Universidad del Bío Bío
Facultad de Educación y Humanidades
Escuela de Pedagogía en Educación Matemática

Reglas de Multiplicadores y el Teorema de Karush Kuhn Tucker

Autor: Juan Ernesto Soto Figueroa
Autor: Valeska Nicol Mena Reyes
Profesor Guía: Roberto Carlos Cabrales

MEMORIA PARA OPTAR AL TÍTULO DE PROFESOR DE EDUCACIÓN MEDIA EN EDUCACIÓN MATEMÁTICA

CHILLÁN, 2008

Índice general

Capítulo 1

Introducción

Antes de comenzar formalmente a hablar de las condiciones del Teorema de Karush Kuhn Tucker, deseamos brindar al lector una breve reseña sobre aquellos matemáticos que dieron vida a este teorema. De esta manera, tenemos que a Albert Tucker, nacido en Canadá y destacado profesor de la Universidad de Princeton, y Harold Kuhn, quien trabajase desde la Universidad con Tucker sin dimensionar - quizás - la teoría que junto a W. Karush lograron realizar.

Los importantes matemáticos mencionados anteriormente trabajaron en la teoria de programación no lineal, es de este modo que surge el teorema que lleva sus nombres.

Este teorema se encuentra precedido por una importante herramienta en el campo de la optimización, en el cual nos podemos encontrar con problemas que necesiten resolverse mediante el Método de los Multiplicadores de Lagrange, como por ejemplo:

Sea

$$\begin{aligned} f(\overrightarrow{x}) : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \overrightarrow{x} &\longmapsto f(\overrightarrow{x}) = x + y + z \end{aligned}$$

Sujeto a las restricciones $g_0(\overrightarrow{x}) = x^2 + y^2 = 2,$ y $g_1(\overrightarrow{x}) = x + z = 1$

Se sabe que para hacer uso del método de los Multiplicadores de Lagrange se tiene que encontrar los gradientes de f y de las restricciones, g_0, g_1. Este método nos indica que existirán $\lambda_0, \lambda_1 \in \mathbb{R}$ de tal forma que

$$\nabla f(\overrightarrow{x}) = \lambda_0 \cdot \nabla g_0(\overrightarrow{x}) + \lambda_1 \cdot \nabla g_1(\overrightarrow{x})$$

y resolviendo las ecuaciones encontradas de esta igualdad se puede concluir que los extremos o puntos donde la función alcanza su máximo son $P_1(0, \sqrt{2}, 1); P_2(0, -\sqrt{2}, 1)$. Cuando tenemos problemas de programación no lineal, el método de los Multiplicadores de Lagrange se extiende en el teorema de Karush Kuhn Tucker.

Es por esto que nuestro Seminario abordará la temática de optimización, de forma particular estudiando el teorema de Karush - Kuhn - Tucker (KKT), que se utiliza para la localización de soluciones óptimas de problemas con restricciones. Además, durante el transcurso de este seminario procederemos a demostrarlo, de acuerdo con los Teoremas de Separación de Hanh Banach principalmente.

Podemos comentar que entre las aplicaciones que se realizan de este teorema se encuentra su uso en los métodos estadísticos como el método del cordón de regresión que trata de ajustar una aplicación lineal a una nube de puntos, estableciendo la mínima distancia entre los puntos y la aplicación lineal de tal manera que la cantidad de elementos que no pertenezcan al recorrido de la aplicación sea extremadamente pequeña. Asimismo podemos mencionar el método de bioequivalencia, que se trata de la comparación de dos tratamientos que tienen similares caracteríaticas de absorción.

Para proceder a dicho estudio es necesario contar con herramientas de cálculo diferencial (derivada direccional, primera variación según Lagrange, derivada de Gâteaux, derivada de Fréchet, ect), conceptos básicos de la teoría de optimización y reafirmar las nociones adquiridas en el desarrollo de los diversos cursos de análisis matemático realizados durante nuestro periodo formativo. Asimismo, necesitaremos conocer estrategias de resolución tales como el método del menor cuadrado.

En el siguiente seminario pretendemos estudiar el teorema de KKT para proceder a su posterior demostración, para lograr una mejor asimilación de los conceptos que están implicados en él, que implica realizar un recuento sobre algunos tópicos de análisis y de álgebra lineal. Asimismo, la demostración de este teorema será de vital importancia para poder dimensionar su real importancia y su aplicabilidad en la resolución de problemas de optimización.

Capítulo 2

Derivadas

Antes de comenzar a hablar acerca del Teorema de Karush Kuhn Tucker y de las Reglas de Multiplicadores relacionadas, es necesario conocer algunas herramientas de cálculo diferencial que serán utilizadas en algunas de las demostraciones presentes en nuestro seminario, puesto que - dependiendo del tipo de derivada que posea - otorgará algunas propiedades que se describen a continuación y que permiten probar algunas propiedades que se encuentran en función de estos. Es fruto de esto que en este capítulo nosotros desarrollaremos los distintos tipos de derivadas que son posibles encontrar en un subconjunto U de $\mathbb{R}^n$.

Por ello, y para todas estas subsecciones, consideraremos U un subconjunto abierto de $\mathbb{R}^n$ y además tendremos que $\mathbf{F} = (F_1, F_2, \ldots, F_m) : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es una función.

2.1. Derivadas direccionales

Definición 2.1.1 *Si el límite*

$$\lim_{\lambda \to 0^+} \frac{\mathbf{F}(\overrightarrow{\mathbf{x_0}} + \lambda \overrightarrow{\mathbf{h}}) - \mathbf{F}\overrightarrow{\mathbf{x_0}}}{\lambda}$$

existe, entonces diremos que $\mathbf{F}$ *tiene derivada direccional en el punto* x_0 *según la dirección* $\overrightarrow{\mathbf{h}}$*, la cual denotaremos como* $\mathbf{F}'(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}})$

Ejemplo 2.1.1 *Consideremos la siguiente función:*

$\mathbf{F} : \mathbb{R} \to \mathbb{R}$

$x \mapsto \mathbf{F}(x) = |x|$

En este caso, tendremos que:

- *Si* $x_0 = 0 \quad \Rightarrow$

$$\mathbf{F}'(0, h) = \lim_{\lambda \to 0^+} \frac{\mathbf{F}(\lambda h) - \mathbf{F}(x_0)}{\lambda} = \lim_{\lambda \to 0^+} \frac{|\lambda h|}{\lambda} = \lim_{\lambda \to 0^+} |h| = |h|$$

- *Si* $x_0 > 0 \quad \Rightarrow$

$$\mathbf{F}'(x_0, h) = \lim_{\lambda \to 0^+} \frac{\mathbf{F}(x_0 + \lambda h) - \mathbf{F}(x_0)}{\lambda} = \lim_{\lambda \to 0^+} \frac{|x_0 + h - \mathbf{F}(x_0)|}{\lambda} =$$

$$\lim_{\lambda \to 0^+} \frac{x_0 + \lambda h - x_0}{\lambda} = h$$

- *Si* $x_0 < 0 \Rightarrow -x_0 > 0$ *y tendremos además que:*

$$\mathbf{F}'(x_0, h) = \lim_{\lambda \to 0^+} \frac{\mathbf{F}(x_0 + \lambda h) - \mathbf{F}(x_0)}{\lambda} = \lim_{\lambda \to 0^+} \frac{|x_0 + \lambda h| - |x_0|}{\lambda} =$$

$$\lim_{\lambda \to 0^+} \frac{-x_0 - \lambda h - (-x_0)}{\lambda} = \frac{-\lambda h}{\lambda} = -h$$

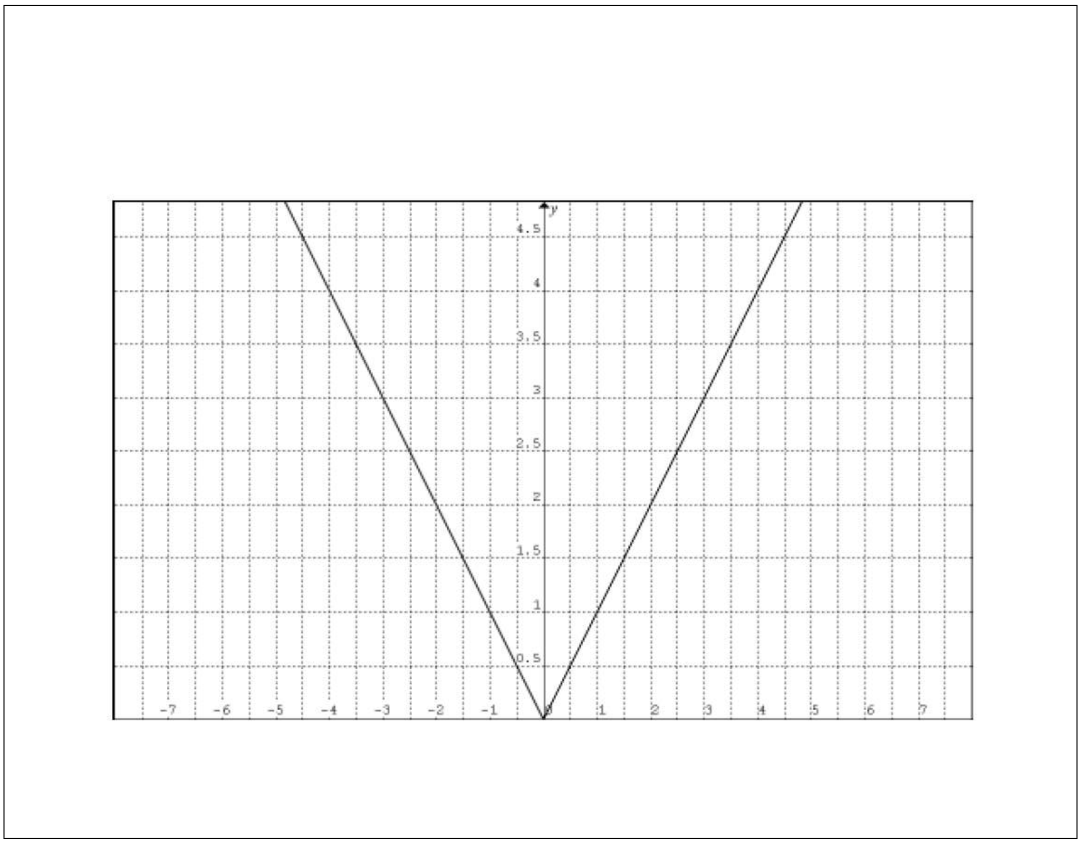

Figura 2.1: Gráfico de la función Valor Absoluto

Luego, al derivar , tendremos que

$$F'(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \\ \nexists & x = 0 \end{cases}$$

Por lo tanto

$$\mathbf{F}'(\mathbf{x}) = \begin{cases} h & x > 0 \\ -h & x < 0 \\ |h| & x = 0 \end{cases}$$

2.2. Primera variación

Definición 2.2.1 *Supongamos que para toda dirección* $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$*, la derivada* $\mathbf{F}'(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}})$ *existe. Si esto ocurre, es posible definir la siguiente aplicación:*

$\delta_+\mathbf{F}(\overrightarrow{\mathbf{x_0}}, \cdot) : \mathbb{R}^n \to \mathbb{R}^m$*, tal que*

$$\overrightarrow{\mathbf{h}} \mapsto \delta_+\mathbf{F}(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}}) = \mathbf{F}'(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}})$$

Se denomina como la primera variación de la aplicación $\mathbf{F}$ *en el punto* $\overrightarrow{\mathbf{x_0}}$*, o de otra forma, se tiene que* $\mathbf{F}$ *posee una primera variación en el punto* $\overrightarrow{\mathbf{x_0}}$

2.2.1. Primera variación según Lagrange

Definición 2.2.2 *Si la aplicación* $\mathbf{F}$*, además de poseer primera variación, cumple la siguiente característica*

$$\delta_+\mathbf{F}(\overrightarrow{\mathbf{x_0}}, -\overrightarrow{\mathbf{h}}) = -\delta_+\mathbf{F}(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}}),$$

para toda dirección $\overrightarrow{\mathbf{h}}$*. O visto de otra manera, si tenemos que para toda dirección* $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$*, el límite*

$$\delta\mathbf{F}(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}}) = \lim_{\lambda \to 0} \frac{\mathbf{F}(\overrightarrow{\mathbf{x_0}} + \lambda\overrightarrow{\mathbf{h}}) - \mathbf{F}(\overrightarrow{\mathbf{x_0}})}{\lambda}$$

existe, entonces nos encontraremos con que $\mathbf{F}$ *es denominada como una aplicación con* <u>*Primera Variación según Lagrange*</u> *de la aplicación* $\mathbf{F}$ *en el punto* $\overrightarrow{\mathbf{x_0}}$

2.3. Derivada de Gâteaux

Definición 2.3.1 *Consideremos la aplicación* $\mathbf{F}$ *posee una primera variación en el punto* $\overrightarrow{\mathbf{x_0}}$ *y que además existe un operador lineal* $\Lambda : \mathbb{R}^n \to \mathbb{R}^m$ *de tal forma que* $\delta_+\mathbf{F}(\overrightarrow{\mathbf{x_0}}, \overrightarrow{\mathbf{h}}) = \Lambda\overrightarrow{\mathbf{h}}$*. Entonces, tendremos que el operador* Λ *es denominado como la derivada de Gâteaux de la aplicación* $\mathbf{F}$ *en el punto* $\overrightarrow{\mathbf{x_0}}$ *y se escribe como* $\mathbf{F}'_G(\overrightarrow{\mathbf{x_0}}) = \Lambda$

De esta forma, tendremos que $\mathbf{F}'_G(\overrightarrow{\mathbf{x_0}})$ *es un elemento que pertenece a los operadores* $\mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m) = \{\phi : \mathbb{R}^n \to \mathbb{R}^m lineal\}$ *tal que para cada* $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$ *se tenga la siguiente relación:*

$$\mathbf{F}(\overrightarrow{\mathbf{x_0}} + \lambda\overrightarrow{\mathbf{h}}) = \mathbf{F}(\overrightarrow{\mathbf{x_0}}) + \mathbf{F}'_G(\overrightarrow{\mathbf{x_0}})\overrightarrow{\mathbf{h}} + o(\lambda)$$

Siendo $o(\lambda)$ *una cantidad sumamente pequeña cuando* λ*, es decir,*$o(\lambda)$ *si* $\lambda \to 0^+$*.*

2.4. Derivada de Fréchet

Definición 2.4.1 *Supongamos que una vecindad del punto* $\overrightarrow{\mathbf{x_0}}$ *puede ser representada de la siguiente forma:*

$$F(\overrightarrow{\mathbf{x_0}} + \overrightarrow{\mathbf{h}}) = F(\overrightarrow{\mathbf{x_0}}) + \Lambda \overrightarrow{\mathbf{h}} + \alpha(\overrightarrow{\mathbf{h}})||\overrightarrow{\mathbf{h}}||_{\mathbb{R}^n}$$

Donde se tiene que Λ *pertenece al conjunto de los operadores lineales de* $\mathbb{R}^n$ *a* $\mathbb{R}^m$*, es decir,* $\Lambda \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$ *y donde también se tiene que*

$$\lim_{||\overrightarrow{\mathbf{h}}||_{\mathbb{R}^n} \to 0} ||\alpha(\overrightarrow{\mathbf{h}})||_{\mathbb{R}^n} = ||\alpha(0)||_{\mathbb{R}^n} = 0$$

Si esto ocurre, entonces tendremos que $F(\cdot)$ *es diferenciable según Fréchet en el punto* $\overrightarrow{\mathbf{x_0}}$*. En cuanto al operador* Λ*, tenemos que este se denomina como Derivada de Fréchet de la aplicación* F *en el punto* $\overrightarrow{\mathbf{x_0}}$ *y se denota como* $F'(\overrightarrow{\mathbf{x_0}})$

Observación 2.4.1 *Si se tiene que una función* F *es diferenciable según Fréchet, entonces se tiene que la función es continua en el punto* $\overrightarrow{\mathbf{x_c}} \in \mathbb{R}^n$

2.5. Derivada Estricta

Definición 2.5.1 *Sea* F *una aplicación y* $\Lambda \in \mathfrak{L}(\mathbb{R}^n, \mathbb{R}^m)$*. Entonces,* F *es denominada estrictamente diferenciable en el punto* $\overrightarrow{\mathbf{x_c}}$ *si existe* Λ *tal que para todo* $\epsilon > 0$*, exista* $\delta > 0$ *tal que para todos los* $\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{x_2}}$ *se verifiquen las desigualdades* $||\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_c}}|| < \delta, ||\overrightarrow{\mathbf{x_2}} - \overrightarrow{\mathbf{x_c}}|| < \delta$ *implican que se cumpla:*

$$||F(\overrightarrow{\mathbf{x_1}}) - F(\overrightarrow{\mathbf{x_2}}) - \Lambda(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}})|| \leq \epsilon||\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}||$$

Observación 2.5.1 *El hecho de que una función* F *sea estrictamente diferenciable en un punto* $\overrightarrow{\mathbf{x_c}} \in U \subset \mathbb{R}^n$ *implica que la función no solamente es continua en el punto* $\overrightarrow{\mathbf{x_c}}$*, sino que también lo es en una vecindad de este.*

Capítulo 3

Teorema de Hanh-Banach

Definamos las siguientes funciones:

$$f_0, f_1, ..., f_p, f_{p+1}, ..., f_{p+q} : U \subset \mathbb{R}^n \to \mathbb{R}$$

donde $U \subset \mathbb{R}^n$ es un conjunto abierto y donde $p, q \in \mathbb{Z}_+ \cup \{0\}$. Además deseamos encontrar $\overrightarrow{\mathbf{x_c}} \in U$ de tal modo que minimice f_0 sujeta a p restricciones de desigualdad $(f_1 \leq 0, ..., f_p \leq 0)$ y a q restricciones de igualdad $(f_{p+1} = 0, ..., f_{p+q} = 0)$.

Estableceremos una función F definida desde $\mathbb{R}^n \to \mathbb{R}^{p+q+1}$ de la siguiente manera

$$F = (f_0, f_1, ..., f_p, f_{p+1}, ..., f_{p+q}) : U \subset \mathbb{R}^n \to \mathbb{R}^{p+q+1}$$

Esta definición nos permitirá establecer la condición de optimalidad de la función objetivo, condición que quedará caracterizada por la intersección de dos conjuntos. De esto se pueden obtener las reglas de multiplicadores separando estos dos conjuntos convexos relacionados.

