambda)u_g + \lambda v_g}{g(x_u \cos\theta + x_v \sin\theta)}$$

con lo cual es posible construir la siguiente función dependiente de θ:

$$\lambda(\theta) = \frac{u_g f(x_u \cos\theta + x_v \sin\theta) - u_f g(x_u \cos\theta + x_v \sin\theta)}{(u_g - v_g) f(x_u \cos\theta + x_v \sin\theta) - (u_f - v_f) g(x_u \cos\theta + x_v \sin\theta)},$$

la cual esta bien definida, para verificar esto, denotemos como $T(\theta)$ al denominador de $\lambda(\theta)$, luego, $T(\theta)$ es cuadrática al ser combinación lineal de cuadráticas por lo que en forma paralela se puede escribir $T(\theta) = \alpha \cos^2\theta + \beta \sin^2\theta + 2\gamma \cos\theta \sin\theta$, por otro lado, $d = T(0) = T(\pm\frac{\pi}{2}) \Rightarrow \alpha = \beta = d \therefore T(\theta) = d + \gamma \sin 2\theta$. Luego observando que,

$$T(\theta) > 0, \text{ si } \big[\gamma \geq 0 \text{ y } \theta \in [0, \frac{\pi}{2}]\big] \text{ ó } \big[\gamma \leq 0 \text{ y } \theta \in [-\frac{\pi}{2}, 0]\big],$$

es que sin perdida de generalidad se supone que $\gamma \geq 0$ y así como se deseaba $\lambda(\theta)$ esta bien definida en $[0, \frac{\pi}{2}]$.

Por otro lado se tiene que $\lambda(\theta)$ es continua al ser composición de funciones cuadráticas, además, $\lambda(0) = 0,\ \lambda(\frac{\pi}{2}) = 1$, así por el teorema del valor intermadio existe $\theta_\lambda \in (0,1)$ de modo que $\lambda(\theta_\lambda) = \lambda$ y así x_λ satisface (6.1).

$\square$

Ejemplo 6.1. *El resultado anterior no es válido para formas cuadráticas no homogéneas. En efecto sean* $f(x,y) = x+y-x^2-y^2-2xy$, $g(x,y) = x^2+y^2+2xy-1$, $F(\mathbb{R}^2) = \{(f(x,y), g(x,y)) :\ (x,y)\ \in\ \mathbb{R}^2\}$.

Consideremos los puntos $(0,0) = (f(0,1), g(0,1))$, $(-2,0) = (f(-1,0), g(-1,0))$ *y* $t = \frac{1}{2}$. *Sin embargo, una vez dicho esto, mostraremos que* $(-1,0) = \frac{1}{2}(0,0) + \frac{1}{2}(-2,0)\ \notin\ F(\mathbb{R}^2)$.

$$\begin{array}{l} f(x,y) = -1 \\ g(x,y) = 0 \end{array} \Leftrightarrow \begin{array}{l} x+y-(x+y)^2 = -1 \\ (x+y)^2 - 1 = 0, \end{array}$$

es decir, $|x+y| = 1$ *y* $x+y = 0$. *Por lo tanto el sistema anterior no tiene solución, con lo cual se concluye que* $(-1,0)\ \notin\ F(\mathbb{R}^2)$.

gráficamente se puede ver la no convexidad de $F(\mathbb{R}^2)$.

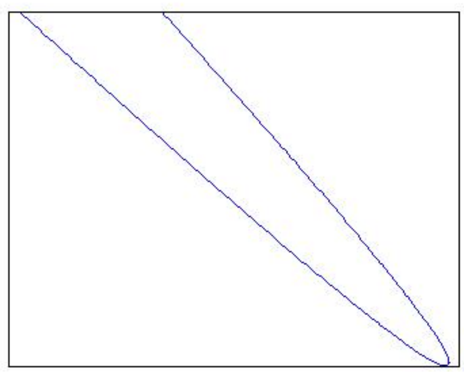


Figura 6.1: Grafica de $F(\mathbb{R}^2)$

Teorema 6.2. [15, Teorema 5.6] (Brickman) *Sean $A,\ B\ \in\ S^n$ y $n \geq 3$. Entonces,*

$$\{(x^\top Ax, x^\top Bx):\ \|x\| = 1\} \ \ \textit{es convexo y compacto en } \mathbb{R}^2.$$

La hipótesis sobre la dimensión de las matrices es necesaria tal cual como se muestra en el siguiente ejemplo.

Ejemplo 6.2. [15, Observación 5.7] *Consideremos las matrices,*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Haciendo la parametrización $x(\theta) = (\cos\theta, \sin\theta)$, $\theta\ \in\ [0, 2\pi]$, se obtiene:

$$\{(x^\top Ax, x^\top Bx):\ \|x\| = 1\} = \{(\cos 2\theta, \sin 2\theta):\ \theta\ \in\ [0, 2\pi]\} = B(0,1) \ \textit{esfera unitaria, en } \mathbb{R}^2,$$

la cual claramente no es convexa.

Definición 6.1. *Dado un subconjunto $K \subseteq \mathbb{R}^n$, diremos que K es un cono regular si $K \cup -K$ es un subespacio vectorial.*

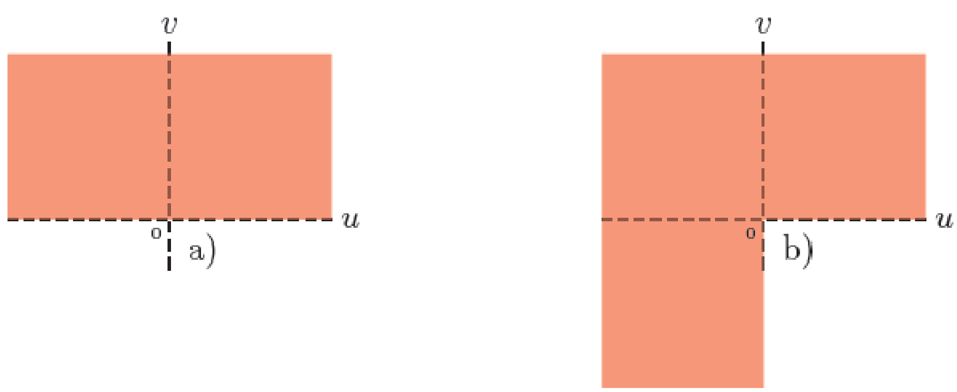


Figura 6.2: Ejemplos de conos regulares: a) $\{(u,v):\ v \geq 0\}$, b) $\{(u,v):\ u \leq 0\} \cup \mathbb{R}^2_+$

Teorema 6.3. [10, Teorema 3.2](Dines Generalizado 2009)
Sean $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas homogéneas, esto es,* $f(x) = \frac{1}{2}x^\top Ax$, $g(x) = \frac{1}{2}x^\top Bx$, $A,\ B\ \in\ S^n$. *Sea* $K \subseteq \mathbb{R}^n$ *un cono regular. Entonces*

$$\{(f(x), g(x)) :\ x\ \in\ K\},\ \textit{es convexo.}$$

Demostración. Sea $F(K) = \{(f(x), g(x)) :\ x\ \in\ K\}$ y $S = K \cup -K$. Se sigue que S es un subespacio de dimension m con $m \leq n$. Además como las formas son homogéneas, $F(K) = F(-K)$. Por lo tanto $F(K) = \{(f(x), g(x)) :\ x\ \in\ S\}$. Como S es un subespacio vectorial de dimensión m, entonces existe una matriz $Q\ \in\ \mathbb{R}^{n\times m}$, de rango completo tal que $S = \{Qy :\ y\ \in\ \mathbb{R}^m\}$. De esta forma

$$\begin{aligned} F(K) &= \{(\tfrac{1}{2}x^\top Ax, \tfrac{1}{2}x^\top Bx) :\ x\ \in\ S\} \\ &= \{(\tfrac{1}{2}y^\top(Q^\top AQ)y, \tfrac{1}{2}y^\top(Q^\top BQ)y) :\ y\ \in\ \mathbb{R}^m\}. \end{aligned}$$

El cual por teorema 6.1 es convexo.

