UNIVERSIDAD CENTROCCIDENTAL
“LISANDRO ALVARADO”

Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

“Dualidad sin Condiciones de Calificación en Optimización no Diferenciable”

Trabajo Especial de Grado presentado por

Br. Luis A. Sánchez H.

como requisito final
para obtener el título de Licenciado
en Ciencias Matemáticas
Área de Conocimiento: **Optimización Multiobjetivo**
Tutor: MsC. Clavel Quintana

Barquisimeto, Venezuela. Julio de 2012

Universidad Centroccidental
"Lisandro Alvarado"
Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

ACTA
TRABAJO ESPECIAL DE GRADO

Los suscritos miembros del Jurado designados por el Jefe del Departamento de Matemáticas del Decanato de Ciencias y Tecnología de la Universidad Centroccidental "Lisandro Alvarado", para examinar y dictar el veredicto sobre el Trabajo Especial de Grado titulado:

"DUALIDAD SIN CONDICIONES DE CALIFICACIÓN EN OPTIMIZACIÓN NO DIFERENCIABLE"

Presentado por el ciudadano BR. LUIS A. SÁNCHEZ H. titular de la Cédula de Identidad Nï¿½ V-10.456.175. Con el propï¿½sito de cumplir con el requisito académico final para el otorgamiento del título de Licenciado en Ciencias Matemáticas.

Luego de realizada la Defensa y en los términos que imponen los Lineamientos para el Trabajo Especial de Grado de la Licenciatura en Ciencias Matemáticas, se procedió a discutirlo con el interesado habiéndose emitido el veredicto que a continuación se expresa:

[1] ______________________

Con una calificación de ______________________ puntos.

En fe de lo expuesto firmamos la presente Acta en la Ciudad de Barquisimeto a los _____ días del mes de ______________________ de __________.

TUTOR	FIRMA
JURADO	FIRMA
JURADO	FIRMA

OBSERVACIONES:

---

[1] Aprobado ó Reprobado

El presente trabajo es una prueba de "la perseverancia de quien quiso llegar a ser lo que alguna vez se propuso", del Apoyo y el Amor que me brindan mis padres y hermanos , del estimulo de mis amigos, mis Ahijadas, de mi tutora la Prof. Clavel y del aprecio de la Familia Campos Márquez.

AGRADECIMIENTOS

A la fuerza motivadora de que todo lo puede: DIOS, por haberme guiado a través de las sendas de este mundo tan complejo y maravilloso que es LA MATEMÁTICA, por darme el conocimiento, la fortaleza y las herramientas necesarias para seguir adelante y vencer el desánimo que en muchas situaciones me embargaban.

Ser hijo de Alicia Herrera y José Sánchez es mi mayor premio en la vida; gracias a su educación, su apoyo, la inculcación a la perseverancia y los 16 hermanos que me dieron pues son los motivos por el cual jamás me rendí. A ti hermano Lisandro que aunque no pudiste ver físicamente la culminación de uno más de mis logros, siempre formaras parte de ellos.

"Hay algo más por construir en tu vida"; fueron las palabras que me motivaron de mis ahijadas que no conformándose con un Padrino T.S.U. querían un Padrino Licenciado. Gracias Marianny, Biyesi, Aliyer, Albany, Naidarly y Mariángel porque con su amor y creencia en Mi hicieron que fuese posible. Así como el amor y aprecio que en Mi depositara Marisol y su Familia, al igual que la Sra. Carmen Carrasco; pues me hicieron sentir parte de su familia por tener la mía tan lejos.

La amistad, el cual es fuente renovadora de muchas grandezas, la pude conseguir de mano de varios de mis profesores y compañeros de clases; compañeros que dejaron de serlo para convertirse en mis buenos amigos: Eleiny Acosta, Ernesto López, Nelson Campos, José Mogollón, Elvis Aponte.

Y eternamente agradecido a quien fortaleció mis conocimientos, vio potencial en Mí para seguir en esta carrera y me dio la oportunidad de trabajar con ella reforzando mis ideas con su conocimiento: A mi tutora y amiga la Prof. Clavel Quintana.

Universidad Centroccidental "Lisandro Alvarado"

"DUALIDAD SIN CONDICIONES DE CALIFICACIÓN EN OPTIMIZACIÓN NO DIFERENCIABLE"

Br. Luis A. Sánchez H.