Lema 3.0.1 $\overrightarrow{\mathbf{x_c}} \in U$ *es una solución óptima sí y solamente sí*

$$F(U) \cap W = \emptyset$$

donde

$$W = \{\overrightarrow{\mathbf{y}} = (\eta_0, \eta_1, ..., \eta_{p+q}) \in \mathbb{R}^{p+q+1} : \eta_0 < f_0(\overrightarrow{\mathbf{x_c}}),\ \eta_i \leq 0,\ i = 1, ..., p, \eta_i = 0,\ i = p+1, ..., p+q\}$$

Demostración: En primer término demostraremos la primera implicancia, es decir, sea $\overrightarrow{\mathbf{x_c}} \in U$ una solución óptima, entonces se ha de demostrar que $F(U) \cap W = \emptyset$. En efecto, supongamos que existe $\overrightarrow{\mathbf{y}}$ perteneciente a la intersección entre $F(U)$ y W entonces existe

$f_0(\overrightarrow{\mathbf{x_1}}) < f_0(\overrightarrow{\mathbf{x_c}})$, $f_i \leq 0$ para $i = 1, \ldots, p$ y $f_i = 0$ para $i = p+1, \ldots, p+q$ con $\overrightarrow{\mathbf{x_1}} \in U$.

Lo anterior es interpretado como la existencia de un punto en U que contradice la optimalidad de $\overrightarrow{\mathbf{x_c}}$, en consecuencia no existe $\overrightarrow{\mathbf{y}} \in F(U) \cap W$, que cumpla con lo anterior. Por lo tanto,

$$F(U) \cap W = \emptyset$$

Ahora demostraremos la segunda inplicancia de este lema, vale decir, sea $F(U) \cap W$ debemos demostrar que $\overrightarrow{\mathbf{x_c}}$ es una solución óptima sobre U.

Para realizar esta demostración procederemos por el principio de contradicción, es decir, supongamos que $\overrightarrow{\mathbf{x_c}}$ no es una solución óptima por lo que existirá algún $\overrightarrow{\mathbf{x_1}} \in U$ de tal modo que este punto sea una solución óptima, por la definición de óptimo se tiene que $f_0(\overrightarrow{\mathbf{x_1}}) \leq f_0(\overrightarrow{\mathbf{x}})\ \forall\ \overrightarrow{\mathbf{x}} \in U$, en particular para $\overrightarrow{\mathbf{x_c}}$, lo que nos indicaría que existe un punto en la intersección de $F(U)$ y de W, lo que contradice la hipótesis, por lo tanto $\overrightarrow{\mathbf{x_c}}$ es una solución óptima. ♦

Observaremos que el conjunto W que hemos definido anteriormente es convexo, vale decir que, para todo $\lambda \in [0,1]$ y todo $\overrightarrow{\mathbf{z}}$, $\overrightarrow{\mathbf{y}} \in W$ se tiene que $\lambda\overrightarrow{\mathbf{z}} + [1-\lambda]\overrightarrow{\mathbf{y}} \in W$. En efecto, consideremos $\overrightarrow{\mathbf{z}} = (\eta_0, \eta_1, ..., \eta_{p+q})$ e $\overrightarrow{\mathbf{y}} = (\tau_0, \tau_1, ..., \tau_{p+q})$ por lo que al multiplicarlos por λ y $[1-\lambda]$ respectivamente y luego sumar ambas expresiones se obtiene lo siguiente

$$(\tau_0 + \lambda[\eta_0 - \tau_0], \ldots, \tau_{p+q} + \lambda[\eta_{p+q} - \tau_{p+q}])$$

Ahora resta verificar que el nuevo vector de R^{p+q+1} pertenece al conjunto W. Efectivamente es así, dado que $\eta_0 < f_0(\overrightarrow{\mathbf{x_c}})\ y\ \tau_0 < f_0(\overrightarrow{\mathbf{x_c}})$, añadido al hecho de que λ y $1-\lambda$ son mayores que 0, podemos efectuar la combinación de las desigualdades mencionadas anteriormente, obteniéndose:

$$\tau_0 + \lambda[\eta_0 - \tau_0] < f_0(\overrightarrow{\mathbf{x_c}})$$

Por medio del mismo procedimiento se puede establecer que

$$\begin{array}{ll} \tau_i + \lambda[\eta_i - \tau_i] \leq 0 & \forall\, i = 1, \ldots, p \\ y & \\ \tau_i + \lambda[\eta_i - \tau_i] = 0 & \forall\, i = p+1, \ldots, p+q \end{array}$$

Por lo que se muestra que W es convexo.

Lo que el lema plantea es que si existe un conjunto que considera todos los valores reales en su primera componente tal que estos sean menores que la función objetivo evaluada en un punto $(\overrightarrow{\mathbf{x_c}}) \in U$, valores negativos (o cero) en $\eta_i\ \ i = 1, ..., p$, y $\eta_i = 0\ \ i = p+1, ..., p+q$ y además si se tiene el conjunto $F(U)$ -originado por la aplicación de la función F sobre U-. Entonces $\overrightarrow{\mathbf{x_c}}$ será necesariamente una solución óptima si la intersección de dichos conjuntos es vacía, hecho que es bastante natural si nos detenemos a pensar en el hecho que la primera componente de cada uno de los vectores de este conjunto será menor que el valor que alcanza f_0 evaluada en $\overrightarrow{\mathbf{x_c}}$.

Como comentario adicional a este lema podemos señalar que la demostración anteriormente realizada muestra una importante caracterización de las condiciones de optimalidad, la que cobrará una gran importancia en las demostraciones concernientes al teorema de Karush-Kuhn-Tuucker.

3.1. Separación de conjuntos convexos

Consideremos $\overrightarrow{\mathbf{l}} = (\lambda_1, ..., \lambda_k)$ y $\overrightarrow{\mathbf{y}} = (\eta_1, ..., \eta_k)$ en $\mathbb{R}^k$, donde

$$\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{y}} = \sum_{i=1}^{k} \lambda_i \eta_i$$

denota el producto interno usual en $\mathbb{R}^k$

Luego, tenemos que $H \subset \mathbb{R}^k$ se dice *hiperplano* si para algún $\overrightarrow{\mathbf{l}} \in \mathbb{R}^k$, con $\overrightarrow{\mathbf{l}} \neq 0$ y un número real γ $H = \left\{ \overrightarrow{\mathbf{y}} \in \mathbb{R}^k / \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{y}} = \gamma \right\}$

Se definen dos *semiespacios* cerrados de $H \subset \mathbb{R}^k$, denotados por H^+, H^- tal que

$$H^+ = \left\{ \overrightarrow{\mathbf{y}} \in \mathbb{R}^k / \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{y}} \geq \gamma \right\}$$

$$H^- = \left\{ \overrightarrow{\mathbf{y}} \in \mathbb{R}^k / \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{y}} \leq \gamma \right\}$$

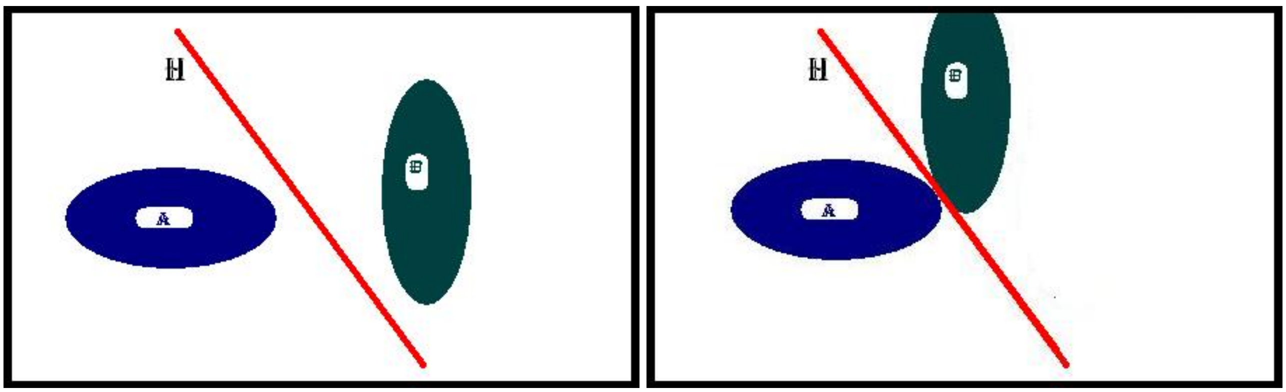

Figura 3.1: En estas imágenes se puede apreciar la diferencia entre un conjunto estrictamente separado (primera figura) y uno separado (segunda figura)

Entonces se dice que dos conjuntos ,$C_1 \subset H^+$ y $C_2 \subset H^-$, son estrictamente separados si $C_1, C_2 \cap \ H = \emptyset$ y son separados cuando $C_1, C_2 \cap \ H$ existe.

Es fácil notar que si C_1, C_2 son estrictamente separados, entonces también son separados. Sin embargo, tenemos que el recíproco es falso, pues si suponemos $\overrightarrow{\mathbf{x}} \in C_1, C_2 \cap H$ y que C_1 y C_2 son estrictamente separados, por lo que se tendría que $C_1, C_2 \cap \ H = \emptyset$, que es una abierta contradicción con la suposición inicial.

Lema 3.1.1 *Consideremos un conjunto convexo, cerrado y distinto de vacío de $\mathbb{R}^k$, al que llamaremos C, con $0 \notin C$. Entonces, C y $\{0\}$ son estrictamente separados.*

Demostración: Consideremos la función $d : \mathbb{R}^k \to \mathbb{R}$ tal que $d(\overrightarrow{\mathbf{x}}) = \|\overrightarrow{\mathbf{x}}\|^2 = \overrightarrow{\mathbf{x}} \cdot \overrightarrow{\mathbf{x}}$. Notar que

$$\lim_{\overrightarrow{\mathbf{x}} \to \infty} d(\overrightarrow{\mathbf{x}}) = \infty$$

Luego, al ser C un conjunto cerrado y no vacío y d continua, tendremos que d es limitada inferiormente y por ello podemos afirmar que existe $\overrightarrow{\mathbf{l}} \in C$ tal que

$$d(\overrightarrow{\mathbf{l}}) = \min_{\overrightarrow{\mathbf{x}} \in C} d(\overrightarrow{\mathbf{x}})$$

Recordemos que C es convexo y que $\overrightarrow{\mathbf{x}} \in C$. De esta forma, podemos fijar $\overrightarrow{\mathbf{x}}$ en C y así afirmar que $\alpha \overrightarrow{\mathbf{x}} + (1-\alpha)\overrightarrow{\mathbf{l}} \in C$ para todo α que se encuentre en el intervalo $[0, 1]$,

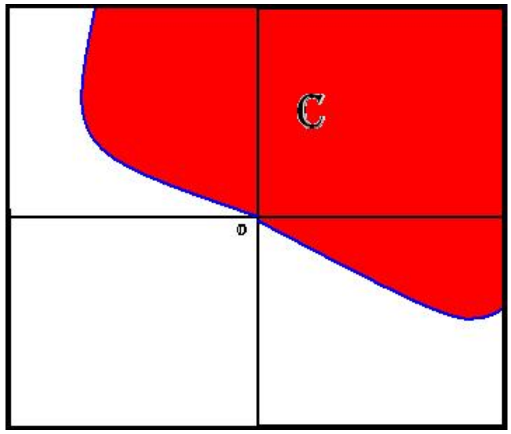

Figura 3.2: Separación de cero y C

producto de que C es convexo. Como el mínimo de d sobre C es obtenido en $\overrightarrow{\mathbf{l}}$, tenemos que $\|\alpha\overrightarrow{\mathbf{x}} + (\mathbf{1}-\alpha)\overrightarrow{\mathbf{l}}\|^{\mathbf{2}} \geq \|\overrightarrow{\mathbf{l}}\|^{\mathbf{2}}$ que es lo mismo que $2\alpha\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{y}} - 2\alpha\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}} + \alpha^2\overrightarrow{\mathbf{y}}\cdot\overrightarrow{\mathbf{y}} - 2\alpha^2\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{y}} + \alpha^2\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}} \geq 0$ o bien $2\alpha\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{l}} + \alpha^2\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{x}} - 2\alpha^2\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{l}} + \alpha^2\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}} \geq 2\alpha\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}}$. Ahora, al dividir por α, se obtiene que $2\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{l}} + \alpha\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{x}} - 2\alpha\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{l}} + \alpha\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}} \geq 2\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}}$ y si consideramos que $\alpha \to 0^+$ entonces $\overrightarrow{\mathbf{x}}\cdot\overrightarrow{\mathbf{l}} \geq \overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{l}} = \|\overrightarrow{\mathbf{l}}\|^2$ y además podremos definir: $\gamma = \frac{1}{2}\cdot\|\overrightarrow{\mathbf{l}}\|^2$ con lo que $\overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{x}} > \gamma > 0$ con lo que obtenemos la separación deseada. ♦

Lema 3.1.2 *Considerar a $C \subset \mathbb{R}^k$, convexo no vacío con $0 \notin C$, entonces C y $\{0\}$ son separados.*

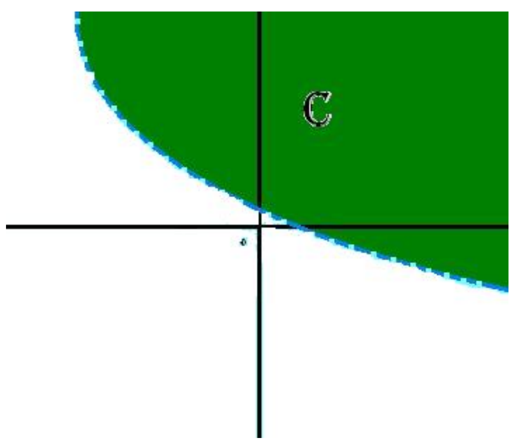

Figura 3.3: Separación de cero y C, siendo C no necesariamente cerrado

Demostración:

Para proceder a demostrar el lema utilizaremos el lema anterior. Con este fin, definamos

$$K(\overrightarrow{\mathbf{x}}) = \left\{\overrightarrow{\mathbf{l}} \in \mathbb{R}^k : \|\overrightarrow{\mathbf{l}}\| = 1 \ \text{ y } \ \overrightarrow{\mathbf{l}}\cdot\overrightarrow{\mathbf{x}} \geq 0\right\},$$

con $\overrightarrow{\mathbf{x}} \in C$.

Es sencillo notar que $K(\overrightarrow{\mathbf{x}}) \subset \overline{\mathbf{B}(0,1)} \subset \mathbb{R}^k$, dado que $\overline{\mathbf{B}(0,1)} = \left\{\overrightarrow{\mathbf{y}} \in \mathbb{R}^k \ : \ \|\overrightarrow{\mathbf{y}}\| \leq 1\right\}$ es decir $K(\overrightarrow{\mathbf{x}})$ se encuentra contenido en la frontera de $\overline{\mathbf{B}(0,1)}$.

En esta ocasión el real buscado para el hiperplano es 0, por lo que bastará que

$$\bigcap_{\overrightarrow{\mathbf{x}} \in C} K(\overrightarrow{\mathbf{x}}) \neq \emptyset, \text{ puesto que para todo } \overrightarrow{\mathbf{l}} \in \bigcap_{\overrightarrow{\mathbf{x}} \in C} K(\overrightarrow{\mathbf{x}}) \text{ se tendrá que } \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}} \geq 0 \ \forall \ \overrightarrow{\mathbf{x}} \in C.$$

Observemos que $K(\overrightarrow{\mathbf{x}})$ es cerrado ($\|\overrightarrow{\mathbf{l}}\| = 1$) dentro de la bola compacta.

Veamos que

$$\bigcap_{\overrightarrow{\mathbf{x}} \in C} K(\overrightarrow{\mathbf{x}}) \neq \emptyset$$

Consideremos $\overrightarrow{\mathbf{x_1}} \dots \overrightarrow{\mathbf{x_p}} \in C$. Entonces se define la siguiente envolvente convexa (menor conjunto convexo que contiene a un conjunto) $co\left\{\overrightarrow{\mathbf{x_1}} \dots \overrightarrow{\mathbf{x_p}}\right\} \subset \mathbb{R}^k$ es un convexo cerrado donde cero no pertenece al conjunto de la envolvente convexa. Por el lema anterior, tenemos que existe $\overrightarrow{\mathbf{l}} \in \mathbb{R}^k$, $\overrightarrow{\mathbf{l}} \neq 0$ tal que $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}} \geq 0 \ \forall \overrightarrow{\mathbf{x}} \in co\left\{\overrightarrow{\mathbf{x_1}} \dots \overrightarrow{\mathbf{x_p}}\right\}$ y con mayor razón para $\overrightarrow{\mathbf{x_1}} \dots \overrightarrow{\mathbf{x_p}}$, probando que

$$\overrightarrow{\mathbf{l}} \in \bigcap_{j=1}^{p} K(\overrightarrow{\mathbf{x_j}}).$$

♦

Observación 3.1.1 *Observemos que el γ que nos permite establecer los hiperplanos que hemos buscado para realizar la demostración anterior es el real cero, vale decir, $\gamma = 0$. De este modo los hiperplanos construidos en el lema 3.1.2 son explícitamente*

$$H^+ = \left\{\overrightarrow{\mathbf{x}} \in \mathbb{R}^k \ : \ \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}} \geq 0\right\} \ y$$

$$H^- = \left\{\overrightarrow{\mathbf{x}} \in \mathbb{R}^k \ : \ \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}} \leq 0\right\}$$

Lema 3.1.3 *Sea $C \subset \mathbb{R}^k, C \neq \emptyset$, convexo, con $0 \notin intC$, entonces C y $\{0\}$ son separados.*

Demostración: En esta ocasión podemos observar que no nos entregan mayor información del conjunto C, por lo que estamos en condiciones de afirmar que este tienen dos grandes posibilidades; la primera que sea cerrado y la segunda es que sea abierto. Asimismo, si es cerrado es factible que cero este o no en la frontera, por otro lado, si es abierto, nos encontramos con una situación similar.

1. De este modo si C es cerrado y cero esta en el complemento, esto coincide con las hipótesis del lema (3.1.1), por lo que de este modo se demostraría lo solicitado, no obstante aún nos falta por analizar otras opciones.

2. Por otro lado si cero no está en la frontera de C y por hipótesis se tiene que no está en el interior de C -considerando a este conjunto como abierto- es posible definir un conjunto al que denotaremos por C_r y es definido como $C_r = Fr(C) \cup C$. Es sencillo notar que C_r es cerrado y convexo. Siendo que cero no está en C_r se afirma que nuevamente tenemos las hipótesis del lema 3.1.1, con lo que bajo estas suposiciones se demuestra el lema.