□

Teorema 6.4. [10, Teorema 3.6](Yuan 2009) *Sean* $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, esto es,* $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$, $A,\ B\ \in\ S^n$, $a,\ b\ \in\ \mathbb{R}^n$, $\alpha,\ \beta\ \in\ \mathbb{R}$. *Sea* $a_0\ \in\ \mathbb{R}^n$ *y sea* $S \subseteq \mathbb{R}^n$ *un subespacio vectorial. Entonces se cumple sólo una de las siguientes alternativas:*

a) *Existe* $x_0\ \in\ a_0 + S$ *tal que*

$$\begin{aligned} f(x) < 0 \\ g(x) < 0 \end{aligned}$$

b) *Existe* $(\gamma^*, \lambda^*)\ \in\ \mathbb{R}^2_+ \setminus \{(0,0)\}$ *tal que*

$$\gamma^* f(x) + \lambda^* g(x) \geq 0,\ \forall\ x\ \in\ a_0 + S.$$

Demostración. $b) \Rightarrow$ no $a)\rfloor$ Es directo.

no $a) \Rightarrow b)\rfloor$ Como el sistema $a)$ no tiene solución entonces

$$x \in S: \ f_1(x) = f(x+a_0) < 0, \ g_1(x) = g(x+a_0) < 0$$

no tiene solución. Definiendo las formas homogéneas asociadas a f_1 y g_2, esto es

$$\widetilde{f}_1(x,t) \doteq \frac{1}{2}x^\top A_1 x + t a_1^\top x + t^2\alpha_1$$

$$\widetilde{g}_1(x,t) \doteq \frac{1}{2}x^\top B_1 x + t b_1^\top x + t^2\beta_1$$

en donde $A_1 = A, \ B_1 = B, \ a_1 = a + Aa_0, \ b_1 = b + Ba_0, \ \alpha_1 = \alpha + \frac{1}{2}a_0^\top Aa_0 + a^\top a_0, \ \beta = \beta + \frac{1}{2}a_0^\top Ba_0 + b^\top a_0$.

De esta manera se tiene que el sistema homogéneo

$$\widetilde{f}_1(x,t) < 0, \ \widetilde{f}_2(x,t) < 0, \ (x,t) \in S \times \mathbb{R}_+$$

es no soluble. En efecto, supongamos que (x,t) es una solución del sistema homogéneo anterior.

Si $t \neq 0$, entonces $f(\frac{x}{t} + a_0) = f_1(\frac{x}{t}) = t^{-2}\widetilde{f}_1(x,t) < 0$ y $g(\frac{x}{t} + a_0) = g_1(\frac{x}{t}) = t^{-2}\widetilde{g}_1(x,t) < 0$ lo cual contradice el hecho que la hipótesis $a)$ no se cumpla.

Si $t = 0$, entonces $\frac{1}{2}x^\top Ax = \widetilde{f}_1(x,0) < 0$ y $\frac{1}{2}x^\top Bx = \widetilde{g}_1(x,0) < 0$, esto implica que

$$f(nx) \longrightarrow -\infty, \ g(nx) \longrightarrow -\infty, \ \text{cuando } n \to \infty.$$

Por lo tanto existe $N \in \mathbb{N}$ de modo que $f(Nx) < 0$ y $g(Nx) < 0$ lo cual es una contradicción.

Como $S \times \mathbb{R}_+$ es un cono regular, $\widetilde{f}_1, \widetilde{g}_1$ son homogéneas en virtud al teorema 6.3 y argumentos de separación de conjuntos convexos, existe $(\gamma^*, \lambda^*) \in \mathbb{R}_+^2 \setminus \{(0,0)\}$ tal que

$$\gamma^* \widetilde{f}_1(x,t) + \lambda^* \widetilde{g}_1(x,t) \geq 0, \ \forall \in (x,t) \ S \times \mathbb{R}^+.$$

En particular para $t = 1$, se cumple $b)$.

$\square$

Corolario 6.1. [15, Teorema 2.2](S-lemma, Yakubovich 1971) *Sean* $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, esto es,* $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$, $A,\ B\ \in\ S^n$, $a,\ b\ \in\ \mathbb{R}^n$, $\alpha,\ \beta\ \in\ \mathbb{R}$. *Supongamos que existe* $\overline{x} \in \mathbb{R}^n$ *tal que* $g(\overline{x}) < 0$. *Entonces se cumple sólo una de las siguientes alternativas:*

a) *Existe* $x\ \in\ \mathbb{R}^n$ *tal que*

$$\begin{aligned} f(x) &< 0 \\ g(x) &\leq 0 \end{aligned}$$

b) *Existe* $\lambda^* \geq 0$ *tal que:*

$$f(x) + \lambda^* g(x) \geq 0\ \forall\ x\ \in\ \mathbb{R}^n.$$

Demostración. $b) \Rightarrow$ no $a)\rfloor$ Es directo.

no $a) \Rightarrow b)\rfloor$ Por hipótesis se tiene que el sistema

$$\begin{aligned} f(x) &< 0 \\ g(x) &< 0 \end{aligned}$$

no es soluble, luego aplicando el teorema 6.4, en donde $S = \mathbb{R}^n$ y $a_0 = 0$, existe $(\lambda_1, \lambda_2)\ \in\ \mathbb{R}^2_+ \setminus \{(0,0)\}$ de modo tal que

$$\lambda_1 f(x) + \lambda_2 g(x) \geq 0,\ \forall\ x\ \in\ \mathbb{R}^n.$$

Por otro lado λ_1 tiene que ser positivo pues en caso contrario $\lambda_2 > 0$ y $\lambda_2 g(x) \geq 0$ para todo $x\ \in\ \mathbb{R}^n$. Sin embargo para $x = \overline{x}$, $\lambda_2 g(\overline{x}) < 0$ lo cual es una contradicción.

Finalmente haciendo $\lambda^* = \frac{\lambda_2}{\lambda_1}$ se cumple $b)$.

□

Como adelantamos a principios de la sección, el cono de la imagen de dos funciones cuadráticas arbitrarias, no necesariamente homogéneas es un conjunto convexo en $\mathbb{R}^2$, el cual, damos su prueba de fácil demostración.

Teorema 6.5. *Sean* $f:\mathbb{R}^n \to \mathbb{R}$ $g:\mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, esto es,* $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$, $A,\ B\ \in\ S^n$, $a,\ b\ \in\ \mathbb{R}^n$, $\alpha,\ \beta\ \in\ \mathbb{R}$. *Además consideramos* μ *finito y* $C = H^{-1}(d)$, *en donde,* $H\ \in\ M^{m\times n}$, $d\ \in\ \mathbb{R}^m$. *Entonces*

$$\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) \text{ es convexo.}$$

Demostración. Dado que μ es finito, no existe $x\ \in\ \mathbb{R}^n$ que resuelva el sistema $g(x) < 0$, $Hx = d$ y $f(x) - \mu < 0$. Por otro lado $C = x_0 + \mathrm{Ker}\ H$ para cada elemento fijo $x_0\ \in\ C$. De esta forma y en virtud del teorema 6.4, asegura la existencia de $(\gamma,\lambda)\ \in\ \mathbb{R}^2_+ \setminus \{(0,0)\}$ tal que

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0\ \forall\ x\ \in\ C = x_0 + \mathrm{Ker}\ H.$$

Aplicando el teorema 4.2 se concluye que $\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ es convexo.

□

6.2. Caso cuadrático no-homogéneo con restricción lineal y cuadrática

Vimos en la sección 4.2 que son equivalentes,

$$\gamma(f(x) - \mu) + \lambda g(x) \geq 0\ \forall\ x\ \in\ C \iff (\gamma,\lambda)\ \in\ [\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})]^*,$$

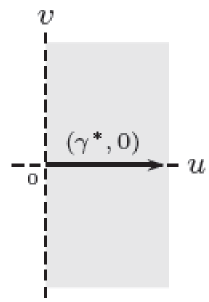
esto es, los multiplicadores (γ,λ) que verifican la desigualdad anterior, pertenecen al polar de $\mathrm{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$.

El siguiente resultado describe todas las situaciones vistos en el teorema 4.3, en este caso, una completa caracterización de los multiplicadores de Lagrange cuando se considera un problema de minimización cuadrática con una restricción cuadrática y lineal.