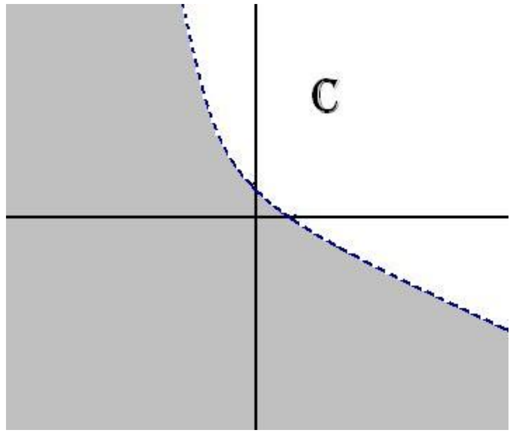

Figura 3.4: Separación de cero y en interior de C

3. Por otro lado si C es abierto y cero está en la frontera de C este lema que demostrado por el lema 3.1.2, dado que si C es abierto, $C = intC$ y $0 \notin C$, dado que $Fr(C) \nsubseteq C$.

4. Finalmente supongamos que $0 \in Fr(C)$, siendo se C abierto, esto quiere decir que $0 \notin C$, por lo que por el lema 3.1.1 se tiene la separación deseada.

5. Ahora supongamos que $0 \in FrC \subset C$, así establezcamos una sucesión de $\overrightarrow{\mathbf{y_n}} \notin C$ de tal modo que cuando $n \to \infty$ entonces, $\overrightarrow{\mathbf{y_n}} \to 0$, en consecuencia, por el lema 3.1.2 es posible afirmar que existirá un $\overrightarrow{\mathbf{l_n}} \in \mathbb{R}^k$ de tal modo que $\|\overrightarrow{\mathbf{l_n}}\| = 1$. Es claro que $\overrightarrow{\mathbf{l_n}} \in \overline{B(0,1)}$, por lo que una subsucesión de $\overrightarrow{\mathbf{l_n}}$, -digamos $\overrightarrow{\mathbf{l_{n_j}}}$- que será convergente, de este forma tenemos que para todo $\overrightarrow{\mathbf{x}} \in C$ se cumple que $\overrightarrow{\mathbf{l_{n_j}}} \cdot \overrightarrow{\mathbf{x}} \geq \overrightarrow{\mathbf{l_{n_j}}} \cdot \overrightarrow{\mathbf{y_{n_j}}}$ $\forall$ $\overrightarrow{\mathbf{x}} \in C$, así cuando $j \to \infty$, $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}} \geq 0$. Obteniendo al desigualdad deseada.

De (1),(2),(3),(4),(5) se consigue la demostración requerida. ♦

Observación 3.1.2 *En el punto (5), la última desigualdad se justifica con el hecho de que si una sucesión es convegente, entonces tendremos que sus subsucesiones también convergerán.*

3.2. Teorema de Separación de Hahn-Banach

Hemos notado que los lemas que tratamos anteriormente nos permitián decidir si un conjunto era o no separable con cero, no obstante, hasta el momento no poseemos una caracterización que nos otorge las herramientas necesarias para decidir acerca de la separabilidad de dos conjuntos cualesquiera. Para solucionar esta problemática es que surgen los teoremas de separación de Hanh-Banach.

Para tratar esta caracterización es imprescindible contar con un conjunto que realice la diferencia entre dos conjuntos dados. En otras palabras, el conjuntos al que hacemos referencia es definido de la siguiente manera:

Definición 3.2.1 *Consideremos dos subconjuntos de* $\mathbb{R}^k$*, digamos* C_1 *y* C_2*, entonces:*

$$C_1 - C_2 = \left\{ \overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}} / \overrightarrow{\mathbf{y_1}} \in C_1 \ y \ \overrightarrow{\mathbf{y_2}} \in C_2 \right\}$$

A partir de la definición anterior podemos realizar la observación de que los conjuntos C_1 y C_2 no tendrán elemetos en común, es decir serán disjuntos, sí y sólo sí $\overrightarrow{\mathbf{0}} \notin C_1 - C_2$, o más formalmente

$$C_1 \cap C_2 = \emptyset \ \Leftrightarrow \ \overrightarrow{\mathbf{0}} \notin C_1 - C_2$$

Hecho que tendrá una notable importancia dentro del teorema de separación de Hanh-Banach, puesto que intuitivamente sabemos que si dos conjuntos no tienen elementos en común, estos dejarán "espacios" estre ellos con lo que se hace mucho más probable una sepación entre ellos.

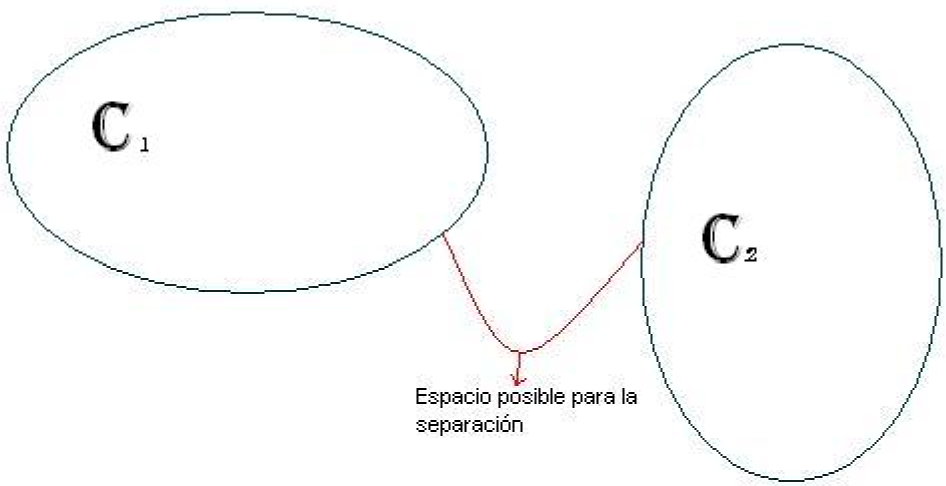

Figura 3.5: Ilustración del teorema de Hanh-Banach

Luego de este preámbulo enunciaremos y demostraremos el teorema a tratar.

Teorema 3.2.1 *Hahn-Banach Supongamos* C_1 *y* $C_2 \subset \mathbb{R}^k$ *convexos no vacíos entonces cada una de las siguientes condiciones nos permite decir que* C_1 *y* C_2 *son separables:*

$$\overrightarrow{\mathbf{0}} \notin C_1 - C_2 \tag{3.1}$$

$$\overrightarrow{\mathbf{0}} \notin int(C_1 - C_2) \tag{3.2}$$

$$int\, C_2 \neq \phi \ y \ \overrightarrow{\mathbf{0}} \notin \ C_1 - int\, C_2 \tag{3.3}$$

Demostración: En primer término defineremos el conjunto $C = C_1 - C_2$. Es sencillo notar que C es convexo, pues como C_1 es convexo se tiene que para todo $\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}} \in C_1$ se tiene que $\lambda \overrightarrow{\mathbf{x_1}} + (1-\lambda)\overrightarrow{\mathbf{y_1}} \in C_1(1)$. Asimismo, para todo $\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}} \in C_2$ se tiene que $\lambda\overrightarrow{\mathbf{x_2}} + (1-\lambda)\overrightarrow{\mathbf{y_2}} \in C_2(2)$.

Ahora al realizar la sustracción entre (1) y (2) y al proceder de manera conveniente obtenemos que

$$\lambda\overrightarrow{\mathbf{x_1}} + (1-\lambda)\overrightarrow{\mathbf{y_1}} - \lambda\overrightarrow{\mathbf{x_2}} - (1-\lambda)\overrightarrow{\mathbf{y_2}} = \lambda(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}) + (1-\lambda)(\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}}) \in C \ \ \forall \ \lambda \in [0,1].$$

En consecuencia C es convexo.

1. Ahora utilizando el hecho recién demostrado estamos en condiciones de decir fehacientemente que $\overrightarrow{\mathbf{0}}$ y C son separables, esto al ser reintrerpretado se dice que el conjunto es no vacío dado que C_1, C_2 no son vacíos y además el conjunto C es convexo, entonces en virtud del 3.1.2 se tiene que $\overrightarrow{\mathbf{0}}$ y C son separados. También el lema nos asegura que el γ que nos permite construir los hiperplanos deseados es cero, así de esta forma tenemos que para todo $\overrightarrow{\mathbf{x}} \in C$ existirán $\overrightarrow{\mathbf{l}} \in \mathbb{R}^k$ de tal manera que $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}} \geq 0$ o bien $\displaystyle\sum_{j=1}^{k} \overrightarrow{\mathbf{l}}_j \cdot \overrightarrow{\mathbf{x_j}} \geq 0$. No obstante se sabe que $\overrightarrow{\mathbf{x}} \in C = C_1 - C_2$, entonces se está en condiciones de afirmar que con $\overrightarrow{\mathbf{x}} = \overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}$ con $\overrightarrow{\mathbf{x_1}} \in C_1$ y con $\overrightarrow{\mathbf{x_2}} \in C_2$, tenemos que:

$$\sum_{j=1}^{k} \overrightarrow{\mathbf{l_j}} \cdot \overrightarrow{\mathbf{x_j}} = \sum_{j=1}^{k} \overrightarrow{\mathbf{l_j}} \cdot \overrightarrow{\mathbf{x_{1_j}}} - \sum_{j=1}^{k} \overrightarrow{\mathbf{l_j}} \cdot \overrightarrow{\mathbf{x_{2_j}}} \geq 0$$

que implica

$\displaystyle\sum_{j=1}^{k} l_j \cdot x_{1_j} \geq \sum_{j=1}^{k} \overrightarrow{\mathbf{l_j}} \cdot \overrightarrow{\mathbf{x_{2_j}}}$ o dicho de otro modo $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_1}} \geq \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_2}}$

En consecuencia, planteamos que si nosotros tenemos que $\gamma_1 = \sup\left\{\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_2}} : \overrightarrow{\mathbf{x_2}} \in C_2\right\}$, y asimismo $\gamma_2 = \text{ínf}\left\{\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_1}} / \overrightarrow{\mathbf{x_1}} \in C_1\right\}$, es decir, $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_2}} \geq \gamma_1$ y $\gamma_2 \geq \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_1}}$ por lo tanto tenemos que $\gamma_1 \geq \gamma_2$ y además, por teorema del punto medio, tenemos que

$$\gamma_1 \leq \frac{\gamma_1 + \gamma_2}{2} \leq \gamma_2$$

A través de este razonamiento obtenemos que $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_2}} \leq \gamma_1 \leq \frac{\gamma_1+\gamma_2}{2} \leq \gamma_2 \leq \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_1}}$

y mediante ello podemos concluir que $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_2}} \leq \dfrac{\gamma_1 + \gamma_2}{2} \leq \overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x_1}}$, con lo que se extrae la separación requerida.

2. En este caso estamos interesados en demostrar, en primer término, que el interior de $C_1 - C_2$, que llamaremos C_i, es convexo. Para este fin procedamos por el principio de contradicción, es decir supongamos que C_i no es convexo, esto implica que $\exists \overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}} \in int(C_1 - C_2) \subset C_1 - C_2$, en consecuencia, tendremos que $\lambda\overrightarrow{\mathbf{x}} + (1-\lambda)\overrightarrow{\mathbf{y}} \notin C$, y por definición de interior nos encontramos con que existen $\mathfrak{B}(x,r)$ y $\mathfrak{B}(y,r) \forall r \in \mathbb{R}$ tal que

$\mathfrak{B}(x,r) \subset C$ y $\mathfrak{B}(y,r) \subset C$ pero $\lambda\overrightarrow{\mathbf{x}} + (1-\lambda)\overrightarrow{\mathbf{y}} \notin C$, lo que resulta contradictorio con el hecho de que C es convexo (demostrado en i)), por lo que C_i es convexo.

En virtud de lo anterior y del lema 3 podemos proceder de manera similar a la demostración i), vale decir, el lema citado nos enseña que $\gamma = 0$ y que existirán $\overrightarrow{\mathbf{l}} \in \mathbb{R}^k$ tal que $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{x}}_{\in C_i} \geq 0$, y al seguir los pasos de i) obtenemos la separación deseada.

3. Anteriormente hemos visto que el interior de un conjunto convexo es convexo, entonces el conjunto $C_0 = C_1 - intC_2$ es convexo (por lo visto en i) sobre la convexidad de $C = C_1 - C_2$). Además se nos señala en las hipótesis que $\overrightarrow{\mathbf{0}} \notin C_0$, entonces podemos aplicar el lema 3 y seguir i), por lo tanto C_1 y C_2 son separados. ♦

Capítulo 4

Reglas de multiplicadores

En el transcurso de este capítulo trataremos las diferentes reglas de multiplicadores (reglas de multiplicadores convexa, regla de John, Carathéodory, Carathéodory-John, etc.) con las que el teorema de Karush-Kuhn-Tucker será tratado en diversas situaciones desmenuzándolo en varios lemas.

4.1. Reglas de Multiplicadores sin restrición de igualdad

Suponga que U es un conjunto convexo abierto sobre $\mathbb{R}^n$, y $f_0, ..., f_p$ son $p+1$ funciones convexas sobre U.

Buscamos un punto $\overrightarrow{\mathbf{x_c}}$ que sea solución del siguiente problema

$$T \begin{cases} \min f_0(\overrightarrow{\mathbf{x}}) \\ \text{sujeto a} \quad f_i(\overrightarrow{\mathbf{x}}) \leq 0 \quad i = 1, ..., p \end{cases}$$

donde f_0 y $f_i : U \subseteq \mathbb{R}^n \to \mathbb{R}$ son funciones dadas.

4.1.1. Reglas de Multiplicadores convexa

Tenemos que si $\overrightarrow{\mathbf{x_c}} \in U$ resuelve T, entonces existe un vector $l = (\lambda_0, ..., \lambda_p) \neq 0 \in \mathbb{R}^{p+1}$ que satisface las siguientes condiciones:

(i) Si $f = \displaystyle\sum_{i=0}^{p} \lambda_i f_i$ con $\overrightarrow{\mathbf{x_c}} \in U$ entonces

$\lambda_0 f_0(\overrightarrow{\mathbf{x_c}}) = f(\overrightarrow{\mathbf{x_c}})$ y también $f(\overrightarrow{\mathbf{x_c}}) \leq f(\overrightarrow{\mathbf{x}})$,

(ii) $\lambda_i \geq 0,$ para todo, $i = 0, 1, ..., p$ y además que se cumpla

(iii) $\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) = 0$, donde $i = 1, ..., p$

Demostración: Para realizar esta demostración recurriremos a los teoremas de Hahn-Banach. Para realizar la demostración en cuentión por medio de este teorema necesitamos asegurar la existencia de dos conjuntos, contenidos en $\mathbb{R}^k$, que sean convexos. En esta línea definamos el conjunto

$$K(\overrightarrow{\mathbf{x}}) = \left\{ \overrightarrow{\mathbf{y}} = (\eta_0, \eta_1, ..., \eta_p) \in \mathbb{R}^{p+1} / f_i(\overrightarrow{\mathbf{x}}) \leq \eta_i \;\; i = 1, \ldots, p \right\}, \text{ con } \overrightarrow{\mathbf{x}} \in U.$$

Ahora consideremos

$$K = \bigcup_{\overrightarrow{\mathbf{x}} \in U} K(\overrightarrow{\mathbf{x}}).$$

Nuestra tarea en este momento será la de demostrar que K es convexo. En efecto, si se tienen $\overrightarrow{\mathbf{y}}_1, \overrightarrow{\mathbf{y}}_2$ pertenecientes a K, se pretende comprobar que

$$\lambda \overrightarrow{\mathbf{y}}_1 + (1-\lambda)\overrightarrow{\mathbf{y}}_2 \text{ pertenece a } K \text{ para todo lambda entre cero y uno} \tag{4.1}$$

De este modo se ve que, siendo $\overrightarrow{\mathbf{y}}_1 = (\alpha_0, \ldots, \alpha_p)$ y $\overrightarrow{\mathbf{y}}_2 = (\beta_0, \ldots, \beta_p)$,tenemos

$$\lambda(\alpha_0, \ldots, \alpha_p) + (1-\lambda)(\beta_0, \ldots, \beta_p) =$$

$$= (\beta_0 + \lambda(\alpha_0 - \beta_0), \ldots, \beta_p + \lambda(\alpha_p - \beta_p))$$

Así resta verificar la condición que nos permitirá observar que este elemento pertenece a K, es decir, necesitamos probar que $f_i(\overrightarrow{\mathbf{x}}) \leq \beta_i + \lambda(\alpha_i - \beta_i)$, con $\overrightarrow{\mathbf{x}}$ en U e i entre uno y p. Además, se sabe que

$$f_i(\overrightarrow{\mathbf{x}}) \leq \alpha_i \quad y \quad f_i(\overrightarrow{\mathbf{x}}) \leq \beta_i$$

Por lo que multiplicando la primera desigualdad por λ y la segunda desigualdad por $1-\lambda$ y sumando ambos resultados se obtiene

$$f_i(\overrightarrow{\mathbf{x}}) \leq \alpha_i + \lambda(\beta_i - \alpha_i)$$

con lo que se prueba 4.1.

i) Para realizar esta demostración definamos $F = (f_0, \ldots, f_p) : U \longrightarrow \mathbb{R}^{p+1}$. Por definición de F, es evidente que $F \subset K$ pues K considera todos los puntos de $\mathbb{R}^{p+1}$ tal que las componentes de sus elementos sean menores o iguales a las componentes de los elementos de $F(U)$. Además, estableceremos el conjunto

$$W = \{\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_p) \in \mathbb{R}^{p+1} / \eta_0 < f(\overrightarrow{\mathbf{x_c}}), \;\; \eta_i \leq 0 \text{ con } i = 1, \; \ldots, p\}.$$

Por el lema 3.0.1 se tiene que

$$F(U) \cap W = \phi \;\text{ y }\; F(\overrightarrow{\mathbf{x_c}}) \in F(U) \cap \overline{W}. \tag{4.2}$$

Por el teorema 3.1, se tiene que los conjuntos $F(U) \subseteq K$ y W, además por 4.2 tenemos que el hiperplano de esta separación pasa por $F(\overrightarrow{\mathbf{x_c}})$ y por el teorema antes mencionado tenemos que existe $l = (\lambda_0, \ldots, \lambda_p) \in \mathbb{R}^{p+1}$ de tal forma que

$$l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{x_c}})] \geq 0 \;\text{ para todo } \overrightarrow{\mathbf{y}} \in K \tag{4.3}$$

$$l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{x_c}})] \leq 0 \text{ para todo } \overrightarrow{\mathbf{y}} \in W \tag{4.4}$$

Tenemos que $F(U) = \left\{F(\overrightarrow{\mathbf{x}}) : \overrightarrow{\mathbf{x}} \in U\right\}$ contenido en K y considerando 4.3, junto con el hecho de que el conjunto K es cerrado, es decir, $\overline{K} = K$, podemos afirmar que la condición citada se cumple en particular para $F(\overrightarrow{\mathbf{x}}) \in F(U) \subseteq K$, y así se obtiene que $\sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}}) \leq \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}})$, para todo $\overrightarrow{\mathbf{x}}$ que pertenece a U. Para finalizar esta demostración debemos comprobar la condición tres, por lo que esta demostración la concluiremos más adelante.

ii) En primer lugar asumamos que $\lambda_0 < 0$, además por la desigualdad b) se sabe que $\sum_{i=0}^{p} \lambda_i \eta_i \leq \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}})$. Así, suponiendo que el resto de los λ_i con $i = 1, \ldots, p$, son mayores que cero, tendremos que $\lambda_0 \eta_0 \leq \lambda_0 f_i(\overrightarrow{\mathbf{x_c}})$ y en consecuencia $\eta_0 \geq f_i(\overrightarrow{\mathbf{x_c}})$ lo que es una contradicción con respecto a la definición de los elementos de W.

Ahora, si suponemos que $\lambda_j < 0$, para algún valor de j entre 1 y p, tendremos que tomar en consideración que, por la desigualdad 4.4, ocurre que para todo $\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_p) \in W$ se da el hecho de que

$$\sum_{i=0}^{p} \lambda_i \eta_i \leq \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}})$$

Sin embargo, si hacemos que $\eta_j \to -\infty$, tendremos que $\lambda_j \eta_j \to +\infty$, por lo que

$$\lambda_j \eta_j + \sum_{i=0}^{j-1} \lambda_i \eta_i + \sum_{i=j+1}^{p} \lambda_i \eta_i \geq \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}})$$

Que contradice lo escrito previamente en relación con la desigualdad para los términos $\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_p) \in W$, puesto que los términos no j-ésimos serán menores que $\lambda_0 f_0(\overrightarrow{\mathbf{x_c}})$, por lo que no se puede acotar el producto $l \cdot \overrightarrow{\mathbf{y}}$ y a través de ello se obtiene que $\lambda_i \geq 0, i = 1, \ldots, p$

iii) Para realizar la demostración de este punto consideramos el vector

$$\overrightarrow{\mathbf{y}} = (f_0(\overrightarrow{\mathbf{x_c}}), \ldots, \frac{1}{2} f_j(\overrightarrow{\mathbf{x_c}}), \ldots, f_p(\overrightarrow{\mathbf{x_c}})) \in \overline{W}$$

Considerando 4.4, y particularmente, tendremos que $l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{x_c}})] \leq 0$ con lo que obtenemos $\sum_{i=0}^{p} \lambda_i \eta_i - \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}}) \leq 0$. Al realizar las operaciones correspondientes, obtendremos que

$$-\frac{1}{2} \lambda_j f_j(\overrightarrow{\mathbf{x_c}}) \leq 0$$

y al multiplicar por -2 la desigualdad obtenida, tendremos que

$$\lambda_j f_j(\overrightarrow{\mathbf{x_c}}) \geq 0 \tag{4.5}$$

Además, por las restricciones del problema 4.1, tenemos que para los valores comprendidos entre 1 y p

$$f_i(\overrightarrow{\mathbf{x}}) \leq 0$$

Ahora, particularmente tendremos $f_i(\overrightarrow{\mathbf{x_c}}) \leq 0$ y si tenemos $\lambda_i \geq 0$, entonces $\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) \leq 0$ y de forma puntual

$$\lambda_j f_j(\overrightarrow{\mathbf{x_c}}) \leq 0 \tag{4.6}$$

De esta forma, y producto de 4.5 y 4.6 ,$\lambda_j f_j(\overrightarrow{\mathbf{x_c}})$ no tiene otra opción que ser igual a cero.

$$\therefore\ \lambda_j f_j(\overrightarrow{\mathbf{x_c}}) = 0, \text{ para } j = 1, \ldots, p$$

♦

Corolario 4.1.1 *(Karush-Kunh-Tucker) Si f_i, $i = 1, \ldots, p$ no son idénticamente nulas sobre un conjunto factible*

$$S = \left\{ \overrightarrow{\mathbf{x}} \in U, f_i(\overrightarrow{\mathbf{x}}) \leq 0,\ i = 1, \ldots, p \right\}$$

entonces las conclusiones de regla de multiplicadores convexa se satisface con $\lambda_0 > 0$. Recíprocamente, si las reglas de multiplicadores convexa se satisface para $\lambda_0 > 0$, entonces $\overrightarrow{\mathbf{x_c}} \in U$ resuelve el problema T que se encuentra en la página 21.

Demostración: En primer lugar demostraremos que se cumplen las reglas de multiplicadores convexa si se tienen las hipótesis de que f_i no son identicamente nulas sobre el conjunto factible S. De esta forma nuestra tarea se enfoca en demostrar que dadas estas condiciones existe un vector al que llamaremos $\overrightarrow{\mathbf{l}} = (\lambda_0, \ldots, \lambda_p)$ y supongamos que $\lambda_0 = 0$, de este modo se obtiene la desigualdad

$$f(\overrightarrow{\mathbf{x}}) = \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}}) \geq 0 \tag{4.7}$$

Además se tiene que λ_i, con $i =, \ldots, p$, es mayor o igual que cero y que $f_i(\overrightarrow{\mathbf{x}})$, con $i = 1, \ldots .p$ y $\overrightarrow{\mathbf{x}} \in U$, es menor o igual que cero, con lo que obtenemos que

$$f(\overrightarrow{\mathbf{x}}) = \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}}) \leq 0 \tag{4.8}$$

A partir de 4.7 y 4.8 concluimos que

$$f(\overrightarrow{\mathbf{x}}) = \sum_{i=0}^{p} \lambda_i f_i = 0$$

lo que implica que cada $\lambda_i f_i(\overrightarrow{\mathbf{x}})$ es igual a cero, no obstante se posee la información de que $f_i(\overrightarrow{\mathbf{x}})$, para $i = 1, \ldots, p$, no son necesariamente nulas, con lo que λ_i es igual a cero, para todo i considerado en el intervalo entre uno y p, lo que es una contradicción con el hecho de que $\overrightarrow{\mathbf{l}}$ es distinto de cero.

Ahora procederemos a demostrar que si las reglas de multiplicadores convexa entonces el punto $\overrightarrow{\mathbf{x_c}}$ resuelve el problema T.

Bajo estas presuposiciones se sabe que se cumplen las siguientes desigualdades

$$\lambda_0 f_0(\overrightarrow{\mathbf{x_c}}) = f(\overrightarrow{\mathbf{x_c}}) \leq f(\overrightarrow{\mathbf{x}}) = \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}}) \text{ para toda } \overrightarrow{\mathbf{x}} \in U \tag{4.9}$$

Además se conoce que $\lambda_i \geq 0$ y que $f_i(\overrightarrow{\mathbf{x}}) \leq 0$, para toda $i = 1, \ldots, p$, por lo que

$$\sum_{i=1}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}}) \leq 0 \tag{4.10}$$

y al sumarle $\lambda_0 f_0(\overrightarrow{\mathbf{x}})$ a 4.10 por se consigue la si

$$\lambda_0 f_0(\overrightarrow{\mathbf{x}}) + \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}}) \leq \lambda_0 f_0(\overrightarrow{\mathbf{x}}) \tag{4.11}$$

Aplicando la propiedad de transitividad sobre 4.9 y 4.11 se obtiene

$$f_0(\overrightarrow{\mathbf{x_c}}) \leq f(\overrightarrow{\mathbf{x}}) \text{ para toda } x \in U$$

por lo que $\overrightarrow{\mathbf{x_c}}$ satisface el problema T. ♦

4.1.2. Reglas de Multiplicadores de John

En el conjunto de reglas de los multiplicadores anteriormente estudiada se nos entrega información acerca de las condiciones que debe cumplir el vector definido en $\mathbb{R}^{p+1}$, al que hemos llamado $\overrightarrow{\mathbf{l}} = (\lambda_0, \ldots, \lambda_p)$, considerando que $\overrightarrow{\mathbf{x_c}}$ resuelve el problema T, y viceversa, si se cumplen las condiciones mencionadas para $\overrightarrow{\mathbf{l}}$ entonces el punto en cuentión es solución del problema T. Asimismo,las condiciones de las reglas de multiplicadores convexa trabajan directamente con las funciones involucradas, vale decir, determinan las condiciones que cumplen las componentes de $\overrightarrow{\mathbf{l}}$ operando sólo con la función objetivo y las funciones de restricción.

En esta sección estudiaremos un tipo de regla de multiplicadores conocida como *multiplicadores de John*, las que esta vez también consideran un vector $\overrightarrow{\mathbf{l}} = (\lambda_0, \ldots, \lambda_p)$ en $\mathbb{R}^{p+1}$ con el que definiremos una función que realiza el producto interno entre $F(U)$ y $\overrightarrow{\mathbf{l}}$ pero que esta vez la información será entregada sobre la diferencial según Fréchet. La condición de optimalidad que esta regla de multiplicadores plantea sobre la función que realiza el producto interno es similar a la caracterización que es utilizada en las funciones definida de los reales a los reales con la derivada usual, vale decir, si x_c soluciona el problema T entonces la derivada de la función ya mencionada es igual a cero. Sumado a lo anterior se deben considarar las condiciones que se establece sobre las componentes de $\overrightarrow{\mathbf{l}}$.

Las reglas de multiplicadores a las que hacemos referencia es la siguiente:

Supongamos que tenemos un conjuntos abierto U subconjunto de $\mathbb{R}^n$, además considere $f_0, f_1, \ldots, f_p$ como $p+1$ funciones reales sobre U, siendo cada una *deferenciable según Fréchet* en el punto $\overrightarrow{\mathbf{x_c}} \in U$. Suponiendo que $\overrightarrow{\mathbf{x_c}} \in U$ resuelve el problema T, entonces existe un vector en $\mathbb{R}^{p+1}$, al que llamaremos $\overrightarrow{\mathbf{l}}$, de tal forma que $\overrightarrow{\mathbf{l}} = (\lambda_0, \ldots, \lambda_p) \neq 0$ que satisface las siguientes condiciones:

(i) Si $f = \sum_{i=0}^{p} \lambda_i f_i$ con $\overrightarrow{\mathbf{x_c}} \in U$ entonces $f'(\overrightarrow{\mathbf{x_c}}) = 0$

(ii) $\lambda_i \geq 0,$ para todo, $i = 0, 1, ..., p$ y además se satiface

(iii) $\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) = 0$, donde $i = 0, 1, ..., p$

Antes de proseguir con la demostración de estas reglas de multiplicadores, mostraremos un sencillo ejemplo de la búsqueda de un mínimo en la que se utiliza una traslación de la función, con el fin de simplificar el problema sin perder generalidad.

Ejemplo 4.1.1 *Sea $f : \mathbb{R} \to \mathbb{R}$ de tal forma que $x \mapsto f(x) = x^2 + x$, como es posible observar en la gráfica que sigue:*

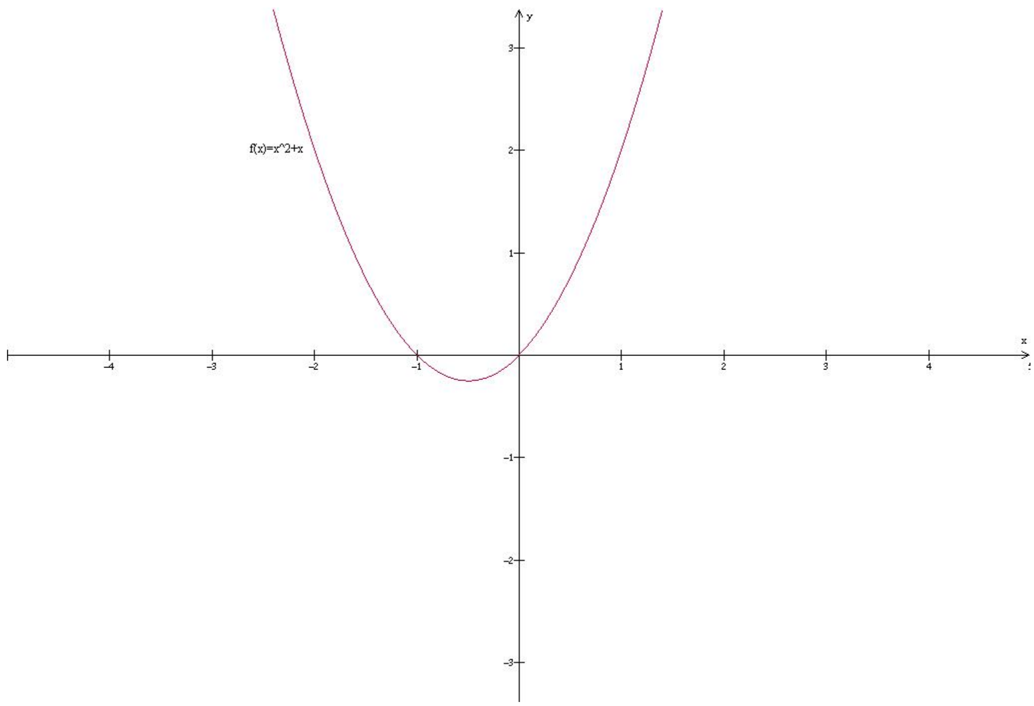

Figura 4.1: Gráfico función a trasladar

El punto en el que $f(x)$ alcanza su mínimo es $x = -0{,}5$ y $f(-0{,}5) = -0{,}25$. No obstante, nosotros trataremos de establecer una función con la que trasladaremos la función f al origen del eje de coordenadas.

De esta forma, tenemos la función

$h : \mathbb{R}^2 \to \mathbb{R}^2$

$(x, y) \mapsto h((x, y)) = (x - 0{,}5, y - 0{,}25)$

Así, reemplazando las coordenadas coorespondientes, la función f se reescribe como sigue:

$$y - \frac{1}{4} = (x - \frac{1}{2})^2 + (x - \frac{1}{2})$$

$$y = x^2$$

De manera más formal, se consigue obtener la nueva función $g : \mathbb{R} \to \mathbb{R}$ donde tendremos que $x \mapsto g(x) = x^2$, que es la traslación antes mencionada.

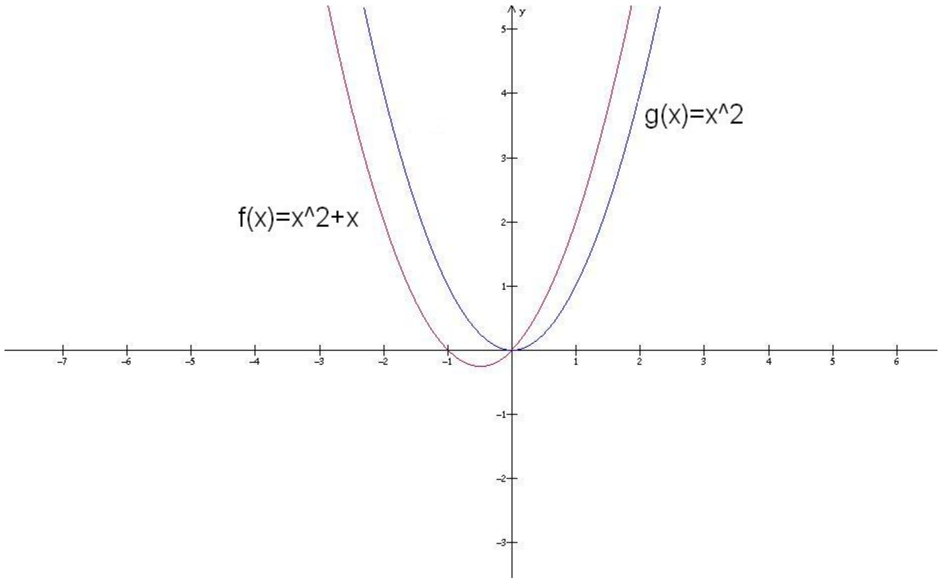

Figura 4.2: Función original y función trasladada

Por lo que es sencillo notar que el mínimo es alcanzado en el punto (0,0) y la nueva función cumple las mismas propiedades de la anterior ($f(x)$), por lo que no se pierde generalidad al considerar el origen para estudiar la optimalidad de una función.

Demostración: Como vimos en el ejemplo anterior, podemos realizar una traslación de la función, dejando el punto mínimo de esta en el origen. De esta manera, sin perder generalidad, podemos considerar $\overrightarrow{\mathbf{x_c}} = \overrightarrow{\mathbf{0}}$ y además $f_0(\overrightarrow{\mathbf{x_c}}) = 0$. Si le añadimos a lo precedente el hecho de que las distintas funciones f_i, con $i = 1, \ldots, p$ son diferenciables según Fréchet en el punto $\overrightarrow{\mathbf{x_c}}$, nosotros podemos aseverar que existe una vecindad de $\overrightarrow{\mathbf{x_c}}$, la cual, por definición, se puede deducir que es convexa. Sin temor a perder generalidad, consideremos a U un conjunto convexo y sea $\overrightarrow{\mathbf{x_c}} \in U$. Además, consideremos $F = (f_0, \ldots, f_p) : U \to \mathbb{R}^{p+1}$ y donde también se tiene

$$W = \left\{ \overrightarrow{\mathbf{y}} = (\eta_0, \eta_1, \ldots, \eta_p) \in \mathbb{R}^{p+1} : \eta_0 < 0,\ \eta_i \leq 0,\ \text{para } i = 1, \ldots, p \right\}$$

Debido a la condición de optimalidad del punto $\overrightarrow{\mathbf{0}} \in U$ garantiza que

$$F(U) \cap W = \emptyset \quad \text{y} \quad F(\overrightarrow{\mathbf{0}}) \in F(U) \cap \bar{W} \text{ condición de optimalidad} \tag{4.12}$$

Ahora, denotamos como L la derivada de F en el punto $\overrightarrow{\mathbf{0}} \in U$, además de definirse un funcional afín A de la siguiente forma:

$$A(\overrightarrow{\mathbf{x}}) = F(\overrightarrow{\mathbf{0}}) + L\overrightarrow{\mathbf{x}}$$

A continuación procederemos a mostrar que $A(U)$ y W son separados. Para este fin supongamos que no lo son, entonces por la condición 3.3 del teorema de separación de Hanh-Banach se tiene que cero, de $\mathbb{R}^{p+1}$, pertenece a $A(U) - intW$,o sea, existirá por lo menos un elemento en común entre ambos conjuntos, así se tiene que existe $\overline{x}$ en U de tal manera que $A(\overline{x})$ pertenece a $intW$. Además se sabe que F es Fréchet diferenciable en el punto cero y considerando un α tendiendo a cero real se puede escribir:

$$F(\overrightarrow{\mathbf{0}} + \alpha\overline{x}) = F(\overrightarrow{\mathbf{0}}) + L(\alpha\overline{x}) + r(\alpha\overline{x})\|\alpha\overline{x}\| \text{ con } r(\alpha\overline{x})\|\alpha\overline{x}\| \to 0 \text{ cuando } \alpha \to \infty \quad (4.13)$$

De esta forma, al considerar el funcional afin, se tiene que la expresion 4.13 podemos reescribirla de la siguiente manera:

$$F(\alpha\overline{x}) = (1-\alpha)F(\overrightarrow{\mathbf{0}}) - \alpha[A(\overline{x}) + r(\alpha\overline{x})\|\overline{x}\|] \quad (4.14)$$

De 4.14 podemos decir que $(1-\alpha)f(\overrightarrow{\mathbf{0}}) \leq 0$. Además $A(\overline{x}) + r(\alpha\overline{x}\|\overline{x}\|)$ pertenece al interior de W, puesto que si no fuera de esta forma se tendría que, al definir los vectores originados por $A(\overline{x}) + r(\alpha\overline{x})\|\overline{x}\|$ como $(a_0 + r_0\|\overline{x}\|, \ldots, a_p + r_p\|\overline{x}\|)$, existiría una componente del vector anteriormente señaldo de tal manera que $a_j + r_j\|\overline{x}\| > 0$. Considerando que $r(\alpha\overline{x})\|\alpha\overline{x}\| \to 0$ cuando $\alpha \to \infty$ y que W es abierto, en consecuencia $W = intW$, $a_j > 0$, lo que contradice el supuesto que $A(\overline{x})$ pertenece al interior de W. Por otro lado, es natural que si fijamos un α conveniente en el intervalo (0,1), $\alpha[A(\overline{x}) + r(\alpha\overline{x})\|\overline{x}\|]$ pertenecerá al interior de W.