Proposición 6.1. *Sean* $f:\mathbb{R}^n \to \mathbb{R}$ $g:\mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, esto es,* $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$, $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$, $A,\ B\ \in\ S^n$, $a,\ b\ \in\ \mathbb{R}^n$, $\alpha,\ \beta\ \in\ \mathbb{R}$. *Además consideramos* μ *finito y* $C = H^{-1}(d)$, *en donde,* $H\ \in\ M^{m\times n}$, $d\ \in\ \mathbb{R}^m$. *Entonces*

a.1) *Si* $\mathrm{argmin}_K f \cap S^-_g(0) \neq \emptyset$ *ó* $S^+_f(\mu) \cap S^-_g(0) \neq \emptyset$, $r = -\infty$, *entonces* $S^-_f(\mu) = \emptyset$ *y como consecuencia*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \ \in \ \mathbb{R}^2 : \ u > 0\}$$

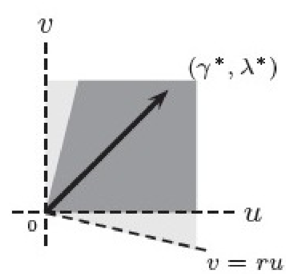


Además,

$$\exists \ (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \setminus \{(0,0)\} : \ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall \ x \ \in \ C \iff \gamma^* > 0, \ \lambda^* = 0;$$

a.2) *Si* $\text{argmin}_K f \cap S_g^-(0) = \emptyset$, $S_f^-(\mu) = \emptyset$ *y* $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, $-\infty < r < 0$, *entonces*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \ \in \ \mathbb{R}^2 : \ v > ru, \ u > 0\}$$

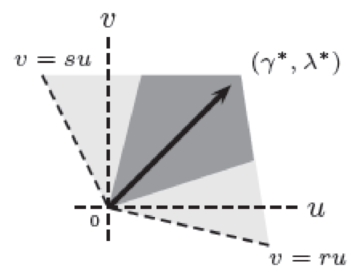


Además,

$$\exists \ (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \setminus \{(0,0)\} : \ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall \ x \ \in \ C \iff \gamma^* > 0, \ 0 \leq \lambda^* \leq -\frac{1}{r}\gamma^*;$$

a.3) *Si* $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$ *y* $S_f^-(\mu) \neq \emptyset$, *entonces* $s \leq r$ *y* $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, *como consecuencia:*

$$\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \ \in \ \mathbb{R}^2 : \ v > ru, \ v > su\}$$

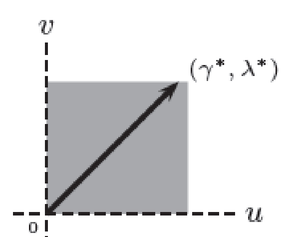


Aquí,

$$\exists\, (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \backslash \{(0,0)\} : \ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall\, x \ \in \ C \iff \gamma^* > 0, \ -\frac{1}{s}\gamma^* \leq \lambda^* \leq -\frac{1}{r}\gamma^*;$$

a.4) *Si* $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, $S_f^+(\mu) \cap S_g^-(0) = \emptyset = S_f^-(\mu)$, *entonces*

$$\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \ \in \ \mathbb{R}^2 : \ u > 0, \ v > 0\}$$

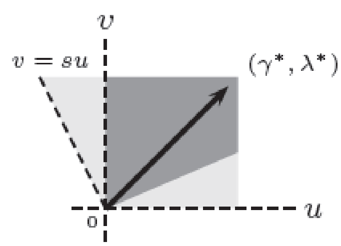


Además,

$$\exists\, (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \backslash \{(0,0)\} : \ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall\, x \ \in \ C \iff (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \backslash \{(0,0)\};$$

a.5) *Si* $S_f^+(\mu) \cap S_g^-(0) = \emptyset,\ S_f^-(\mu) \neq \emptyset,\ s < 0,$ *entonces* $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$, *como consecuencia,*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \ \in \ \mathbb{R}^2 : \ v > su, \ v > 0\}$$

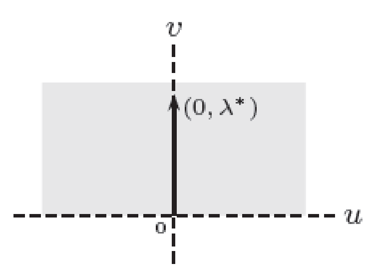


Además,

$$\exists \ (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \setminus \{(0,0)\} : \ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall \ x \ \in \ C \iff \gamma^* \geq 0, \ \lambda^* > -\frac{1}{s}\gamma^*;$$

a.6) *Si* $S_f^+(\mu) \cap S_g^-(0) = \emptyset$, $S_f^-(\mu) \neq \emptyset,\ s = 0,$ *entonces* $\operatorname{argmin}_K f \cap S_g^-(0) = \emptyset$ *como consecuencia,*

$$\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) = \{(u,v) \ \in \ \mathbb{R}^2 : \ v > 0\}$$



Además,

$$\exists \ (\gamma^*, \lambda^*) \ \in \ \mathbb{R}^2_+ \setminus \{(0,0)\} : \ \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \ \forall \ x \ \in \ C \iff \gamma^* = 0, \ \lambda^* > 0;$$

Demostración. **a.1)**

i) Si $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$. Entonces de la convexidad de $\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ se tiene que $S_f^-(\mu) = \emptyset$, luego

$$\begin{aligned} \operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \operatorname{cone}_+(F(\operatorname{argmin}_K f \cap S_g^-(0)) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) \in \mathbb{R}^2 : v > 0\} \end{aligned}$$

$\Rightarrow\rfloor$ Razonando por contradicción, supongamos que $\gamma^* = 0$, ó $\lambda^* > 0$. Como sigue que $\operatorname{argmin}_K f \cap S_g^-(0) \neq \emptyset$, existe $\overline{x}$ de modo tal que $g(\overline{x}) < 0$ y $f(\overline{x}) = \mu$. Luego $\gamma^*(f(\overline{x}) - \mu) + \lambda^* g(\overline{x}) = \lambda^* g(\overline{x}) \geq 0$, es decir, $g(\overline{x}) \geq 0$ lo cual es una contradicción.

$\Leftarrow\rfloor$ Supongamos que $\gamma^* > 0,\ \lambda^* = 0$. Luego probar que

$$\gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0, \forall\, x \in C,$$

se reduce a demostrar que $f(x) - \mu \geq 0,\ \forall\, x \in C$, lo cual es directo aplicando la proposición 4.1, pues como ya se demostró $S_f^-(\mu) = \emptyset$.

ii) Si $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$ con $r = -\infty$. En este caso la convexidad de $\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ asegura nuevamente que $S_f^-(\mu) = \emptyset$, de esta manera

$$\begin{aligned} \operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \operatorname{cone}_+(F(S_f^+(\mu) \cap S_g^-(0)) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) \in \mathbb{R}^2 : v > 0\} \end{aligned}$$

$\Rightarrow\rfloor$ Razonando por contradicción supongamos en un primer lugar que $\gamma^* = 0$. Entonces $\lambda^* > 0$, sigue que $\lambda^* g(x) \geq 0,\ \forall\, x \in C$, es decir, $g(x) \geq 0,\ \forall\, x \in C$. Por lo tanto $S_g^-(0) = \emptyset$ lo cual es una contradicción.

Si $\lambda^* > 0$, se sigue de la convexidad de $\operatorname{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$,

$$\frac{g(x)}{f(x) - \mu} \geq -\frac{\gamma^*}{\lambda^*},\ \forall\, x \in S_f^+(\mu) \cap S_g^-(0),$$

por lo cual, $r \geq -\dfrac{\gamma^*}{\lambda^*}$, lo cual es una contradicción.

$\Leftarrow\rfloor$ Bajo el supuesto que $\gamma^* > 0,\ \lambda^* = 0$. Ya sea $\text{argmin}_K f \cap S_g^-(0) \neq \emptyset$ ó $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ r = -\infty$, la convexidad de $\text{cone}(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ permite asegurar que $S_f^-(\mu) = \emptyset$. Por lo tanto de la proposición 4.1 $f(x) \geq \mu$, para todo $x \in C$. Así, $\gamma^*(f(x) - \mu) \geq 0,\ \forall\, x \in C$.

a.2) Suponiendo que $\text{argmin}_K f \cap S_g^-(0) = \emptyset$, $S_f^-(\mu) = \emptyset$ y $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset,\ -\infty < r < 0$, entonces

$$\begin{aligned} \text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \text{cone}_+(F(S_f^+(\mu) \cap S_g^-(0)) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) \in \mathbb{R}^2 :\ v > ru,\ u > 0\} \end{aligned}$$

$\Leftarrow\rfloor$ Supongamos que $\gamma^* > 0$ y $0 \leq \lambda^* \leq -\dfrac{1}{r}\gamma^*$. En el caso en que $\lambda^* = 0$ y usando el hecho que $S_f^-(\mu) = \emptyset$ se tiene que

$$\gamma^*(f(x) - \mu) + \lambda^* g(x) = \gamma^*(f(x) - \mu) \geq 0,\ \forall\, x \in C.$$

Si $0 < \lambda^* \leq -\dfrac{1}{r}$, consideremos $x \in C$ arbitrario. Sabemos que $S_f^-(\mu) = \emptyset$ por lo que $f(x) - \mu \geq 0$ por proposición 4.1. Se sigue que $g(x) \geq 0$ entonces $\gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0$.