Con lo anterior estamos en condiciones de afirmar que $F(\alpha\overline{x})$ está en el interior de W-con mayor razón en W-, y se sabe que $F(\alpha\overline{x})$ también está en $F(U)$, de esta forma la intersección entre $F(U)$ y W es distinta de vacío, lo que contradice la condición 4.12, así que $A(U)$ y W son necesariamente separables.

Además se tiene que $F(\overrightarrow{\mathbf{0}}) = A(\overrightarrow{\mathbf{0}})$ perteneciente a $A(U) \cap \overline{W}$, con lo que se puede señalar que existe un hiperplano que pasa por $F(\overrightarrow{\mathbf{0}})$. Asimismo, existe un vector $\overrightarrow{\mathbf{l}} = (\lambda_0, \ldots, \lambda_p)$ en $\mathbb{R}^{p+1}$ que satisface:

$$\overrightarrow{\mathbf{l}} \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] \geq 0 \text{ para todo } \overrightarrow{\mathbf{y}} \in A(U) \quad (4.15)$$

$$\overrightarrow{\mathbf{l}} \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] \leq 0 \text{ para todo } \overrightarrow{\mathbf{y}} \in W. \quad (4.16)$$

De 4.15 tenemos que para todo $\overrightarrow{\mathbf{x}}$ en U esta desigualdad se reescribe de la siguiente forma

$$\overrightarrow{\mathbf{l}} \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] = \overrightarrow{\mathbf{l}} \cdot L\overrightarrow{\mathbf{x}} \geq 0 \text{ para todo } \overrightarrow{\mathbf{y}} \in A(U) \quad (4.17)$$

Por otro lado se tiene que $L\overrightarrow{\mathbf{x}}$ es un operador lineal y además se esta considerando una vecindad de cero por lo que existirá una bola $B(0,r)$ contenida en U, para todo $r > 0$; por lo que por definición se está en condiciones de plantear la siguiente expresión:

$$L(\overrightarrow{\mathbf{x}} - 0)$$

pero se sabe que la norma con la que se define la bola cumple con la propiedad simétrica por lo que concluimos que $\|\overrightarrow{\mathbf{x}} - 0\| = \|0 - \overrightarrow{\mathbf{x}}\|$, con lo que la aplicación lineal L se puede evaluar de la forma $L0 - \overrightarrow{\mathbf{x}}$, así, por propiedad de linealidad tenemos que

$$L(\overrightarrow{\mathbf{x}}) = L(0) + L(-\overrightarrow{\mathbf{x}}) = 0 - L(\overrightarrow{\mathbf{x}}) = -L(\overrightarrow{\mathbf{x}})$$

de lo que unido a 4.17 se concluye que $\overrightarrow{\mathbf{l}} \cdot L\overrightarrow{\mathbf{x}} = 0$, lo que implica

$$\sum_{i=0}^{p} \lambda_i f_i'(0)\overrightarrow{\mathbf{x}} = 0 \text{ para todo } \overrightarrow{\mathbf{x}} \in \mathbb{R}^n$$

o dicho de otro modo

$$f'(0) = 0$$

con lo que se demuestra la condición (i).

ii) Supongamos $\lambda_j < 0$ para algún valor de j entre 1 y p. De esta forma, tendremos que $\lambda_j \cdot \eta_j \geq 0$. Ahora, consideremos $\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_p) \in W$ de tal forma que

$$\sum_{i=0}^{p} \lambda_i \eta_i \leq \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}})$$

Que es la desigualdad producto de 4.4. Pero, si nosotros consideramos η_j tal que $\eta_j \rightarrow -\infty \Rightarrow \lambda_j \eta_j \rightarrow +\infty$, nos encontraremos con que

$$\lambda_j \eta_j + \sum_{i=0}^{j-1} \lambda_i \eta_i + \sum_{i=j+1}^{p} \lambda_i \eta_i \geq \sum_{i=0}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x_c}}),$$

que se opone al hecho de que $\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_p)$ se encuentre en el conjunto W, dado que en este, por definición, $\overrightarrow{\mathbf{l}}[\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{x_c}})] \leq 0$, para todo $\overrightarrow{\mathbf{y}} \in W$, lo que implica, a su vez, que $\overrightarrow{\mathbf{l}} \cdot \overrightarrow{\mathbf{y}} \leq \overrightarrow{\mathbf{l}} \cdot F(\overrightarrow{\mathbf{x_c}})$, lo cual es una clara contradicción con lo demostrado previamente.

iii) Sea $\overrightarrow{\mathbf{y}} = (f_0(\overrightarrow{\mathbf{x_c}}), \ldots, \frac{1}{2} f_j(\overrightarrow{\mathbf{x_c}}), \ldots, f_p(\overrightarrow{\mathbf{x_c}})) \in \overline{W}$ con j entre 1 y p. Ahora, utilizando 4.16, obtenemos que $\lambda_0 f_0(\overrightarrow{\mathbf{x_c}}) + \ldots + \frac{1}{2}\lambda_j f_j(\overrightarrow{\mathbf{x_c}}) + \ldots + \lambda_p f_p(\overrightarrow{\mathbf{x_c}}) - [\lambda_0 f_0(\overrightarrow{\mathbf{x_c}}) + \ldots + \lambda_p f_p(\overrightarrow{\mathbf{x_c}})] \leq 0$ y de esta forma $-\frac{1}{2}\lambda_j f_j(\overrightarrow{\mathbf{x_c}}) \leq 0$.De este modo se tiene que $\lambda_j f_j(\overrightarrow{\mathbf{x_c}}) \geq 0$.

Por otro lado ,de ii), se tiene que cualquier j entre cero y p será positivo y las funciones que definen los conjuntos de restricción son negativas por lo que el producto de los lambda y estas funciones serán negativos.

De todo lo anterior se obtiene que $\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) = 0$ para $i = 1, \ldots, p$

♦

Corolario 4.1.2 *(Karush-Kuhn-Tucker). Si existe $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$ de tal forma que $f_i'(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} < 0$, para todo $i = 1, \ldots, p$ de manera que $f_i(\overrightarrow{\mathbf{x_c}}) = 0$, entonces las conclusiones de la regla de multiplicadores de John son satisfechas para $\lambda_0 > 0$*

Demostración: Definamos los siguientes conjuntos:

$$I = \{i \in \ \{1, \ldots, p\} : \ f_i(\overrightarrow{\mathbf{x_c}}) < 0\}$$
$$J = \{i \in \ \{1, \ldots, p\} : \ f_i(\overrightarrow{\mathbf{x_c}}) = 0\}$$

por lo que dentro del contexto de nuestro problema la unión de los conjuntos definidos constituyen los subíndices {1,...,p}. Asumamos que las reglas de multiplicadores de John se cumple para $\lambda_0 = 0$, con lo que de i) obtenemos que

$$\sum_{i=1}^{p} \lambda_i f_i'(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} = 0$$

De lo anteriormente enunciado, en conjunto con la condición iii) de las reglas de multiplicadores, podemos establecer que si $i \in I$, $\lambda_i = 0$.

Por otro lado, si $i \in J$, entonces tendremos que

$$\sum_{i=1}^{p} \lambda_i f_i'(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} = 0$$

con lo se sabe que todo término de la sumatoria será no positivo dado que $\lambda_i \geq 0$ y $f_i'(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} < 0$. En consecuencia, la sumatoria obliga a que $\lambda_i = 0$, para todo i en la unión de J e I, y además se tiene como supuesto que $\lambda_0 = 0$. De todo lo anterior se concluye que $l = (\lambda_0, \ldots, \lambda_p) = 0$, lo que contradice la hipótesis de que $l \neq 0$. ♦

4.2. Regla de multiplicadores con restriciones de igualdad y desigualdad

4.2.1. Reglas de multiplicadores de Carathéorory-John

En las reglas de multiplicadores vistas previamente sólo hemos considerado restricciones de desigualdad, no obstante, en este apartado añadiremos restricciones de igualdad. La regla que considdera estas condiciones es conocida como regla de multiplicadores de Carathéodory - Jonh y nos entrega una caracterización de la optimalidad de un punto en términos de la derivada estricta y sujeta a las restricciones establecidas. Sin embargo, para proceder a demostrar estas reglas de multiplicadores, utilizaremos el siguiente lema, el cual no será demostrado en este seminario puesto que escapa de nuestros objetivos principales.

Lema 4.2.1 *Sea $U \subset \mathbb{R}^n$ un conjunto abierto y además se tiene que $C \subset U$, donde C es convexo. Ahora definimos la función*

$$F : U \to \mathbb{R}^k,$$

tal que L sea la derivada estricta de F en el punto $\overrightarrow{\mathbf{x}_\alpha}$, el cual pertenece a $U \cap \overline{C}$ y también con $L(\overrightarrow{\mathbf{x}_\alpha}) \in intL(C)$. De esta manera, se tendrá que

$$F(\overrightarrow{\mathbf{x}_\alpha}) \in intF(C)$$

Supondremos que existe un subconjunto abierto U de $\mathbb{R}^n$, y además consideremos $f_0, f_1, \ldots, f_p$, $f_{p+1}, \ldots, f_{p+q}$ funciones reales definidas sobre U, estrictamente diferenciables en un punto $\overrightarrow{\mathbf{x_c}} \in U$. Si $\overrightarrow{\mathbf{x_c}}$ minimiza a f_0 sujeta a las restricciones de desigualdad $f_1 \leq 0, \ldots, f_p \leq 0$ y las q restricciones de igualdad $f_{p+1}, \ldots, f_{p+q}$, entonces existe $l = (\lambda_0, \ldots, \lambda_p, \lambda_{p+1}, \ldots, \lambda_{p+q})$ pertenecientes a $\mathbb{R}^{p+q+1}$, de tal manera que $l \not\equiv 0$, de tal forma que se cumple que:

(i) Si $f = \sum_{i=0}^{p} \lambda_i f_i$ con $\overrightarrow{\mathbf{x_c}} \in U$ entonces $f'(\overrightarrow{\mathbf{x_c}}) = 0$

(ii) $\lambda_i \geq 0,$ para todo, $i = 0, 1, ..., p$ y asimismo se satisface que

(iii) $\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) = 0$, donde $i = 0, 1, ..., p$

Demostración: Consideremos, en primer lugar, que sin perder generalidad podemos tener que $\overrightarrow{\mathbf{x_c}} = 0$, que el conjunto U es convexo y que además se tiene

$$F(\overrightarrow{\mathbf{x_c}}) = 0 \,\text{con}\, F = (f_0, \ldots, f_{p+q}) : U \to \mathbb{R}^{p+q+1}$$

Ahora, se define el siguiente conjunto:

$$W = \{\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_p + q) \in \mathbb{R}^{p+q+1} : \eta_0 < 0, \eta_i \leq 0 \; i = 1, \ldots, p \text{ y } \eta_i = 0 \; i = p+1, \ldots, p+q\}$$

Por la optimalidad de $\overrightarrow{\mathbf{x_c}} = 0$ en el conjunto U, se tiene que

$$F(U) \cap W = \emptyset \text{ y además } F(0) = 0 \in F(U) \cap \overline{W}$$

Asimismo, tenemos que L es la derivada estricta de F en el punto $F(\overrightarrow{\mathbf{x_c}}) = F(0) = 0 \in U$ y definimos la aplicación afín A dada por

$$A(\overrightarrow{\mathbf{x}}) = F(0) + L\overrightarrow{\mathbf{x}} = L\overrightarrow{\mathbf{x}}$$

En primer lugar, veamos que los conjuntos $L(U)$ y W son separados. En efecto, supongamos que $L(U)$ y W no son separados, en ese caso lo que obtenemos es, por el teorema de Hanh-Banach, $0 \in \; int[L(U) - W]$. Ahora definamos la siguiente función:

$$\psi : \mathbb{R}^n \times \mathbb{R}^{p+q+1} \longrightarrow \mathbb{R}^{p+q+1}$$

por la forma

$$\psi(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) = F(\overrightarrow{\mathbf{x}}) - \overrightarrow{\mathbf{y}},$$

con el objetivo de obtener una contradicción con la caracterización de optimalidad de $\overrightarrow{\mathbf{x_c}} = 0$. Para este fin procederemos a utilizar el lema 4.2.1, para lo cual veremos, en primer lugar, que la derivada estricta de ψ existe y que $U \times W$ es convexo.

Demostraremos a continuación que $U \times W$ es convexo. En efecto, para que esto ocurra, se debe verificar que para todo $\overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{b}} \in U \times W$, se cumple que $\lambda \cdot \overrightarrow{\mathbf{a}} + (1-\lambda)\overrightarrow{\mathbf{b}} \in U \times W$.Es más, notemos que $\overrightarrow{\mathbf{a}} = (\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}})$ y $\overrightarrow{\mathbf{b}} = (\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}})$, de esta manera tenemos que

$$\lambda\overrightarrow{\mathbf{a}} + (1-\lambda)\overrightarrow{\mathbf{b}} = (\lambda\overrightarrow{\mathbf{x_1}} + (1-\lambda)\overrightarrow{\mathbf{y_1}}, \lambda\overrightarrow{\mathbf{x_2}} + (1-\lambda)\overrightarrow{\mathbf{y_2}})$$

y se sabe que cada componente del vector anterior pertenecen a U y a W respectivamente. De esta forma,

$$\lambda\overrightarrow{\mathbf{a}} + (1-\lambda)\overrightarrow{\mathbf{b}} \in U \times W$$

por lo que $U \times W$ es convexo.

Por la diferenciabilidad estricta de F podemos escribir

$$||F(\overrightarrow{\mathbf{x_1}}) - F(\overrightarrow{\mathbf{x_2}}) - L(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}})|| \leq \epsilon \|\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}\|$$

y al sumar convenientemente cero, podemos reescribirlo como:

$$\|F(\overrightarrow{\mathbf{x_1}}) - F(\overrightarrow{\mathbf{x_2}}) - (\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}}) - [L(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}) - (\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}})]\| \leq \epsilon \|\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}\|$$

y esto implica que, dado que $\psi(\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}}) = F(\overrightarrow{\mathbf{x_1}}) - \overrightarrow{\mathbf{y_1}}$ y análogamente para $\psi(\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}})$, entonces obtendremos

$$\|\psi(\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}}) - \psi(\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}}) - [L(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}) - (\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}})]\| \leq \epsilon \|(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}})\|.$$

Entonces ψ es estrictamente diferenciable en el punto $(0,0)$, donde $\psi'(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})\mid_{(0,0)} = [L, -I]$, siendo I la aplicación identidad. Producto del lema 4.2.1,

$$0 = \psi((0,0)) \in \ int\psi(U \times W) \subseteq \psi(U \times W),$$

lo cual a su vez implica que

$$0 = F(\overrightarrow{\mathbf{x}}) - \overrightarrow{\mathbf{y}} \ \text{ para algún} \overrightarrow{\mathbf{x}} \in U$$

y con $\overrightarrow{\mathbf{y}} \in W$, que a su vez significa que $\overrightarrow{\mathbf{y}} \in (F(U) \cap W)$, por lo que $F(U) \cap W \neq \emptyset$, lo que contradice la caracterización de optimalidad en el punto $\overrightarrow{\mathbf{x_c}} = 0$

De todo lo anterior tenemos que existe $l = (\lambda_0, \ldots, \lambda_p + q) \in \mathbb{R}^{p+q+1}$ que verifica

$(a) l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] \geq 0 \ \forall \ \overrightarrow{\mathbf{y}} \in L(U)$

$(b) l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] \leq 0 \ \forall \ \overrightarrow{\mathbf{y}} \in W$

Ahora probaremos i), para esto veamos que en a) tenemos que

$$l \cdot [L\overrightarrow{\mathbf{x}} - F(0)] = l \cdot [L\overrightarrow{\mathbf{x}}] \geq 0,$$

lo que implica que $l \cdot L(\overrightarrow{\mathbf{x}}) \geq 0$.

Además, por la definición de la derivada estricta podemos considerar una vecindad de $\overrightarrow{\mathbf{x_c}} = 0$, junto con la propiedad de las normas en $\mathbb{R}^m$ tendremos que $||\overrightarrow{\mathbf{x}} - 0|| = ||0 - \overrightarrow{\mathbf{x}}||$, por lo que tendremos $L\overrightarrow{\mathbf{x}} = L - \overrightarrow{\mathbf{x}}$, y de esta manera logramos tener que $l \cdot L - \overrightarrow{\mathbf{x}} \geq 0$ lo que directamente implica que $l \cdot L\overrightarrow{\mathbf{x}} \leq 0$. Finalmente, de las afirmaciones anteriores obtenemos que $l \cdot L\overrightarrow{\mathbf{x}} = 0$, o dicho de otro modo, lo expuesto previamente sólo se cumple cuando $f'(\overrightarrow{\mathbf{x_c}}) = 0$

Prosiguiendo con la demostración, probaremos la condición ii). Para este fin supongamos que existe algún λ_j, con j entre cero y p, que cumpla $\lambda_j < 0$. De este modo el término $\lambda_j \eta_j$ es positivo y si observamos la suma $\displaystyle\sum_{i=0}^{j-1} \lambda_i \eta_i + \sum_{i=j+1}^{p} \lambda_i \eta_i$ tenemos que está acotada por cero. Consiguientemente al hacer tender a $\eta_j \to -\infty$, además $\lambda_j \eta_j \to +\infty$. Por último, se obtiene que $\displaystyle\sum_{i=0}^{p+q} \lambda_i \eta_i$ no esta acotada, lo que contradice la igualdad b).

Para concluir la demostración, mostraremos la prueba de la condición iii), para esto consideremos el vector

$$\overrightarrow{\mathbf{y}} = (f_0(\overrightarrow{\mathbf{x_c}}), \ldots, \frac{1}{2} f_j(\overrightarrow{\mathbf{x_c}}), \ldots, f_{p+q}(\overrightarrow{\mathbf{x_c}}))$$

con lo que podemos escribir la siguiente expresión basados en la desigualdad b):

$$\lambda_0 f_0(0) + \ldots + \frac{1}{2}\lambda_j f_j(0) + \ldots + \lambda_{p+q} f_{p+q} - [\lambda_0 f_0(0) + \ldots + \lambda_j f_j(0)] \leq 0,$$

con lo que se consigue

$$-\frac{1}{2}\lambda_j f_j(0) \leq 0,$$

lo que implica que

$$\lambda_j f_j(0) \geq 0,$$

vale decir

$$f_j(0) \geq 0,$$

dado que para cualquier λ_i, con $i = 1, \ldots, p$, son mayores o iguales que cero, por lo que esto contradice las condiciones de la hipótesis. Luego, $\lambda_i \cdot f_i(\overrightarrow{\mathbf{x_c}}) = 0$, para $i = 1, \ldots, p$.

Por lo anteriormente expuesto se demuestran las reglas de multiplicadores de Carathéodory-John. ♦

Corolario 4.2.1 *(Karush-Kuhn-Tucker) Si $f'_{p+1}(\overrightarrow{\mathbf{x_c}}), \ldots, f'_{p+q}(\overrightarrow{\mathbf{x_c}})$ son linealmente independientes y algún $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$ satisfaciendo*

(a)$f'_i(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} < 0$ para los $i \in \{1, \ldots, p\}$ tales que $f_i(\overrightarrow{\mathbf{x_c}}) = 0$

(b)$f'_i(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} = 0$ para los $i \in \{p+1, \ldots, p+q\}$

Entonces son satisfechas las reglas de multiplicadores de Carathéodory-John si $\lambda_0 > 0$

Demostración: Para demostrar este corolario procederemos por el principio de contradicción, vale decir, supongamos que con $\lambda_0 = 0$ se cumplen las reglas de multiplicadores de Carathéodory-John. Ahora supongamos $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$ donde se tenga que aquel vector satisface las condiciones a) y b). Entonces podemos afirmar que $\sum_{i=1}^{p} \lambda_i f'_i(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} = 0$, dado que $\sum_{i=p+1}^{p+q} f'_i(\overrightarrow{\mathbf{x_c}}) = 0$, por la condición b) y por el primer enunciado de las reglas de Carathéodory-John.

A continuación definamos

$$\left.\begin{array}{l} I = \{i \in \{1, \ldots, p\} : f_i(\overrightarrow{\mathbf{x_c}}) < 0\} \\ J = \{i \in \{1, \ldots, p\} : f_i(\overrightarrow{\mathbf{x_c}}) = 0\} \end{array}\right. \Longrightarrow I \cup J = \{1, \ldots, p\}$$

Dado que por iii) se tiene $\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) = 0$ y si además se tiene que si i pertenece a I cumple que $f_i(\overrightarrow{\mathbf{x_c}}) < 0$, entonces λ_i no tiene otra opción más que ser cero. De esto se desprende $\lambda_i = 0$ para $i \in I$.

Por otra parte se sabe que toda derivada estricta es de Fréchet por lo que por

$\sum_{i=1}^{p} \lambda_i f_i'(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} = 0$ en esta parte la demostración en curso se puede reducir a las reglas de multiplicadores de John, esto equivale a señalar que si i pertenece a J se tendrá que cada $f_i'(\overrightarrow{\mathbf{x_c}})\overrightarrow{\mathbf{h}} < 0$, así cada λ_j que forma parte de la suma $\sum_{j \in J} \lambda_j f_j(\overrightarrow{\mathbf{x_c}}) = 0$ se ve obligada a ser cero.

De todo lo anterior podemos concluir que $\lambda_0 = 0$ y $\lambda_i = 0$ para todo $i \in I \cup J$ contradiciendo el hecho de que $\overrightarrow{\mathbf{l}} = (\lambda_0, \ldots, \lambda_p) \neq 0$. ♦

Capítulo 5

Lagrangeano y Reglas de Multiplicadores

5.1. Formalización del Lagrangeano

En el capítulo anterior hemos estudiado problemas en optimización que consideran por un lado restricciones de desigualdad, y por otra parte, problemas con restricciones de igualdad. Si ponemos atención en todas las reglas de multiplicadores definimos un funcional al que llamamos f, que considera el producto interno entre un vector de $\mathbb{R}^{k+1}$ y la función que nos entrega vectores cuyas componentes son la función objetivo f_0 y las k-funciones que restringen al problema inicial. Y si observamos detenidamente, podremos ver que f - en todos los casos - depende de los vectores del conjunto U, de las componentes del vector $\overrightarrow{\mathbf{l}}$ y con especial atención del λ_0 presente en los corolarios de Karush Kuhn Tucker.

Siguiendo esta misma lógica podemos definir un funcional al que denominaremos ***lagrangeano***. Para esto, considérese, en primer lugar

$$\left.\begin{array}{ll} & \min f_0(\overrightarrow{\mathbf{x}}) \\ \text{sujeto a} & \\ & f_i(\overrightarrow{\mathbf{x}}) \leq 0 \quad i = 1, ..., p \end{array}\right\} T$$

con lo que el ***lagrangeano*** queda definido de la siguiente forma;
$\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}$

$$(\overrightarrow{\mathbf{x}}, \overrightarrow{\lambda_{\mathbf{0}}}, \overrightarrow{\lambda}) \mapsto \mathcal{L}((\overrightarrow{\mathbf{x}}, \overrightarrow{\lambda_{\mathbf{0}}}, \overrightarrow{\lambda})) = \sum_{k=0}^{p} \lambda_k f_k(\overrightarrow{\mathbf{x}})$$

con $\overrightarrow{\lambda} = (\lambda_1, \ldots, \lambda_p)$.

$$\left.\begin{array}{ll} & \min f_0(\overrightarrow{\mathbf{x}}) \\ \text{sujeto a} & \\ & f_i(\overrightarrow{\mathbf{x}}) = 0 \quad i = 1, \ldots, q \end{array}\right\} T_I$$

así nuestro funcional ***lagrangeano*** es
$\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^q \to \mathbb{R}$

$$(\overrightarrow{\mathbf{x}}, \overrightarrow{\lambda_{\mathbf{0}}}, \overrightarrow{\lambda}) \mapsto \mathcal{L}((\overrightarrow{\mathbf{x}}, \overrightarrow{\lambda_{\mathbf{0}}}, \overrightarrow{\lambda})) = \sum_{k=0}^{q} \lambda_k f_k(\overrightarrow{\mathbf{x}})$$

considerando a $\overrightarrow{\lambda} = (\lambda_1, \ldots, \lambda_q)$.

$$\left.\begin{array}{ll} & \text{mín } f_0(\overrightarrow{\mathbf{x}}) \\ \text{sujeto a} & \\ & f_i(\overrightarrow{\mathbf{x}}) \leq 0 \qquad i = 1, ..., p \\ & f_i(\overrightarrow{\mathbf{x}}) = 0 \quad i = p+1, \ldots, p+q \end{array}\right\} T_G$$

caso en el que el ***lagrangeano*** es definido de la siguiente manera:
$\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{p+q} \to \mathbb{R}$

$$(\overrightarrow{\mathbf{x}}, \overrightarrow{\lambda_{\mathbf{0}}}, \overrightarrow{\lambda}) \mapsto \mathcal{L}((\overrightarrow{\mathbf{x}}, \overrightarrow{\lambda_{\mathbf{0}}}, \overrightarrow{\lambda})) = \sum_{k=0}^{p+q} \lambda_k f_k(\overrightarrow{\mathbf{x}})$$

con $\overrightarrow{\lambda} = (\lambda_1, \ldots, \lambda_{p+q})$.

Si observamos con detención nos percataremos que λ_0 puede asumir cualquier valor real. Así, tendremos que si este valor real es igual a cero, en cualquiera de los problemas propuestos anteriormente, entonces se dice que el problema es ***anormal***. En cambio, cuando λ_0 es distinto de cero, se dice que es ***normal***.

5.2. Reglas de Multiplicadores Convexa

Supongamos U convexo abierto $f_0, \ldots, f_p$, p+1 funciones convexas sobre U. Además se tiene que $\overline{\mathbf{x}}$ soluciona el problema T, entonces existen multiplicadores de Lagrange $\hat{\lambda}_0, \hat{\lambda}$ que no se anulan simultáneamente de tal forma que se cumple

$a) min_{x\in U}\mathcal{L}(\overrightarrow{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda}) = \mathcal{L}(\overline{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda})$ (Principio de mínimo)

$b) \hat{\lambda}_i \geq 0 \quad i = 1, \ldots, p$ (Condición de no negatividad)

$c) \hat{\lambda}_i f_i(\overline{\mathbf{x}}) = 0 \quad i = 1, \ldots, p$ (Condición de no rigidez complementar)

Notemos que esta es la reescritura inmediata de la regla de multiplicadores convexa en términos del lagrangeano, dado que, reinterpretando lo enunciado anteriormente con el lagrangeano se tiene

$$\mathcal{L}(\overline{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda}) = \lambda_0 f_0(\overline{\mathbf{x}}) \leq \mathcal{L}(\overrightarrow{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda}) \quad \forall \quad \overrightarrow{\mathbf{x}} \in U$$

En consecuencia, para probar esta regla escrita de este modo, se realiza de manera análoga a la realizada en el capítulo anterior.

Corolario 5.2.1 *(Karush Kuhn Tucker)*

1. *Sí $\hat{\lambda}_0 \neq 0$ las condiciones a), b) y c) son suficientes para que un punto admisible sea solución del problema T.*

2. *Para tener $\hat{\lambda}_0 \neq 0$ es suficiente el que exista un punto $\overrightarrow{\mathbf{y}} \in U$ tal que se verifique $f_i(\overrightarrow{\mathbf{y}}) < 0$ con $i = 1.\ldots,p$ (Condición de Slater)*

Demostración:

1. Sabiendo que $\lambda_0 \neq 0$ se puede decir, por a), que $\min\limits_{x\in U} \mathcal{L}(\overrightarrow{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda}) = \mathcal{L}(\overline{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda})$, los que implica por la definición de optimalidad que
$$\lambda_0[f_0(\overline{\mathbf{x}}) - f_0(\overrightarrow{\mathbf{x}})] \leq \sum_{k=1}^{p} \lambda_k f_k(\overrightarrow{\mathbf{x}}).$$
Analizando el caso de que $\lambda_0 > 0$ se puede obtener que
$$f_0(\overline{\mathbf{x}}) \leq f_0(\overrightarrow{\mathbf{x}}) \ \ \forall \ \ \overrightarrow{\mathbf{x}} \in U$$
En caso de que $\lambda_0 < 0$ tendremos que
$$f_0(\overline{\mathbf{x}}) \geq f_0(\overrightarrow{\mathbf{x}}) \ \ \forall \ \ \overrightarrow{\mathbf{x}} \in U,$$
lo que evidentemente responde a una solución máxima, no obstante sabemos que si definimos una función real que tenga como dominio a U, a la que llamaremos $\tilde{f}_0$, tal que $\tilde{f}_0 = f_0$, entonces el asunto en cuestión se reducirá a un problema de minimización, con lo que se concluye que bajo las hipótesis el punto $\overline{x}$ es una solución al problema T.

2. Supongamos que $\hat{\lambda}_0 = 0$. Entonces, se tiene que
$$0 = \mathcal{L}(\overline{\mathbf{x}}, 0, \hat{\lambda}).$$
De esto tendremos que el lagrangeano
$$\mathcal{L}(\overline{\mathbf{x}}, 0, \hat{\lambda}) \leq \mathcal{L}(\overrightarrow{\mathbf{x}}, 0, \hat{\lambda}) \ \ \ \forall \ \ \overrightarrow{\mathbf{x}} \in U$$
Producto de lo anterior, obtendremos que $0 \leq \mathcal{L}(\overrightarrow{\mathbf{x}}, 0, \hat{\lambda})$.
Ahora, por la condición de Slater, existe $\overrightarrow{\mathbf{y}} \in U$ tal que $f_i(\overrightarrow{\mathbf{y}}) < 0$. Si consideramos que $\lambda_i \geq 0$ con $i = 1,\ldots,p$, entonces
$$\sum_{i=1}^{p} \lambda_i f_i(\overrightarrow{\mathbf{y}}) \leq 0.$$
Si escribimos la primera desigualdad como una sumatoria, tendremos que
$$0 \leq \sum_{i=1}^{p} \lambda_i f_i(\overrightarrow{\mathbf{x}}),$$
para todo $\overrightarrow{\mathbf{x}} \in U$, lo que contradice lo mencionado por la condición de Slater, por lo tanto
$$\hat{\lambda}_0 \neq 0.$$

♦

5.3. Reglas de Carathéodory-John

De igual forma como la Regla de Multiplicadores Convexa, tenemos que las Reglas de Carathéodory-John pueden ser reinterpretadas al incorporar la definición del Lagrangeano en su definición.

De esta manera, sea $U \subseteq \mathbb{R}^n$ un subconjunto abierto, y considérese $f_0, \ldots, f_p, \ldots, f_{p+q}$, $p + q + 1$ funciones reales definidas sobre U, donde cada una de aquellas funciones son estrictamente diferenciables en un punto $\overline{\mathbf{x}} \in U$. Si además se tiene que el punto $\overline{\mathbf{x}}$ del subconjunto U resuelve el problema T_G, entonces, existen los Multiplicadores de Lagrange $\hat{\lambda}_0$ y $\hat{\lambda}$ que no se anulan simultáneamente de tal forma que

$a)\nabla_x \mathcal{L}(\overline{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda}) = 0$ (Condición estacionaria)

$b)\hat{\lambda}_k \geq 0 \quad k = 1, \ldots, p$ (Condición de no negatividad)

$c)\hat{\lambda}_k f_k(\overline{\mathbf{x}}) = 0 \quad k = 1, \ldots, p$ (Condición de no rigidez)

Demostración:

Observación 5.3.1 *Antes de comenzar con la demostración es posible observar que*

$$\mathcal{L}(\overrightarrow{\mathbf{x}}, \hat{\lambda}_0, \hat{\lambda}) = \sum_{k=0}^{p+q} \lambda_k f_k(\overrightarrow{\mathbf{x}}) = (\lambda_0, \ldots, \lambda_{p+q}) \cdot (f_0(\overrightarrow{\mathbf{x}}), \ldots, f_{p+q}(\overrightarrow{\mathbf{x}})),$$

En otras palabras se puede decir que el lagrangeano está definido por el producto interno de dos vectores de $\mathbb{R}^{p+q+1}$. Bajo esto procederemos a realizar la demostración de manera análoga a la regla de Carathéodory-John estudiada previamente en 4.2.

a)Definamos:

$F(U) = \{F = (f_0, \ldots, f_{p+q}) \in \mathbb{R}^{p+q+1} :$ para $\overrightarrow{\mathbf{x}} \in U, F(\overrightarrow{\mathbf{x}}) = (f_0(\overrightarrow{\mathbf{x}}), \ldots, f_{p+q}(\overrightarrow{\mathbf{x}}))\}$ y además consideremos

$$W = \{\overrightarrow{\mathbf{y}} = (\eta_0, \ldots, \eta_{p+q}) \in \mathbb{R}^{p+q+1} : \eta_0 < f_0(\overline{\mathbf{x}}), \eta_i \leq 0,\ i = 1, \ldots, p \text{ y } \eta_i = 0\ i = p+1, \ldots, p+q\}$$

Suponiendo que $\overline{\mathbf{x}} = 0$ y además $F(\overline{\mathbf{x}}) = 0$, entonces por la optimalidad de $\overline{\mathbf{x}}$ se tiene que

$$F(U) \cap W = \emptyset \text{ y además } F(0) = 0 \in F(U) \cap \overline{W}$$

Asimismo, tenemos que L es la derivada estricta de F en el punto $F(\overrightarrow{\mathbf{x_c}}) = F(0) = 0 \in U$ y definimos la aplicación afín A dada por

$$A(\overrightarrow{\mathbf{x}}) = F(0) + L\overrightarrow{\mathbf{x}} = L\overrightarrow{\mathbf{x}}$$

En primer lugar, veamos que los conjuntos $L(U)$ y W son separados. En efecto, supongamos que $L(U)$ y W no son separados, en ese caso lo que obtenemos es, por el teorema de Hanh-Banach, $0 \in\ int[L(U) - W]$. Ahora definamos la siguiente función:

$$\psi : \mathbb{R}^n \times \mathbb{R}^{p+q+1} \longrightarrow \mathbb{R}^{p+q+1}$$

por la forma

$$\psi(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) = F(\overrightarrow{\mathbf{x}}) - \overrightarrow{\mathbf{y}},$$

con el objetivo de obtener una contradicción con la caracterización de optimalidad de $\overline{\mathbf{x}} = 0$. Para este fin procederemos a utilizar el lema 4.2.1, para lo cual veremos, en primer lugar, que la derivada estricta de ψ existe y que $U \times W$ es convexo.