Si $g(x) < 0$ entonces $f(x) - \mu > 0$ pues $\text{argmin}_K f \cap S_g^-(0) = \emptyset$. Luego

$$r \leq \frac{g(x)}{f(x) - \mu} \Leftrightarrow \frac{\gamma^*(f(x) - \mu)}{-g(x)} \geq -\frac{1}{r}\gamma^* \geq \lambda^*,$$

y así $\gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0,\ \forall\, x \in C$.

$\Rightarrow\rfloor$ Bajo las presentes hipótesis, sabemos que

$$\begin{aligned} \text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \text{cone}_+(F(S_f^+(\mu) \cap S_g^+(0)) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) \in \mathbb{R}^2 :\ v > ru,\ v > 0\} \end{aligned}$$

por lo cual $(\gamma^*, \lambda^*) \in \text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})^* = \{(u,v) \in \mathbb{R}^2_{++} :\ v \leq -\frac{1}{r}u\}$.

Entonces $\gamma^* > 0$ y $0 \leq \lambda^* \leq -\frac{1}{r}\gamma^*$.

a.3) Bajo el supuesto $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, $S_f^-(\mu) \neq \emptyset$, entonces de la convexidad de $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ se tiene que $\text{argmin}_K f \cap S_g^-(0) = \emptyset$ y $s \leq r$. Además

$$
\begin{aligned}
\text{cone}_+(F(C)-\mu(1,0)+\mathbb{R}^2_{++}) &= \text{cone}_+(F((S^+_f(\mu)\cap S^-_g(0))\cup S^-_f(\mu))-\mu(1,0)+\mathbb{R}^2_{++}) \\
&= \{(u,v)\ \in\ \mathbb{R}^2:\ v>ru,\ v>su\}
\end{aligned}
$$

y así $(\gamma^*,\lambda^*)\ \in\ \text{cone}_+(F(C)-\mu(1,0)+\mathbb{R}^2_{++})^*=\{(u,v)\ \in\ \mathbb{R}^2:\ -\frac{1}{s}u\leq v\leq -\frac{1}{r}u\}$.

Con lo cual $\gamma^*>0$ y $-\frac{1}{s}\gamma^*\leq\lambda^*\leq-\frac{1}{r}\gamma^*$.

$\Rightarrow\rfloor$ Supongamos que $\gamma^*>0$ y $-\frac{1}{s}\gamma^*\leq\lambda^*\leq-\frac{1}{r}\gamma^*$. Sea $x\ \in\ C$ arbitrario, con lo cual se dan los siguientes casos.

Si $f(x)-\mu\geq 0$ y $g(x)\geq 0$ entonces $\gamma^*(f(x)-\mu)+\lambda^*g(x)\geq 0$.

Si $f(x)-\mu<0$ y $g(x)>0$ entonces $x\ \in\ S^-_f(\mu)$, luego

$$\frac{g(x)}{f(x)-\mu}\leq s\leq-\frac{\gamma^*}{\lambda^*}\Leftrightarrow\gamma^*(f(x)-\mu)+\lambda^*g(x)\geq 0.$$

Si $f(x)-\mu>0$ y $g(x)<0$ entonces $x\ \in\ S^+_f(\mu)\cap S^-_g(0)$, luego

$$-\frac{\gamma^*}{\lambda^*}\leq r\leq\frac{g(x)}{f(x)-\mu}\Leftrightarrow\gamma^*(f(x)-\mu)+\lambda^*g(x)\geq 0.$$

a.4) Bajo los supuestos $\text{argmin}_K f\cap S^-_g(0)=\emptyset$, $S^+_f(\mu)\cap S^-_g(0)=\emptyset=S^-_f(\mu)$, entonces

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

y así $(\gamma^*,\lambda^*)\ \in\ \text{cone}_+(F(C)-\mu(1,0)+\mathbb{R}^2_{++})^*=\mathbb{R}^2_{++}$.

Con lo cual $\gamma^*\geq 0$ y $\lambda^*\geq 0$.

$\Leftarrow\rfloor$ Supongamos que $(\gamma^*,\lambda^*)\ \in\ \mathbb{R}^2_+\setminus\{(0,0)\}$ y sea $x\ \in\ C$ arbitrario. Se tiene que $f(x)-\mu\geq 0$ y $g(x)\geq 0$. De esta manera

$$\gamma^*(f(x)-\mu)+\lambda^*g(x)\geq 0,\ \forall\ x\ \in\ C$$

a.5) Bajo el supuesto $S_f^+(\mu) \cap S_g^-(0) = \emptyset$, $S_f^-(\mu) \neq \emptyset$, $s < 0$, entonces de la convexidad de $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ se tiene que $\text{argmin}_K f \cap S_g^-(0) = \emptyset$. Además

$$\begin{aligned}\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \text{cone}_+(F(S_f^-(\mu)) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) \in \mathbb{R}^2 : \ v > su, \ v > 0\}\end{aligned}$$

y así $(\gamma^*, \lambda^*) \in \text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})^* = \{(u,v) \in \mathbb{R}^2 : \ -\frac{1}{s}u \leq v, \ u > 0\}$.

Con lo cual $\gamma^* > 0$ y $-\frac{1}{s}\gamma^* \leq \lambda^*$.

$\Leftarrow\rfloor$ Supongamos que $\gamma^* > 0$ y $-\frac{1}{s}\gamma^* \leq \lambda^*$. Sea $x \in C$ arbitrario, con lo cual se dan los siguientes casos.

Si $f(x) - \mu \geq 0$ y $g(x) \geq 0$ entonces $\gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0$.

Si $f(x) - \mu < 0$ y $g(x) > 0$ entonces $x \in S_f^-(\mu)$, luego

$$\frac{g(x)}{f(x) - \mu} \leq s \leq -\frac{\gamma^*}{\lambda^*} \Leftrightarrow \gamma^*(f(x) - \mu) + \lambda^* g(x) \geq 0.$$

a.6) Bajo el supuesto $S_f^+(\mu) \cap S_g^-(0) = \emptyset$, $S_f^-(\mu) \neq \emptyset$, $s = 0$, entonces de la convexidad de $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ se tiene que $\text{argmin}_K f \cap S_g^-(0) = \emptyset$. Además

$$\begin{aligned}\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++}) &= \text{cone}_+(F(S_f^-(\mu)) - \mu(1,0) + \mathbb{R}^2_{++}) \\ &= \{(u,v) \in \mathbb{R}^2 : \ v > 0\}\end{aligned}$$

y así $(\gamma^*, \lambda^*) \in \text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})^* = \{(u,v) \in \mathbb{R}^2 : \ v > 0, \ u = 0\}$.

Con lo cual $\gamma^* = 0$ y $\lambda^* > 0$.

$\Leftarrow\rfloor$ Supongamos que $\gamma^* > 0$ y $\lambda^* > 0$. Sea $x \in \mathbb{R}^n$ arbitrario, con lo cual se dan los siguientes casos.

$\gamma^*(f(x) - \mu) + \lambda^* g(x) = \lambda^* g(x) \geq 0$.

$\square$

Con el objetivo de entregar en algún sentido una caracterización de la propiedad de Strong Duality para el problema considerado en esta sección, es que, previamente damos unos resultados que aprovechan las propiedades de las formas cuadráticas.

Proposición 6.2. *Supongamos que* $K = \{x \in C : g(x) \leq 0\} \neq \emptyset$*,* μ *finito. Sean* $f : \mathbb{R}^n \to \mathbb{R}$*,* $g : \mathbb{R}^n \to \mathbb{R}$ *funciones cuadráticas, en donde* $f(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$*,* $g(x) = \frac{1}{2}x^\top Bx + b^\top x + \beta$*,* $A,\ B\ \in S^n,\ a,\ b\ \in\ \mathbb{R}^n,\ \alpha,\ \beta\ \in \mathbb{R}$*. Entonces*

$$\forall\ v\ \in\ Ker\ H,\ v^\top Av \geq 0\ \text{ó}\ v^\top Bv \geq 0.$$

Demostración. Supongamos que existe $v \in \text{Ker } H$ tal que $v^\top Av < 0$ y $v^\top Bv < 0$. Entonces para algún $t_0 > 0$ $x + tv \in S_f^-(\mu) \cap S_g^-(0)$ para todo $|t| \geq t_0$, $x \in H^{-1}(d)$, lo que no puede ser pues $g(x + tv) < 0$, $H(x + tv) = Hx = d$ implica $f(x + tv) \geq \mu$.