Demostraremos a continuación que $U \times W$ es convexo. En efecto, para que esto ocurra, se debe verificar que para todo $\overrightarrow{\mathbf{a}}\,\overrightarrow{\mathbf{b}} \in U \times W$, se cumple que $\lambda \cdot \overrightarrow{\mathbf{a}} + (1-\lambda)\overrightarrow{\mathbf{b}} \in U \times W$.Es más, notemos que $\overrightarrow{\mathbf{a}} = (\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}})$ y $\overrightarrow{\mathbf{b}} = (\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}})$, de esta manera tenemos que

$$\lambda\overrightarrow{\mathbf{a}} + (1-\lambda)\overrightarrow{\mathbf{b}} = (\lambda\overrightarrow{\mathbf{x_1}} + (1-\lambda)\overrightarrow{\mathbf{y_1}}, \lambda\overrightarrow{\mathbf{x_2}} + (1-\lambda)\overrightarrow{\mathbf{y_2}})$$

y se sabe que cada componente del vector anterior pertenecen a U y a W respectivamente. De esta forma,

$$\lambda\overrightarrow{\mathbf{a}} + (1-\lambda)\overrightarrow{\mathbf{b}} \in U \times W$$

por lo que $U \times W$ es convexo.

Si consideramos que F posee derivada estricta en el punto $\overline{\mathbf{x}}$ podemos escribir

$$||F(\overrightarrow{\mathbf{x_1}}) - F(\overrightarrow{\mathbf{x_2}}) - L(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}})|| \leq \epsilon\|\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}\|$$

y al sumar convenientemente cero, podemos reescribirlo como:

$$\|F(\overrightarrow{\mathbf{x_1}}) - F(\overrightarrow{\mathbf{x_2}}) - (\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}}) - [L(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}) - (\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}})]\| \leq \epsilon\|\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}\|$$

y esto implica que, dado que $\psi(\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}}) = F(\overrightarrow{\mathbf{x_1}}) - \overrightarrow{\mathbf{y_1}}$ y análogamente para $\psi(\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}})$, entonces obtendremos

$$\|\psi(\overrightarrow{\mathbf{x_1}}, \overrightarrow{\mathbf{y_1}}) - \psi(\overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_2}}) - [L(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}) - (\overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}})]\| \leq \epsilon\|(\overrightarrow{\mathbf{x_1}} - \overrightarrow{\mathbf{x_2}}, \overrightarrow{\mathbf{y_1}} - \overrightarrow{\mathbf{y_2}})\|.$$

Entonces ψ es estrictamente diferenciable en el punto $(0,0)$, donde $\psi'(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}})\mid_{(0,0)} = [L, -I]$, siendo I la aplicación identidad. Producto del lema 4.2.1,

$$0 = \psi((0,0)) \in \ int\psi(U \times W) \subseteq \psi(U \times W),$$

lo cual a su vez implica que

$$0 = F(\overrightarrow{\mathbf{x}}) - \overrightarrow{\mathbf{y}} \ \text{ para algún} \overrightarrow{\mathbf{x}} \in U$$

y con $\overrightarrow{\mathbf{y}} \in W$, que a su vez significa que $\overrightarrow{\mathbf{y}} \in (F(U) \cap W)$, por lo que $F(U) \cap W \neq \emptyset$, lo que contradice la caracterización de optimalidad en el punto $\overline{\mathbf{x}} = 0$

De todo lo anterior tenemos que existe $l = (\lambda_0, \ldots, \lambda_p + q) \in \mathbb{R}^{p+q+1}$ que verifica

$(1) l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] \geq 0 \ \forall \ \overrightarrow{\mathbf{y}} \in L(U)$

$(2) l \cdot [\overrightarrow{\mathbf{y}} - F(\overrightarrow{\mathbf{0}})] \leq 0 \ \forall \ \overrightarrow{\mathbf{y}} \in W$

Para proseguir con esta demostración veamos que en 1) tenemos que

$$l \cdot [L\overrightarrow{\mathbf{x}} - F(0)] = l \cdot [L\overrightarrow{\mathbf{x}}] \geq 0,$$

lo que implica que $l \cdot L(\overrightarrow{\mathbf{x}}) \geq 0$.

Además, por la definición de la derivada estricta podemos considerar una vecindad de $\overline{\mathbf{x}} = 0$, sumado a la propiedad de las normas en $\mathbb{R}^m$ tendremos que $||\overrightarrow{\mathbf{x}} - 0|| = ||0 - \overrightarrow{\mathbf{x}}||$, por lo que tendremos $L\overrightarrow{\mathbf{x}} = L - \overrightarrow{\mathbf{x}}$, y de esta manera logramos tener que $l \cdot L - \overrightarrow{\mathbf{x}} \geq 0$ lo que directamente implica que $l \cdot L\overrightarrow{\mathbf{x}} \leq 0$. Finalmente, tenemos que de las afirmaciones anteriores obtenemos que $l \cdot L\overrightarrow{\mathbf{x}} = 0$, o dicho de otro modo, lo expuesto previamente sólo se cumple cuando $\nabla_x \mathcal{L}(\overline{\mathbf{x}}, \lambda_0, \lambda) = 0$

Prosiguiendo con la demostración, probaremos la condición b)suponiendo un λ_0 fijo. Para este fin supongamos que existe algún λ_j, con j entre uno y p, que cumpla $\lambda_j < 0$. De este modo el término $\lambda_j \eta_j$ es positivo y si observamos la suma $\sum_{k=0}^{j-1} \lambda_k \eta_k + \sum_{k=j+1}^{p} \lambda_k \eta_k$ tenemos que está acotada por cero. Consiguientemente al hacer tender a $\eta_j \to -\infty$, además $\lambda_j \eta_j \to +\infty$. Por último, se obtiene que $\sum_{k=0}^{p+q} \lambda_k \eta_k$ no esta acotada, lo que contradice la igualdad 2).

Para concluir la demostración, mostraremos la prueba de la condición c). Con este propósito, consideremos el vector

$$\overrightarrow{\mathbf{y}} = (f_0(\overline{\mathbf{x}}), \ldots, \frac{1}{2} f_j(\overline{\mathbf{x}}), \ldots, f_{p+q}(\overline{\mathbf{x}}))$$

con lo que podemos escribir la siguiente expresión basados en la desigualdad 2):

$$\lambda_0 f_0(0) + \ldots + \frac{1}{2}\lambda_j f_j(0) + \ldots + \lambda_{p+q} f_{p+q} - [\lambda_0 f_0(0) + \ldots + \lambda_j f_j(0)] \leq 0,$$

con lo que se consigue

$$-\frac{1}{2}\lambda_j f_j(0) \leq 0,$$

lo que implica que

$$\lambda_j f_j(0) \geq 0,$$

vale decir

$$f_j(0) \geq 0,$$

dado que para cualquier λ_k, con $k = 1, \ldots, p$, son mayores o iguales que cero, por lo que esto contradice las condiciones de la hipótesis. Luego, $\lambda_k \cdot f_k(\overline{\mathbf{x}}) = 0$, para $k = 1, \ldots, p$. ♦

Antes de continuar resaltemos el hecho que la demostración realizada anteriormente es homóloga a la demostración de las reglas de multiplicadores de Carathéodory John vista en el capítulo anterior, con lo que por extensión se puede decir que la demostración del corolario que escribiremos a continuación es análoga a la efectuada en el capítulo anterior.

Corolario 5.3.1 *(Karush-Kuhn-Tucker) Si $f'_{p+1}(\overline{\mathbf{x}}), \ldots, f'_{p+q}(\overline{\mathbf{x}})$ son linealmente independientes y además existe algún $\overrightarrow{\mathbf{h}} \in \mathbb{R}^n$ que satisface*

a)$f'_k(\overline{\mathbf{x}})\overrightarrow{\mathbf{h}} < 0$ *para los* $k \in \{1, \ldots, p\}$ *tales que* $f_k(\overline{\mathbf{x}}) = 0$

b)$f'_k(\overline{\mathbf{x}})\overrightarrow{\mathbf{h}} = 0$ *para* $k \in \{p+1, \ldots, p+q\}$

Capítulo 6

Aplicación del Teorema de Karush-Kuhn-Tucker

En este capítulo, procederemos a mostrar diversas aplicaciones relacionados con el teorema de Karush Kuhn Tucker. Por este motivo es que comenzaremos resolviendo un problema de una función de una sola variable, el cual no presenta ninguna restricción de igualdad y/o de desigualdad. Esto, dado que será una introducción a un problema de optimización. Luego, plantearemos y resolveremos un problema relacionado con los multiplicadores de Lagrange, para finalizar con la aplicación de las condiciones del problema de KKT.

Aplicación 6.1 *Supongamos que tenemos dos rectas que se cortan perpendicularmente. Por cada una avanzan, simultáneamente dos móviles con velocidades v_A y v_B. Se dirigen al punto de encuento de ambas rectas partiendo de unas distancias, con respecto a la intersección, a y b respectivamente. Resolveremos el problema que consiste en encontrar el instante en que la distancia entre los móviles es mínima.*

Solución 6.1 *El problema que se nos presenta se grafica en la siguiente figura*

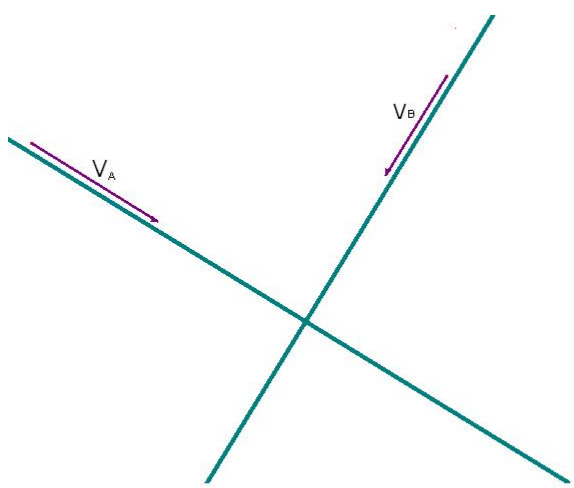

Figura 6.1: Móviles que se trasladan sobre las rectas perpendiculares

Sabemos que la fórmula de velocidad es $v = \frac{d}{t}$, donde v es la velocidad, d la distancia y t el tiempo. De aquí obtenemos que de manera general la distancia que enunciada en términos *de velocidad y tiempo según la siguiente expresión:* $d = v \cdot t$*, y así, cuando tenemos una distancia inicial -*d_1*- preestablecida podemos escribir que* $d = d_1 - v \cdot t$*, en nuestro caso tendremos que la distancia que le resta al móvil A para llegar a la intersección está dada por* $d_A = a - v_A \cdot t$ *y la del móvil B es* $d_B = b - v_B \cdot t$*, finalmente la distancia entre los móviles se puede determinar por medio del teorema de Pitágoras, vale decir,*

$$d = \sqrt{(a - v_A \cdot t)^2 + (b - v_B \cdot t)^2}$$

entonces esta longitud se puede expresar en función del tiempo, es decir

$$\begin{aligned} d : U \subset \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto d(t) = \sqrt{(a - v_A \cdot t)^2 + (b - v_B \cdot t)^2} \end{aligned}$$

Además, $d(t)$ *puede ser reescrita como*

$$d(t) = [a^2 - 2av_At + v_A^2t^2 + b^2 - 2bv_Bt + v_B^2t^2]^{\frac{1}{2}}$$

Notemos que la función no tiene restricciones, además es una función real y sin lugar a dudas es de clase C^1*, por lo que podemos utilizar el criterio de la primera derivada, que implica que cuando la primera derivada de la función es cero o no se encuentra determinada, estamos en presencia de*

un punto mínimo (en nuestro caso en particular). De esa forma, tendremos que:

$$d'(t) = \frac{1}{2} \cdot \frac{2(-av_A + v_A^2t - bv_B + v_B^2t)}{[a^2 - 2av_At + v_A^2t^2 + b^2 - 2bv_Bt + v_B^2t^2]^{\frac{1}{2}}} = 0$$

Lo que implica necesariamente que $-av_A + v_A^2t - bv_B + v_B^2t = 0$*, y entonces tendremos que* $t(v_A^2 + v_B^2) = av_A + bv_B$*, por lo que las distancias serán mínimas en el tiempo t igual a*

$$t = \frac{av_A + bv_B}{v_A^2 + v_B^2}$$

Anteriormente hemos estudiado un problema en el que se debe minimizar una función sin restricciones, abordado según el criterio de la primera derivada. No obstante, existen casos en los que existen funciones que poseen restricciones, uno de los eventos más sencillos es cuando la función tiene restricciones de igualdad. Para resolver este problema recurriremos a lo que se conoce como multiplicadores de Lagrange, planteando la función lagrangeana. Así, a continuación estudiaremos un problema en el que se requiere encontrar un mínimo sujeto a algunas restricciones de igualdad.

Aplicación 6.2 *Hallar el punto del paraboloide*

$$\begin{aligned} z : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto z(x, y) = (x - 2)^2 + \frac{1}{4}(y - 3)^2 + 5 \end{aligned}$$

más próximo al plano $x + y + z = 0$*.*

Solución 6.2 *Observemos que se desea minimizar una función real que calcula la distancia entre el punto y el plano en cuestión, a esta función la llamaremos* Φ *y estará definida de la siguiente manera:*

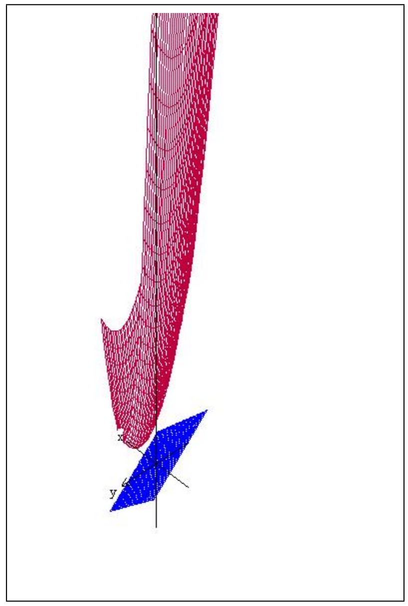

Figura 6.2: Gráfico del paraboloide y del plano x+y+z=0

$$\begin{array}{rcl} \Phi: \quad \mathbb{R}^3 \times \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ ((x_1, y_1, z_1), (x_2, y_2, z_2)) & \longmapsto & \Phi((x_1, y_1, z_1), (x_2, y_2, z_2)) = ||((x_1, y_1, z_1), (x_2, y_2, z_2))||_2 \end{array}$$

Además, supongamos que - para la resolución de esta aplicación - los puntos son (a, b, c) en el paraboloide y (d, e, f) en el plano, por lo que es posible reescribir la función de distancia de la siguiente forma:

$$\text{mín}\, \Phi = \sqrt{(a-d)^2 + (b-e)^2 + (c-f)^2}$$

sujeto a

$$g_1(a, b, c) = -(a-2)^2 - \frac{1}{4}(b-3)^2 + c = 5$$

$$g_2(d, e, f) = d + e + f = 0$$

Para poder resolver este problema, y producto de las características de la función Φ, podemos definir $\phi_2 = (a-d)^2 + (b-e)^2 + (c-f)^2$. De este modo plantearemos la función lagrangeana de la siguiente manera:

$$\begin{array}{rcl} \Phi: \quad \mathbb{R}^6 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (\overrightarrow{\mathrm{h}}, \overrightarrow{\lambda}) & \longmapsto & \mathcal{L}((a, b, c, d, e, f), \overrightarrow{\lambda}) = \phi_2 + \lambda_1 g_1 + \lambda_2 g_2 \end{array}$$

Ahora buscaremos $\nabla_{\overrightarrow{\mathrm{h}}}\mathcal{L}$, el cual se escribe -considerando las bases canónicas de $\mathbb{R}^6$- como sigue:

$$\nabla_{\vec{\mathrm{h}}}\mathcal{L} = \begin{pmatrix} 2(a-d) \\ 2(b-e) \\ 2(c-f) \\ -2(a-d) \\ -2(b-e) \\ -2(c-f) \end{pmatrix} + \begin{pmatrix} -2\lambda_1(a-2) \\ -0{,}5\lambda_1(b-3) \\ \lambda_1 \\ \lambda_2 \\ \lambda_2 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

Consecuentemente podemos ver, de lo anterior, que $\lambda_1 = -\lambda_2$, de este modo, reemplazando convenientemente en el resto de las ecuaciones obtenemos como solución el punto $(a, b, c) = (\frac{3}{2}, 1, \frac{25}{4})$

Aplicación 6.3 *En este apartado culminaremos por utilizar una aplicación del Teorema de Karush Kuhn Tucker. Consideremos la función*

$$\begin{array}{rccl} \mathcal{V}: & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & (x, y, z) & \longmapsto & \mathcal{V}(x, y, z) = xyz \end{array}$$

y la función $\Phi(x, y, z) = x^2 + y^2 + z^2 - r^2 > 0$ con $r > 0$. De este modo nos enfocaremos sobre el problema:

$$\left.\begin{array}{ll} & \text{mín}\,\mathcal{V}(x, y, z) \\ \textit{sujeto a} & \\ & \Phi(x, y, z) = 0 \end{array}\right\} (*)$$

Solución 6.3 *Construyamos la función lagrangeana para este problema, así*

$$\mathcal{L} = \lambda_0 V + \lambda_1 \Phi$$

.

Estamos seguros de que el problema (*) posee solución, pues $\mathcal{V}(x, y, z)$ es, sin lugar a dudas, continua y $\{(x, y, z) : \Phi(x, y, z) = 0\}$ es un compacto en $\mathbb{R}^3$ y sabemos que toda función continua sobre un conjunto compacto -de cerrado y acotado- alcanza sus extremos sobre dicho conjunto.

De esta forma, para utilizar Karush Kuhn Tucker, debemos calcular la derivada del lagrangeano, para esto procedamos a calcular las derivadas parciales de $\mathcal{V}$, obteniendo como resultado:

$V_x(x, y, z) = yz$
$V_y(x, y, z) = xz$
$V_z(x, y, z) = xy$

Notamos que las derivadas parciales de ésta son continuas, por lo que es evidente que $\nabla\mathcal{V}$ existe. Por otra parte, siguiendo igual razonamiento para la función Φ concluimos que existe $\nabla\Phi$.

De lo anterior se tiene que la derivada del lagrangeano existe, por lo que nuestro problema de optimización se reduce a:

$$\nabla_{(x,y,z)}\mathcal{L}((x, y, z), \lambda_0, \lambda_1) = \lambda_0 \nabla_{(x,y,z)}\mathcal{V} + \lambda_1 \nabla_{(x,y,z)}\Phi = 0$$

Sea (m, g, z) el punto que satiface el problema (*). Antes de continuar observemos que si $\lambda_0 = 0$ tendríamos que $\lambda_1 \neq 0$ y así llegaríamos a que $\lambda_1 \nabla_{(x,y,z)}\Phi = 0$, lo que a su vez *produciría que* $\nabla_{(x,y,z)}\Phi = 0$*, lo que no es posible, pues en caso de que esto fuese cierto tendríamos que :*

$\begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ *con lo que se obtiene que* $\Phi(0,0,0) = -r^2$.