□

Proposición 6.3. *Sea* $v \in Ker\ H$*,* $v^\top Bv < 0$*,* $v^\top Av = 0$*. Entonces*

a) *Existe* $t_1 \geq 0$ *tal que* $x + tv \in S_g^-(0)$*,* $\forall\ x \in H^{-1}(d)$*,* $\forall\ |t| \geq t_1$*.*

b) $\nabla f(x)^\top v = 0$*,* $\forall\ x \in H^{-1}(d)$ *y así* $f(x + tv) = f(x)\ \forall\ x \in H^{-1}(d)$*,* $\forall\ |t| \geq t_1$*.*

c) $S_f^-(\mu) = \emptyset$*.*

d) *Si* $S_f^+(\mu) \neq \emptyset \Longrightarrow r = -\infty$*.*

Demostración. **a)** Sea $x \in H^{-1}(d)$, luego para todo $t \in \mathbb{R}$ $g(x+tv) = g(x)+t\nabla g(x)^\top v+\frac{t^2}{2}v^\top Bv$. Como $v^\top Bv < 0$ entonces $g(x + tv) \longrightarrow -\infty$, cuando $|t| \to +\infty$.

En consecuencia, existe $t_1 \geq 0$ tal que $x + tv \in S_g^-(0)$, $\forall\ |t| \geq t_1$.

b) De a), $f(x + tv) \geq \mu\ \forall\ |t| \geq t_1$, pues $x + tv \in H^{-1}(d)$ y $g(x + tv) < 0$. Además $\mu \leq f(x+tv) = f(x)+t\nabla f(x)^\top v$, $\forall\ |t| \geq t_1$ entonces $\nabla f(x)^\top v = 0$ y así $f(x+tv) = f(x)$, $\forall\ |t| \geq t_1$.

c) Si $S_f^-(\mu) \neq \emptyset$. Sea $x_0 \in S_f^-(\mu)$, esto es, $f(x_0) < \mu$. De a), existe $t_1 > 0$ tal que $g(x_0+tv) < 0$ para todo $|t| \geq t_1$ y por lo tanto $f(x_0) = f(x_0 + tv) \geq \mu$, lo cual es una contradicción.

d) Sea $x_0 \in S_f^+(\mu)$, luego existe $t_1 \geq 0$ tal que $f(x_0 + tv) = f(x_0) > \mu$ para todo $|t| \geq t_1$. Además de a) sabemos que $x_0 + tv \in S_f^+(\mu) \cap S_g^-(0)$. De esta manera cuando $t \to +\infty$

$$r \leq \frac{g(x_0) + t\nabla g(x_0)^\top v + \frac{t^2}{2} v^\top B v}{f(x_0) - \mu} \longrightarrow -\infty.$$

$\square$

Proposición 6.4. *Sea $v \in Ker\ H$, $v^\top B v = 0$, $v^\top A v < 0$. Entonces*

a) *Existe $t_1 \geq 0$ tal que $x + tv \in S_f^-(\mu) \cap S_g^+(0)$, $\forall\ x \in H^{-1}(d)$, $\forall\ |t| \geq t_1$.*

b) *$\nabla g(x)^\top v = 0$, $\forall\ x \in H^{-1}(d)$ y así $g(x + tv) = g(x)$ $\forall\ x \in H^{-1}(d)$, $\forall\ |t| \geq t_1$.*

c) *$K = \emptyset$.*

d) *$s = 0$.*

Demostración. **a)** Sea $x \in H^{-1}(d)$, luego para todo $t \in \mathbb{R}$ $f(x+tv) = f(x) + t\nabla f(x)^\top v + \frac{t^2}{2} v^\top A v$. Como $v^\top A v < 0$ entonces $f(x + tv) \longrightarrow -\infty$, cuando $|t| \to +\infty$.

En consecuencia, existe $t_1 \geq 0$ tal que $x + tv \in S_f^-(\mu) \subseteq S_f^-(\mu) \cap S_g^+(0)$, $\forall\ |t| \geq t_1$.

b) De a), $0 < g(x + tv) = g(x) + t\nabla g(x)^\top v$, $\forall\ |t| \geq t_1$, entonces $\nabla g(x)^\top v = 0$, $\forall\ x \in H^{-1}(d)$ y así $g(x + tv) = g(x)$, $\forall\ |t| \geq t_1$.

c) Si $K \neq \emptyset$. Sea $x_0 \in K$, esto es, $g(x_0) \leq 0$. De a), existe $t_1 > 0$ tal que $f(x_0 + tv) < \mu$ para todo $|t| \geq t_1$ y por lo tanto $g(x_0) = g(x_0 + tv) > 0$, lo cual es una contradicción.

d) Sea $x_0 \in S_f^-(\mu)$, de a) existe $t_1 \geq 0$ tal que $x_0 + tv \in S_f^-(\mu) \cap S_g^+(0)$ para todo $|t| \geq t_1$. Además de b) sabemos que $g(x_0 + tv) = g(x_0)$. De esta manera cuando $t \to +\infty$

$$s \geq \frac{g(x_0)}{f(x_0) + t\nabla f(x_0)^\top v + \frac{t^2}{2} v^\top A v - \mu} \longrightarrow 0.$$

$\square$

El siguiente resultado, entrega por un lado una condición suficiente para Strong Duality y por otro, una condición necesaria, en términos de sistemas de desigualdades y en donde no depende de la condición de Slater.

Teorema 6.6. *Sea K no vacío y μ finito.*

a) $S_f^-(\mu) = \emptyset$ *ó* $[\mathrm{argmin}_K f \cap S_g^=(0) = \emptyset$, *para todo* $x_k \in H^{-1}(d) \cap S_f^-(\mu)$, *tal que* $\frac{x_k}{\|x_k\|} \to v$, *se tiene* $v^\top Bv = 0 \implies v^\top Av > 0]$;

b) *S.D se cumple;*

c) $S_f^-(\mu) = \emptyset$ *ó* $[S_f^-(\mu) \neq \emptyset,\ s < 0]$*;*

d) $\inf\limits_{x\in C} f(x) = \mu$ *ó* $[v \in \text{Ker } H,\ v^\top Bv < 0 \implies v^\top Av > 0]$.

Entonces, se satisface la siguiente relación:

$$a) \implies b) \iff c) \implies d).$$

Demostración. $b \Leftrightarrow c)\rfloor$ Es directo del teorema 4.4.

$a) \Rightarrow c)\rfloor$ Si $S_f^-(\mu) = \emptyset$, entonces no hay nada que demostrar. Si $S_f^-(\mu) \neq \emptyset$, Razonando por contrarecíproco. Supongamos que $s = 0$, entonces existe sucesión $\{x_n\}_{n\in\mathbb{N}}$ en $H^{-1}(d) \cap S_f^-(\mu)$, tal que

$$\lim_{n\to+\infty} \frac{g(x_n)}{f(x_n) - \mu} = 0.$$

Luego para cada $n \in \mathbb{N}$, $f(x_n) < \mu$ y $g(x_n) > 0$, con lo cual se presentan dos casos.

Caso 1. $\sup\limits_{n\in\mathbb{N}} \|x_n\| < +\infty$. Entonces existe una subsucesión convergente $\{x_{n_k}\}_{k\in\mathbb{N}}$ de $\{x_n\}_{n\in\mathbb{N}}$ a un punto x_0 en $\mathbb{R}^n$, para lo cual en el límite, $g(x_0) \geq 0$ y $f(x_0) \leq \mu$. Bajo estas condiciones tenemos que, si $g(x_0) = 0$ y $f(x_0) < \mu$ contradice el hecho que μ sea el ínfimo sobre K, además

$$\lim_{n\to+\infty} \frac{g(x_n)}{f(x_n) - \mu} = \begin{cases} \dfrac{g(x_0)}{f(x_0) - \mu} \neq 0 & , \text{ si } g(x_0) > 0 \text{ y } f(x_0) < \mu \\[2ex] -\infty & , \text{ si } g(x_0) > 0 \text{ y } f(x_0) = \mu \end{cases}$$

Por lo que necesariamente $g(x_0) = 0$ y $f(x_0) = \mu$.

Caso 2. $\sup\limits_{n\in\mathbb{N}} \|x_n\| = +\infty$. Entonces podemos suponer sin pérdida de generalidad que

$$\lim_{n\to+\infty} \frac{x_n}{\|x_n\|} = v.$$

Por lo cual, dividiendo por $\|x_n\|^2$,

$$\lim_{n\to+\infty} \frac{f(x_n)}{\|x_n\|^2} = v^\top Av$$

$$\lim_{n\to+\infty} \frac{g(x_n)}{\|x_n\|^2} = v^\top Bv$$

y además, en el límite $v^\top Bv \geq 0$ y $v^\top Av \leq 0$. De modo similar al caso anterior tenemos que

$$\lim_{n\to+\infty} \frac{g(x_n)}{f(x_n)-\mu} = \begin{cases} \dfrac{v^\top Bv}{v^\top Av} \neq 0 & \text{, si } v^\top Bv > 0 \text{ y } v^\top Av < 0 \\ \\ +\infty & \text{, si } v^\top Bv > 0 \text{ y } v^\top Av = 0 \end{cases}$$

Por lo tanto necesariamente se debe cumplir que $v^\top Bv = 0$ y $v^\top Av \leq 0$.

c) $\Rightarrow$ *d)*$\rfloor$ Si $S_f^-(\mu) = \emptyset$ entonces $f(x) \geq \mu$ para todo $x \in C$ por proposición 4.1. Asumimos entonces que $S_f^-(\mu) \neq \emptyset$ y $s < 0$. De la convexidad de $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$, se tiene que $\text{argmin}_K f \cap S_g^-(0) = \emptyset$. Consideramos dos casos: $S_g^-(0) = \emptyset$ ó $S_g^-(0) \neq \emptyset$. Obviamente en el primer caso, la condición c) se cumple vacuamente. Si $S_g^-(0) \neq \emptyset$, se sigue que $S_f^+(\mu) \cap S_g^-(0) \neq \emptyset$, pues en caso contrario se sigue de la proposición 4.1 que $S_g^-(0) = \emptyset$. Por lo tanto tenemos que $-\infty < s \leq r < 0$, esto, nuevamente gracias a la convexidad de $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$.

Razonando por contradicción. Supongamos que existe $v \in \text{Ker } H$, de modo que $v^\top Bv < 0$ y $v^\top Av \leq 0$. Por proposición 6.2, se considera el caso $v^\top Bv < 0$ y $v^\top Av = 0$. Para $x_0 \in H^{-1}(d)$ se obtiene

$$g(x_0 + tv) = g(x_0) + t\langle \nabla g(x_0), v\rangle + \frac{1}{2}t^2 v^\top Bv \longrightarrow -\infty, \text{ cuando } t \longrightarrow +\infty,$$

entonces existe $\bar{t} > 0$ de modo que $g(x_0 + tv) < 0$ para todo $t > \bar{t}$, así $x_0 + tv$ es factible y $f(x_0 + tv) - \mu > 0$ pues $\text{argmin}_K f \cap S_g^-(0) = \emptyset$. Por lo tanto, para todo $t > \bar{t}$, $tv \in S_f^+(\mu) \cap S_g^-(0)$. Sin embargo,

$$\frac{g(x_0 + tv)}{f(x_0 + tv) - \mu} = \frac{g(x_0) + t\langle \nabla g(x_0), v\rangle + \frac{1}{2}t^2 v^\top Bv}{f(x_0) + t\langle \nabla f(x_0), v\rangle - \mu} \longrightarrow -\infty, \text{ cuando } t \longrightarrow +\infty,$$

lo cual contradice el hecho que r sea finito.

$\square$

Los siguientes ejemplos, relacionados con el teorema 6.6, justifican el porqué con las proposiciones presentadas, no se tienen las equivalencias.

Ejemplo 6.3. *La implicancia c) $\Longrightarrow$ a) del teorema 6.6 no se tiene en general. Consideremos el ejemplo 6.1, $f(x,y) = x + y - (x+y)^2$, $g(x,y) = (x+y)^2 - 1$. En este caso, como parte de un problema de minimización*

$$\mu = \inf_{\substack{g(x,y)\leq 0 \\ (x,y)\in\mathbb{R}^2}} f(x,y),$$

y en donde primero mostraremos que μ es real. En efecto, para $\overline{\mu} = -2$ y $\lambda = \frac{3}{2}$ se tiene:

$$f(x,y)-\overline{\mu}+\lambda g(x,y) = x+y-(x+y)^2+2+\frac{3}{2}((x+y)^2-1) = x+y+\frac{1}{2}(x+y)^2+\frac{1}{2} = 2(x+y+1)^2 \geq 0,\ \forall\,(x,y)\ \in\ \mathbb{R}^2,$$

en particular para $(x,y) \in K$ se obtiene que $\mu \geq -2$. Además $f(-1,0) = \overline{\mu} = -2$ donde $(-1,0) \in S_g^{=}(0)$ De esta manera se concluye que $\mu = -2$ y $\operatorname{argmin}_K f \cap S_g^{=}(0) \neq \emptyset$. Por otro lado notemos que

$$\begin{aligned} S_f^-(\mu) &= \{(x,y) :\ (x+y)^2 - (x+y) - 2 > 0\} \\ &= \{(x,y) :\ x+y > 2\} \cup \{(x,y) :\ x+y < -1\} \neq \emptyset \end{aligned}$$

Con lo cual por teorema 6.1 a.3) se tiene que $s \leq r < 0$.

Ejemplo 6.4. *La implicancia d) $\Longrightarrow$ c) no se da en en general. Consideremos $f(x,y) = x + y$ y $g(x,y) = (x+y)^2$. Se tiene que $K = \{(0,0)\}$, $C = \mathbb{R}^2$, además $S_g^-(0) = \emptyset$ y de esta manera la condición d) se cumple vacuamente. Por otro lado $S_f^-(\mu) = \{(x,y) :\ x+y < 0\}$. Entonces*

$$\lim_{\substack{(x,y)\to(0,0) \\ (x,y)\in S_f^-(\mu)}} \frac{g(x,y)}{f(x,y)-\mu} = 0,$$

lo que implica que $s = 0$.

El siguiente ejemplo muestra que aún a falta de la condición de Slater, como esperabamos, Strong Duality se cumple para el caso cuadrático.

Ejemplo 6.5. *Consideramos* $H(x,y) = x - y,\ d = 0,\ f(x,y) = 2x^2 - y^2,\ g(x,y) = x^2 - y^2$. *Se tiene que* $K = \{(x,y) :\ x = y,\ x^2 - y^2 \leq 0\} = \{(t,t) :\ t \in \mathbb{R}\}$. *Claramente* $S_g^-(0) = \emptyset = S_f^-(\mu)$ *y* $\mu = 0$ *con* $\operatorname{argmin}_K f\{(0,0)\}$ *y de acuerdo con teorema 6.1 a.4) se tiene Strong Duality para todo* $\lambda^* \geq 0$.

El siguiente resultado, el cual es nuevo, entrega condiciones de optimalidad para el problema (P) analizado en esta sección en el cual abarca el caso no convexo.

Teorema 6.7. *Sean f y g funciones cuadráticas, μ finito $C = H^{-1}(d)$ y $\overline{x}$ un punto factible para* P. *Entonces las siguientes condiciones son equivalentes:*

a) $\overline{x}$ *es solución de* (P) *y* $S_f^-(\mu) = \emptyset$ *ó* $[S_f^-(\mu) \neq \emptyset,\ s < 0]$ *se cumple.*

b) *Existen* $\lambda^* \geq 0$ *e* $y \in \mathbb{R}^m$ *tales que* $\nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) + H^\top y = 0,\ \lambda^* g(\overline{x}) = 0,\ A + \lambda^* B$ *es copositiva en* $\operatorname{Ker} H$.

Demostración. $a) \Rightarrow b)\rfloor$ Del teorema 6.6, se cumple strong duality, esto es, existe $\lambda^* \geq 0$ tal que

$$f(\overline{x}) + \lambda^* g(\overline{x}) \leq f(\overline{x}) = \inf_{x \in C} (f(x) + \lambda^* g(x)).$$

Lo cual implica que $\lambda^* g(\overline{x}) = 0$ y además $\overline{x}$ es un mínimo para $L(x) = f(x) + \lambda^* g(x)$ sobre C, el cual, al ser afín se cumple

$$\langle \nabla f(\overline{x}) + \lambda^* g(\overline{x}), x - \overline{x} \rangle \geq 0,\ \forall\ x \in C.$$

Como $x - \overline{x} \in \operatorname{Ker} H$ para todo $x \in C$, se sigue que $\langle \nabla f(\overline{x}) + \lambda^* g(\overline{x}), x - \overline{x} \rangle = 0$ para todo $x - \overline{x} \in \operatorname{Ker} H$. Entonces $\nabla f(\overline{x}) + \lambda^* g(\overline{x}) \in (\operatorname{Ker} H)^\perp = H^\top(\mathbb{R}^m)$. Por lo tanto, existe $y \in \mathbb{R}^m$ tal que $\nabla f(\overline{x}) + \lambda^* g(\overline{x}) + H^\top y = 0$. Por otro lado del desarrollo de Taylor se tiene que $L(x) = L(\overline{x}) + \langle \nabla f(\overline{x}) + \lambda^* g(\overline{x}), x - \overline{x} \rangle + \frac{1}{2}(x - \overline{x})^\top (A + \lambda^* B)(x - \overline{x})$. Para todo $x \in C$. Como $L(\overline{x}) \leq L(x)$ y $\nabla f(\overline{x}) + \lambda^* g(\overline{x}) = -H^\top y$, se obtiene

$$0 \leq L(x) - L(\overline{x}) = -\langle H^\top y, x - \overline{x} \rangle + \frac{1}{2}(x - \overline{x})^\top (A + \lambda^* B)(x - \overline{x}) = \frac{1}{2}(x - \overline{x})^\top (A + \lambda^* B)(x - \overline{x}),$$

lo cual implica que $A+\lambda^* B$ es copositiva en Ker H.