Consecuentemente, $\lambda_0 > 0$*, por lo que consideremos* $\lambda_0 = 1$*, con lo que tendremos*

$$\nabla_{(x,y,z)}\mathcal{V}(m,g,z) + \lambda_1 \nabla_{(x,y,z)}\Phi = \overrightarrow{\mathbf{0}}$$

que es equivalente a $(0,0,mg) + \lambda_1(2x,2y,2z) = \overrightarrow{\mathbf{0}}$*, lo que a su vez implica que*

$$\left.\begin{array}{cc} x = 0 & y = 0 \\ -\lambda_1 2z = mg & \\ & \end{array}\right\} (\diamondsuit)$$

Además, no se debe olvidar que λ_1 *es distinto de cero (condición de las reglas de multiplicadores), asimismo recordemos que* $x^2 + y^2 + z^2 = r^2$*. Utilizando* $(\diamondsuit)$*, tendremos que* $z = \pm r$*, con lo que -realizando los cálculo pertinentes-* $\lambda_1 = \pm\frac{mg}{2r}$.

Capítulo 7

Conclusiones

Luego de un detallado tratamiento de las temáticas relacionadas a las condiciones de Karush Kuhn Tucker, tales como algunas nociones de cálculo en varias variables y los importantes teoremas de separación de Hahn Banach.

Una de las nociones importantes que se debe considerar es la convexidad de conjuntos. El teorema de Hahn Banach nos entrega las condiciones necesarias para que dos conjuntos convexos sean separables, vale decir, existirá un hiperplano que los dejará contenidos en dos semiespacios disjuntos, valiéndose de que los dos conjuntos anteriores sean disjuntos, condición dada por medio de la expresión $\overrightarrow{\mathbf{0}} \notin C_1 - C_2$.

En cuanto a los teoremas mencionados anteriormente concluimos que su importancia radica en que nos permiten encontrar los λ necesarios para las diversas reglas de multiplicadores, asegurándonos no sólo su existencia sino que también las propiedades que estos multiplicadores deben cumplir, a saber:

$$\lambda_i \geq 0 \quad i = 0, 1, ..., p$$
$$\lambda_i f_i(\overrightarrow{\mathbf{x_c}}) = 0, \quad \text{donde} \quad i = 1, ..., p,$$

a lo que se suma la primera condición de cada regla de multiplicadores. Todo lo anterior viene asegurado por la definición de producto interno.

Por otra parte cobra vital importancia en nuestro trabajo las diferentes reglas de multiplicadores (Convexa, John y Carathéodory-John), las cuales consisten principalmente en entregar características acerca de los multiplicadores que forman parte de la función lagrangeana, de algunos términos de esta función o bien de las propiedades del gradiente del lagrangeano, como también de los multiplicadores.

Ahora nos referiremos al objetivo principal de nuestro seminario, las condiciones de Karush Kuhn Tucker. En general la importancia de la función lagrangeana viene dada por la razón de que reduce un problema de optimización sujeto a restricciones (tanto de igualdad como de desigualdad, o ambas a la vez) a un problema de optimización no condicionado a restricciones, más ampliamente la función lagrangeana puede ser vista como una función que realiza el producto interno entre un vector de $\mathbb{R}^{p+q+1}$, constituido por las función objetivo y por las funciones restricciones, con un vector del mismo espacio pero esta vez formado por los multiplicadores ($\lambda_i \quad i = 0, 1, ..., p+q$). Además de esta perspectiva puede hacerse notar que el lagrangeano es una función definida desde el producto cartesiano $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{p+q}$ -donde $\mathbb{R}^n$ es el conjunto de definición de la función objetivo y las funciones restricción, $\mathbb{R}$ es el conjunto donde encontramos a λ_0 y $\mathbb{R}^{p+q}$ es el conjunto definición del resto de los λ- a los reales. El teorema de Karush Kuhn Tucker nos permite establecer la presencia del multiplicador de la función objetivo de tal modo que éste sea estrictamente positivo.

Las aplicaciones de este teorema se remiten a diferentes áreas del quehacer científico, entre las que podemos mencionar la teoría de control, medicina (problemas de bioequivalencia), estadística aplicada (estimadores de regresión, curvas de ajuste, métodos de mínimos cuadrados), física (mecánica de fluidos), economía, estudio de poblaciones, etc.

Se debe hacer notar que las condiciones de Karush Kuhn Tucker no son universalmente aplicables a todas las funciones, por ejemplo

$$\text{mín } f_0(x_1, x_2) = x_1$$

sujeto a

$$f_1(x_1, x_2) = x_1^3 - x_2^2 = 0$$

$$U \equiv \mathbb{R}^2,$$

considerando a $\lambda_0 = 1$.

Por lo que aún restan grandes problemáticas por resolver mediante la utilización la técnica de multiplicadores, consiguiendo una estrategia más general para aplicarla a un espectro más amplio de problemas.

Bibliografía

[1] . Jahn, *Introduction to the Theory of Nonlinear Optimization*. Ed. Springer, Tercera Edición, Berlín, 2007

[2] . Grossman, *Álgebra Lineal*. Ed. M^{c}Graw-Hill , Quinta Edición, México, 1996.

[3] . Daridov, *Constrained Estimation and the Theorem of Kuhn-Tucker*, Journal of Applied Mathematics and Decision Sciences, 2005.

[4] . Rojas, A. Moretti, V. Antunes de Oliveira (†) ,*Introducción a Programación Matemática*, Departamento de Matemática Aplicada.

Capítulo 8
Currículum autores

CURRÍCULUM VITAE JUAN SOTO FIGUEROA

1. ANTECEDENTES PERSONALES.
   - Nombre : Juan Ernesto Soto Figueroa.
   - Fecha de nacimiento : 12 de agosto de 1985
   - R.U.T: 16.220.121-2
   - Dirección: Población Santa Elvira calle Luís Vicentini ♯120
   - Teléfono: 09 - 82122016
   - Nacionalidad : Chileno
   - Título: "Profesor en Educación Matemática"(en ejecución)
   - Grado académico: Licenciado en Educación.
2. Antecedentes Laborales:
   - Agosto de 2006 : Trabaja para la Universidad de Concepción como Examinador de pruebas para el campeonato de matemáticas.
   - Septiembre de 2006: Realiza funciones para la universidad de Concepción Como Examinador en las pruebas de eliminación para el Campeonato Nacional de Matemáticas
   - Noviembre de 2006: Presta servicios al Ministerio de Educación como Examinador SIMCE.
   - Abril 2007 - Agosto 2007 : Presta servicios a la Universidad de Concepción como Examinador de pruebas para el Campeonato de Matemáticas.
   - Agosto de 2007 : Realiza funciones para la Universidad de Concepción como Examinador en las pruebas de eliminación para el Campeonato Nacional de Matemáticas en la ciudad de Chillán.
   - Mayo de 2008 : Realiza funciones para la Universidad de Concepción como Examinador en las pruebas de eliminación para el Campeonato Nacional de Matemáticas en la ciudad de Chillán.

- Abril de 2007 a Agosto de 2007: Trabaja en la Universidad del Bío-Bío desempeñando la ayudantía de Estructuras Algebraicas en la Carrera de pedagogía en Edicación Matemática.
- Septiembre de 2007 a Diciembre de 2007: En la Universidad del Bío-Bío realiza la ayudantía de Álgebra Lineal en la carrera de Ingeniería en Alimentos.
- Septiembre de 2007 a Diciembre de 2007: Realiza la ayudantía de Ecuaciones Diferenciales en la carrera de Ingeniería en Alimentos en la Universidad de Bío-Bío
- Septiembre de 2007 a Diciembre de 2007: Esta a cargo de la ayudantía de Álgebra en la carrera de Ingeniería Civil en Informática en la Universidad del Bío-Bío.
- Enero 2008: Presta servicios a la Universidad del Bío Bío como funcionario de registro académico en el proceso de matrícula de estudiantes, proceso de admisión 2008.
- Abril de 2008 a Agosto de 2008: Realiza la ayudantía de Estructuras algebraicas en la Carrera de Pedagogía en Educación Matemática en la Universidad del Bío-Bío.
- Septiembre 2008 a la fecha: Realiza ayudantía de Álgebra en la Universidad del Bío-Bío a la carrera de Pedagogía en Educación Matemática. Octubre 2008: Presta servicios a la Corporación Educacional Colegio Concepción-Ñuble como docente de aula.
- Noviembre 2008: Presta servicios al Ministerio de Educación como Examinador SIMCE. Diciembre 2008: Presta servicios al Departamento de Evaluación, Medición y Registro (DEMRE) como examinador PSU.

3. ANTECEDENTES ACADÉMICOS.
   - Enseñanza Básica:
     - 1992 - 1994 : Ingresa al primer año de Educación General Básica en la Escuela D - 203 República de los Estados Unidos de América, Chillán.
     - 1995 - 1999 : Ingresa al cuarto año de Educación General Básica a la Escuela E - 235 República Árabe Unida, Chillán.
   - Enseñanza Media:
     - 2000 - 2003 : Ingresa al primer año de Enseñanza media al Liceo A - 7 Narciso Tondreau, Chillán.
   - Enseñanza Superior:
     - 2004 - A la fecha : Ingresa a la carrera de Pedagogía en Educación Matemática en la Universidad del Bío - Bío, Chillán.
4. Otros Cursos:
   - Julio de 2006 : Participa en el Tercer Taller de Capacitación "Métodos de Enseñanza de la Matemáticarealizado en el Colegio Alemán, Concepción.
   - Octubre de 2007 : Participa en XXXIV Semana de la Matemática organizada por la Universidad Católica de Valparaíso.

CURRÍCULUM VITAE VALESKA MENA REYES

1. ANTECEDENTES PERSONALES.

   - Nombre : Valeska Nicol Mena Reyes.
   - Fecha de nacimiento : 23 de agosto de 1985
   - R.U.T : 16.217.142-9
   - Dirección : Independencia Nº 717
   - Teléfono : 042-327331 /82490484
   - Nacionalidad : Chilena
   - Grado Académico : Licenciado en Educación
   - Título : "Profesor en Educación Matemática" (en Proceso de Titulación)

2. ANTECEDENTES ACADÉMICOS.

   - Enseñanza Básica:
     - 1992 – 1999 : Cursa desde Primer a Octavo año de Enseñanza Básica en el Colegio de la Purísima Concepción de Chillán.
   - Enseñanza Media:
     - 2000 – 2003 : Cursa desde Primer a Cuarto año de Enseñanza Media en el Colegio de la Purísima Concepción de Chillán.
   - Enseñanza Superior:
     - 2004 – A la fecha : Ingresa a la carrera de Pedagogía en Educación Matemática en la Universidad del Bío – Bío, Chillán.

3. Otros Cursos:

   - Julio de 2006 : Participa en el Tercer Taller de Capacitación "Métodos de Enseñanza de la Matemática", Colegio Alemán de Concepción.
   - Octubre de 2007 : Participa en el Primer Encuentro del Aprendizaje Significativo en la Ciencia y Matemática, organizado por la Universidad Adventista de Chile, sede Chillán.
   - Octubre de 2007 : Participa en XXXIV Semana de la Matemática organizada por la Universidad Católica de Valparaíso

4. ANTECEDENTES LABORALES.

   - Agosto de 2006 : Se desempeña como Examinadora de pruebas para el campeonato de matemáticas desarrollado por la Universidad de Concepción, Chillán.
   - Septiembre de 2006 : Realiza funciones de Examinadora en las pruebas de eliminación para el Campeonato Nacional de Matemáticas, desarrollado en Chillán

- Noviembre de 2006 : Se desempeña como Examinadora SIMCE, desarrollado a nivel nacional.
- Abril 2007 - Agosto 2007 : Presta servicios a la Universidad de Concepción como Examinadora de pruebas para el Campeonato de Matemáticas.
- Agosto de 2007 : Realiza funciones para la Universidad de Concepción como Examinadora en las pruebas de eliminación para el Campeonato Nacional de Matemáticas en la ciudad de Chillán.
- Marzo a Septiembre 2008 : Realiza práctica profesional en el Colegio Chillán de Chillán
- Octubre 2008: Presta servicios a la Corporación Educacional Colegio Concepción-Ñuble como docente de aula.
- Noviembre 2008: Presta servicios al Ministerio de Educación como Examinador SIMCE.
- Diciembre 2008: Presta servicios al Departamento de Evaluación, Medición y Registro (DEMRE) como examinador PSU.

Capítulo 9

Agradecimientos

"Detrás de cada línea de llegada, hay una de partida. Detrás de cada logro, hay otro desafío. Mientras estés vivo, siéntete vivo. Si extrañas lo que hacías vuelve a hacerlo. No vivas de fotos amarillas... Sigue aunque todos esperen que abandones. No dejes que se oxide el hierro que hay en ti... "(Madre Teresa de Calcuta)

Cuando un ciclo comienza su fin, son demasiados los sentimientos que surgen y también con ello muchas las palabras que quedan sin decir. Sin embargo, creo que mucho de lo vivido hasta llegar al día de hoy es fruto de lo vivido anteriormente. Por ello, agradezco a Dios por cada momento vivido, pues ante cada momento - bueno o malo - su amor pudo más que todo.

Gracias a mi familia, por brindarme siempre el apoyo necesario, sabiendo lo difícil de la misión de ser profesor. A ellos, además, por ser mi soporte desde mi infancia y por enseñarme que en nuestros caminos todo se puede cumplir, porque fueron comprensivos en cada instante y porque permitieron que este sueño fuese realidad. A mi padre, por su vocación y sus sueños; a mi madre, por su cariño y su apoyo; a mi abuelita, mis hermanos y a mi sobrina, que con su sonrisa ayudó cuando el camino era más difícil.

A mis profesores, que siempre estuvieron apoyando nuestro seminario, como a aquellos que acompañaron nuestro camino. A nuestro profesor guía Roberto Cabrales, por su paciencia y por sus palabras; al profesor Marko Rojas, por permitirnos aprender de él; al profesor Luis Friz por ser un profesor preocupado por nosotros y por su gran capacidad de diálogo; al profesor Anibal Coronel e Ivo Basso, por estar siempre en contacto con nosotros y a la profesora Nelly Lagos, por su trato amable y por el apoyo en momentos complejos.

Finalmente, agradezco a mi compañero de tesis, mi amigo, mi cómplice y mi novio, por todos los momentos buenos y malos vividos durante este proceso. Por aquellos días en los cuales las fuerzas flaqueaban y por aquellos en los cuales cansados por nuestras prácticas seguíamos trabajando. Gracias por hacerme creer que era posible cumplir este sueño y por ser tú. Juan Ernesto, siempre apoyaré tus sueños, que también son nuestros. . .

Gracias a quienes hicieron posible esta historia

Valeska Mena Reyes

Es un largo camino y una gran parte de mi vida la cobijan estas aulas, estos prados y las almas de las personas con las que compartí y de las que aprendí mucho más de lo que se espera encontrar en un libro. En este lugar, el que me ha cambiado sin retorno y me ha hecho conocer lo mejor de mí, me ha abierto nuevos horizontes y me permitió creer en los sueños.

De este hogar llevo como propiedades inalienables las riquezas de las experiencias, lo valioso de un apoyo y lo importarte de tener fe en lo que se cree y arriesgarse a vivir la aventura vivir, pues la universidad es mucho más que conocimientos encerrados en algún antiguo libro, es un punto de encuento donde convergen muchas historias de vida y en el que lo más valioso que uno acuña es lo aprendido gracias a las muchas personas que nos acompañaron en el proceso de crecer, de atrevernos a ser más y mejores, a decir sí a un gran sueño. El diploma que nos reconoce como profesionales no alcanza a manifestar lo que de verdad fue esta experiencia de vida y nosotros aun lo logramos dimensionar lo distinto que somos ahora.

No obstante, a esta casa no hemos llegados sólos, al menos mi estadía se la debo a muchas personas que si no hubieran tenido la fe y la fuerza de creer en mí esta historia sería muy distinta. En este instante recuerdo a una profesora que dio el primer paso, el primer acto de convicción que me permitió tener la confianza de que podría lograrlo y aunque las circunstancia nos mantienen lejos, gracias profesora Elizabeth Soto, por estar presente en uno de los momentos en los que mi vida comenzó a tomar un rumbo distinto.

Son muchas las personas que me han ayudado a crecer, sin embargo, quisiera agradecer a los profesores que me comenzaron a guiar en esta senda; Sra. Yuri Haraguchi, Sr. Luis Fritz, Sr. Anibal Coronel, Sr. Ivo Basso, gracias por estar ahí y por mostrarnos parte del camino a seguir. No puedo olvidar a nuestro guía de tesis Sr. Roberto Carlos Cabrales, que a pesar de lo difícil que ha sido este camino y la paciencia que hemos demandado de su parte siempre estuvo allí para apoyarnos, asimismo no podemos dejar fuera de nuestro pensamiento al Sr. Marcos Rojas, que aunque poco nos hemos conocido a sido un gran aporte para nuestra formación y crecimiento en este arte maravilloso que es la matemática. A todo ustedes profesores muchas gracias por darnos las herramientas necesarias para emprender el vuelo.

Como anteriormente he señalado, en este camino son muchas las personas a las que encontramos, pero sólo unas pocas son las que se quedan con nosotros y son menos las que permanecen en nosotros, en nuestra alma; Valeska Mena gracias por aventurarte y tomar mi mano, por mostrame un sueño que no me atrevía a vivir, este tiempo juntos a sido el recuerdo más hermoso que guardo en mi ser, gracias por quedarte en mi corazón y por ayudarme a continuar soñando con un futuro sin fronteras, eres lo que no encontré y lo que jamás encontraré en ningún libro, eres, sin lugar a dudas, lo más valioso que me llevo de aquí.

En este punto me imposible evadir el profundo agradecimiento a mis seres más queridos, mi familia, los que sin importar el sacrificio me ayudó a cumplir uno de mis sueños. Y en este lugar deseo realizar un especial reconociemto a mi madre, Olga Figueroa, que siempre ha luchado por llevarnos mucho más lejos, que siempre creyó en mi y que a pesar de todos los problemas me impulsó a seguir sin estimar en lo que ella debiera entregar. Mamá gracias por estar ahí cuando más te he necesitado, gracias por tener la convicción de que los sueños se pueden alcanzar y gracias por tu incondicional entrega.

Finalmente hay una persona que a su modo me demostró su amor y que dio más de lo que a veces podía dar y sé que te hubiera encantado estar a mi lado en este momento celebrando con tus lágrimas, tal como lo hiciste el día en que supiste que de mi admisión en la universidad, sé que te hubiera fascinado estar conmigo por todo lo que esto significa para ti, gracias Nana, y estés donde estés sé que estarás conmigo. GRACIAS.

Juan Soto Figueroa