b) ⇒ a)⌋ $L(x) = f(x) + \lambda^* g(x),\ x \in C$ entonces,

$$L(x) - L(\overline{x}) = \langle \nabla f(\overline{x}) + \lambda^* g(\overline{x}), x - \overline{x} \rangle + \frac{1}{2}(x-\overline{x})^\top (A+\lambda^* B)(x-\overline{x}).$$

Por hipótesis, existe $y \in \mathbb{R}^m$ tal que $\nabla f(\overline{x}) + \lambda^* g(\overline{x}) = H^\top y$, $\lambda^* g(\overline{x}) = 0$ y $A+\lambda^* B$ es copositiva en Ker H. Con estos supuestos la diferencia anterior es positiva con lo cual, para todo $x \in C$

$$0 \leq L(x) - L(\overline{x}) = f(x) + \lambda^* g(x) - f(\overline{x}) \leq f(x) - f(\overline{x}).$$

Lo cual demuestra que $\overline{x}$ es solución de (P).

□

Aplicando el teorema 6.6, volvemos a obtener la versión original [12, Teorema 3.4], dada por Moré, bajo las hipótesis de Slater y que la matriz asociada a la forma cuadrática de la restricción B se no nula.

Corolario 6.2. *Sean f y g funciones cuadráticas, μ finito. Asuma que $g(x_0) < 0$ para algún $x_0 \in \mathbb{R}^n$ y $H = 0, d = 0$. Si $\overline{x}$ es solución del problema* (P)*, entonces existe $\lambda^* \geq 0$ tal que $\nabla f(\overline{x}) + \lambda^* \nabla g(\overline{x}) = 0$, $\lambda^* g(\overline{x}) = 0$ y $A + \lambda^* B \succeq 0$.*

Demostración. En el caso $S_f^-(\mu) = \emptyset$, el resultado es consecuencia del teorema 6.7. Si $S_f^-(\mu) \neq \emptyset$, por chequear que $s < 0$. En efecto supongamos que $s = 0$. Entonces de la convexidad de $\text{cone}_+(F(C) - \mu(1,0) + \mathbb{R}^2_{++})$ se tiene (ver proposición 6.6 a.6)) que $S_f^+(\mu) \cap S_g^-(0) = \text{argmin}_K f \cap S_g^-(0) = \emptyset$, lo cual implica (ver proposición 4.1) que $S_g^-(0) = \emptyset$. Contradiciendo la hipótesis de Slater. Por lo tanto $s < 0$ y se concluye usando el teorema 6.7.

□

6.3. Existencia de mínimos para un problema cuadrático.

El siguiente resultado, entrega una versión particular del teorema dada por Frank and Wolfe [6], comprende un demostración corta, similar a la versión dada por Blum and Oettli [1]. En la versión original dada por Frank and Wolfe requiere el teorema de descomposición para poliedros convexos. En nuestro caso usamos un poliedro particular.

Teorema 6.8. *Sea* $h(x) = \frac{1}{2}x^\top Ax + a^\top x + \alpha$ *con* $A \in S^n$, $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. *Las siguientes proposiciones son equivalentes:*

a) $-\infty < \nu \doteq \inf\limits_{x \in H^{-1}(d)} h(x)$;

b) *A es copositiva en* Ker H *y* $[v^\top Av = 0,\ v \in \text{Ker } H \implies (Ax+a)^\top v = 0\ \forall\ x \in H^{-1}(d)]$;

c) *A es copositiva en* Ker H *y* $\text{argmin}_{H^{-1}(d)} h \neq \emptyset$;

d) *A es copositiva en* Ker H *y existe* $\overline{x} \in H^{-1}(d)$, $y \in \mathbb{R}^m$ *tal que* $A\overline{x} + b + H^\top y = 0$.

Demostración. *a)* $\Rightarrow$ *b)*$\rfloor$ Primero probemos que A es copositiva en Ker H. Sea $x \in H^{-1}(d)$ y $v \in$ Ker H cualquiera. Por hipótesis $-\infty < \nu \doteq \inf\limits_{x \in H^{-1}(d)} h(x)$. Haciendo el desarrollo de Taylor, tenemos

$$h(x+tv) = h(x) + t\langle \nabla h(x), v\rangle + \frac{t^2}{2}v^\top Av \geq \nu,\ \forall\ t \in \mathbb{R}. \tag{6.2}$$

Dividiendo por $t^2 > 0$ en (6.2) se obtiene,

$$\frac{1}{t^2}h(x) + \frac{1}{t}\langle \nabla h(x), v\rangle + \frac{1}{2}v^\top Av \geq \frac{\nu}{t^2},$$

haciendo tender $t \to +\infty$, se obtiene $v^\top Av \geq 0$, $\forall\ v \in$ Ker H, lo cual demuestra la copositividad de A. Por otro lado sea $v \in$ Ker H tal que $v^\top Av = 0$. Entonces de (6.2) se tiene

$$\left.\begin{array}{l} \frac{1}{t}h(x) + \langle \nabla h(x), v\rangle \geq \frac{\nu}{t},\ \forall\ t > 0 \\[2ex] \frac{1}{t}h(x) + \langle \nabla h(x), v\rangle \leq \frac{\nu}{t},\ \forall\ t < 0 \end{array}\right. \implies (Ax+a)^\top v = 0$$

b) $\Rightarrow$ *c)*$\rfloor$ Para cada $k \in \mathbb{N}$, se considera el problema

$$
\begin{array}{ll}
(\mathrm{P_k}) & \text{mín}\, f(x) \\
 & x \in B_k \doteq \{x \ \in \ H^{-1}(d) : \ \|x\| \leq k\}
\end{array}
$$

el cual posee solución para cada $k \ \in \ \mathbb{N}$. Sea $x_k \ \in \ \text{argmin}_{B_k} h$ de norma mínima, es decir $\|x_k\| = \text{mín}\{\|x\| \ \in \ \text{argmin}_{B_k} h\}$.

Caso 1. $\sup\limits_{k\in\mathbb{N}} \|x_k\| < \infty$.

Entonces existe $x_{k_l} \longrightarrow \overline{x}$, cuando $l \to \infty$. Veamos que $\overline{x} \ \in \ \text{argmin}_{H^{-1}(d)} h$.

Sea $x \ \in \ H^{-1}(d)$ y $l_0 \ \in \ \mathbb{N}$ tal que $\|\overline{x}\| < k_{l_0}$. Luego $\|x\| < k_l$ para todo $l \geq l_0$. Esto implica

$$
h(x_{k_l}) \leq h(x), \ \forall \ l \geq l_0.
$$

De donde, debido a la continuidad de h, se obtiene

$$
h(\overline{x}) = \lim_{l\to\infty} h(x_{k_l}) \leq h(x),
$$

y como $x \ \in \ H^{-1}(d)$ es arbitrario, se concluye $\overline{x} \ \in \ \text{argmin}_{H^{-1}(d)} h$.

Caso 2. $\sup\limits_{k\in\mathbb{N}} \|x_k\| = \infty$.

Podemos suponer que $\|x_k\| \longrightarrow \infty$ y $\dfrac{x_k}{\|x_k\|} \longrightarrow v \ \in \ \text{Ker}\, H$ cuando $k \to \infty$.

Obiamente, dado $x \ \in \ H^{-1}(d)$, existe $k_0 \ \in \ \mathbb{N}$ tal que $h(x_k) = \frac{1}{2} x_k^\top A x_k + a^\top x_k + \alpha \leq h(x), \ \forall \ k \geq k_0$.

De donde, $v \top Av \leq 0$. Como A es copositiva, se obtiene $v^\top Av = 0$. Por hipótesis b) se tiene que $(Ax+v)^\top v = 0$ para todo $x \ \in \ H^{-1}(d)$. Por otro lado, $\dfrac{x_k}{\|x_k\|} \longrightarrow v$ implica que para algún $k_1 \ \in \ \mathbb{N}$

$$
\|x_k - \|x_k\| v\| < \|x_k\|, \ \forall \ k \geq k_1.
$$

Sea $u_k \doteq x_k - \|x_k\| v$, luego $u_k \ \in \ H^{-1}(d)$, $\|u_k\| < \|x_k\|$ y

$$
\begin{aligned}
h(u_k) &= h(x_k - \|x_k\|v) \\
&= h(x_k) - \|x_k\|(Ax_k + a)^\top v + \|x_k\|^2 v^\top Av \\
&= h(x_k) - \|x_k\|(x_k^\top Av + a^\top v) \\
&= h(x_k) - \|x_k\| a^\top v \\
&= h(x_k)
\end{aligned}
$$

De aquí, $u_k \in \operatorname{argmin}_{B_k} h$ para un k suficientemente mayor, lo cual no puede suceder pues

$$x_k \in \operatorname{argmin}_{\operatorname{argmin}_{B_k} h} \|\cdot\| \text{ y } \|u_k\| < \|x_k\|.$$

En consecuencia el caso 2 no puede ocurrir y por lo tanto $\operatorname{argmin}_{H^{-1}(d)} h \neq \emptyset$.

c) $\Rightarrow$ *d)*$\rfloor$ Sea $\overline{x} \in \operatorname{argmin}_{H^{-1}(d)} h$. Entonces de la condición necesaria de optimalidad, tenemos que $\langle \nabla h(\overline{x}), x - \overline{x} \rangle \geq 0$ para todo $x \in H^{-1}(d)$. Dado que $x - \overline{x} \in \text{Ker } H$ para todo $x \in H^{-1}(d)$, tenemos que $A\overline{x} + a = \nabla h(\overline{x}) \in (\text{Ker } H)^\perp = H^\top(\mathbb{R}^m)$. Por lo tanto existe $y \in \mathbb{R}^m$ tal que $A\overline{x} + a + H^\top y = 0$.

d) $\Rightarrow$ *a)*$\rfloor$ Es directo. Notemos que $H^{-1}(d) = \overline{x} + \text{Ker } H$ y

$$h(x + \overline{x}) = h(\overline{x}) + \langle \nabla h(\overline{x}), x \rangle + \frac{1}{2} x^\top Ax, \ x \in \text{Ker } H.$$

$\square$

Cuando H es la matriz nula y $d = 0$, el teorema previo admite una formulación más precisa, expresado en el siguiente corolario.

Corolario 6.3. *Sea* $h(x) = \frac{1}{2} x^\top Ax + a^\top x + \alpha$ *con* $A \in S^n$, $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. *Las siguientes condiciones son equivalentes:*

a) $-\infty < \nu \doteq \inf_{x \in \mathbb{R}^n} h(x)$*;*

b) $A \succeq 0$ *y* $[v \in \text{Ker } A \implies a^\top v = 0]$;

c) $A \succeq 0$ *y* $\operatorname{argmin}_{\mathbb{R}^n} h \neq \emptyset$;

d) $A \succeq 0$ *y existe* $\overline{x} \in \mathbb{R}^n$, *tal que* $A\overline{x} + b = 0$

Capítulo 7

Conclusiones y perspectivas

En cuanto a los objetivos planteados en este trabajo, en lo que respecta a la propiedad de Strong Duality el Teorema 6.6, da evidencia de que existe una relación con la solubilidad de un sistema de inecuaciones, gracias a una condiciones suficientes para la proposiciones $s < 0$ y $r = -\infty$. Queda propuesto unificar en solo una condición la caracterización, lo cual significa caracterizar el hecho que $s < 0$.

Las condiciones de optimalidad obtenidas en el Teorema 6.7, nos permiten extender desde el caso homogéneo [5, Teorema 5.1] al caso cuadrático no-homogéneo. Además generaliza el resultado obtenido por Moré [12, Teorema 3.4], en el sentido que se obtienen las mismas condiciones de optimalidad sin pedir la hipótesis de no nulidad a la matriz B.

Relevante para la obtención de nuestros objetivos, resultó el Teorema 6.5, pues la convexidad del conjunto $\mathrm{cone}(F(C) - (\mu, 0) + \mathbb{R}^2_{++})$ es una hipótesis necesaria y suficiente para asegurar Strong Duality, además permitió obtener una descripción más eficiente, del conjunto solución de los multiplicadores de Lagrange.

Cuando aseguramos que el teorema 6.8 entrega una versión más fuerte, ya que, la existencia de mínimos se hace sobre el subespacio vectorial afín $H^{-1}(d)$, que a diferencia del resultado obtenido sobre un poliedro convexo por Frank and Wolfe en [6], entrega mayor información.

Si bien es cierto, las condiciones de optimalidad obtenidas por V. Jeyakumar and Guoyin Li en [11, Teorema 2.1], caracteriza Strong Duality asociado al Lagrangiano regularizado, en el caso $\mathrm{argmin}_K f \neq \emptyset$ para un problema con m restricciones cuadráticas y una lineal, por consiguiente uno podría entender que este caso esta ya esta resuelto. Sin embargo, dado el éxito obtenido en [5] en el cual se obtiene una completa caracterización de Strong Duality para un problema general con una restricción, es aún, un problema abierto el resolver el caso general con m restricciones debido a la complejidad (a priori) en describir el conjunto $\mathrm{cone}(F(C) - (\mu, 0) + \mathbb{R}^{m+1}_{++})$.

Bibliografía

[1] E. Blum, W. Oettli, Direct proof of the existence theorem for quadratic programming, *Operation Research*, **20** (1972), 165–167.

[2] R.I. Bot, G. Wanka, An alternative formulation for a new closed cone constraint qualification, *Nonlinear Analysis* **64** (2006), 1367–1381.

[3] F. Flores-Bazán, Existence Theory for Finite-Dimensional Pseudomonotone Equilibrium Problems, *Acta Applicandae Mathematicae,* **77** (2003), 249–297.

[4] 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, 2012, Vol 1, 29–59.

[5] F. Flores-Bazán, Fernando Flores-Bazán, Cristián Vera, A complete characterization of strong duality in nonconvex optimization with a single constraint, *J. Global Optim.,* DOI 10.1007/s10898-011-9673-6(2011).

[6] M. Frank, P. Wolfe, An algorithm for quadratic programming, *Naval Res. Log. Quart.*, **3** (1956), 95–110.

[7] V. Jekakumar, Constraint qualification characterizing Lagrangian duality in convex optimization, *J. of Optim. Theory and Appl.,* **136** (2008), 31–41.

[8] V. Jekakumar, G. M. Lee, Complete characterizations of stable Farkas'lemma and cone-convex programming duality, *Math. Program.* **114** (2008), 335–347.

[9] V. Jekakumar, N.Q. Huy, G, Y. LI, Necessary and sufficient conditions for S-lemma and nonconvex quadratic optimization, *Optim. Eng.,* **10** (2009), 491–503.

[10] V. Jekakumar, G. M. Lee, and G. Y. LI Alternative theorems for quadratic inequality systems and global quadratic optimization, *SIAM, J. Optim.,* **20** (2009), 983–1001.

[11] V. Jeyakumar, Guoyin Li, Regularized Lagrangian duality for linearly constrained quadratic optimization and trust-region problems, *J. Global Optim.,* **49** (2011), 1–14.

[12] J. J. Moré, Generelization of the trust region problem, *Optim. Methods Software*, **2** (1993), 189–209.

[13] E. Matskani, N. D. Sidiropoulos, Z.-Q. Lou, L. Tassiulas, Convex Approximation Techniques for Joint Multiuser Downlink Beamforming and Admission Control, *IEEE Trans. on Wireless Communications,* **7** (2008), 2682–2693.

[14] JI-Ming Peng and Ya-Xiang Yuan, Optimality Conditions for the Minimization of a quadratic with two quadratic constraints, *SIAM, J. Optim.,* **2** (1997), 579–594.

[15] Imre Pólik, Tamás Terlaky, A Survey of the S-lemma, *SIAM, J. Optim.,* **49** (2007), 371–418.

[16] X. J. Zheng, X. L. Sun, D. Li, Y. F. Xu, On zero duality gap in nonconvex quadratic programming problems, *J. Global Optim.,* (2011) **52(2)** (2012), 229–242.

[17] Qingwei Jina, Shu-Cherng Fangab, Wenxun Xingb, On the global optimality of generalized trust region subproblems, *Optimization,* **59** (2010), 1139–1151.

[18] N. D. Sidiropoulos, T. N. Davidson, Z.-Q. Lou, *Transmit Beamforming for Physical-Layer Multicasting, IEEE Trans. Signal Process, Wireless Communications,* **54** (2006), 2239–2251.