Tutor: MsC. Clavel Quintana

RESUMEN

Nuestro trabajo está enfocado en la propuesta hecha por la Dra. Soghra Nobaskstian (10) que le permitió dar respuesta a un problema de optimización multiobjetivo no diferenciable con restricciones de desigualdad sin condiciones de calificación. Con el fin de obtener los resultados principales, presentamos las definiciones de las funciones convexas generalizadas sobre la base de la derivada direccional generalizada. Bajo la supuesta convexidad generalizada anterior, las condiciones necesarias y suficientes para la optimalidad, se dan sin la necesidad de las condiciones de calificación. Luego, se formula el problema dual correspondiente al problema primal, y algunos resultados de la dualidad que se obtienen sin condiciones de calificación.

La estructura del presente trabajo se realizó siguiendo una revisión detallada de los conceptos y caracterizaciones del Análisis Convexo, el Análisis Matemático y la Optimización Multiobjetivo el cual nos permitieron obtener una base teórica necesaria para el análisis de los aspectos fundamentales de la presente investigación.

Se Desarrollaron las ideas establecidas por S. Nobaskstian siguiendo una serie de definiciones y resultados en el Espacio Vectorial $\mathbb{R}^n$, así como el estudio de aspectos fundamentales de la convexidad, diferenciablidad, subdiferenciablidad, las restricciones de calificación y demás teorías necesarias presentes en la investigación de operaciones.

La teoría Multobjetivo ha ido revolucionando el mundo de la ingeniería y demás ciencias aplicadas, es por ende que el matemático busca dar respuesta a las situaciones que puedan presentarseles a la hora de dar una conclusión en base a los resultados obtenidos. Por tal motivo, mientras mayor sea una base teórica que refuerce los cálculos realizados mayor será la seguridad con que el ingeniero pueda emitir una conclusión.

Palabras claves: condiciones de calificación, optimalidad no suave, convexidad generalizada, cuasiconvexidad generalizada, derivada direccional generalizada, subgradiente generalizado.

ÍNDICE

Índice de Figuras

Introducción

Un problema multiobjetivo es un problema en el que dos o más funciones objetivo son reducidas al mínimo en un conjunto limitado implícitamente factible. En tal problema de condiciones de optimalidad y los resultados de la dualidad, a menudo se enfrentan con la calificación de las restricciones. Una calificación de las restricciones supone una cierta regularidad de las funciones de restricción cerca de la solución óptima. Las restricciones de calificación son a menudo supuestas pero no siempre las tienen. Algunas de las calificaciones mas conocidas son la calificación de Kuhn-Tucker, calificaciones de Slater (5), y calificación de Mangasarian-Fromovitz (9).

Weir y Mond (11) definieron los problemas duales para los problemas de programación con valores escalares dónde el requisito de convexidad usual para la dualidad estaba relajado y una calificación de las restricciones no fue necesitada. Ellos establecieron sus resultados usando diferenciabilidad. Entonces en (6) Weir y Mond extendieron sus resultados para los programas multiobjectivos. Ellos definieron los problemas duales para una programación multiobjectivo y un problema de programación fraccional multiobjectivo convexa/cóncava dónde las funciones involucradas no son asumidas por ser diferenciable y donde una calificación de las restricciones no se requiere. Egudo et al. (9) ha definido los problemas duales para problemas programación multiobjectivo diferenciable dónde una calificación de las restricciones no es asumida; su enfoque era diferente a la de (6) centrándose en la eficiencia en lugar de la eficiencia adecuada. Este enfoque tiene la ventaja de ser adecuado para definir duales para los problemas de programación no convexo.

La presente investigación esta basada en los resultados obtenidos por S. Nobakhtian (10), donde introdujo ciertas funciones convexas generalizadas y obtuvo las condiciones de optimalidad necesarias y suficientes para los problemas multiobjectivos no diferenciable sin la necesidad de restricciones de calificación. Además veremos como define los problemas duales que corresponden a los problemas primales y luego demostramos algunos resultados de dualidad sin la necesidad de restricciones de calificación y sin diferenciabilidad. También en su enfoque el requisito de convexidad usual para las funciones está relajado.

Se derivan las condiciones necesarias para problemas de optimización multiobjetivo con una función objetivo escalarizada bajo las mismas restricciones de calificación del problema no lineal monoobjetivo, etc.

Objetivo General

Estudiar el problema dual sin condiciones de calificación en optimización multiobjetivo de funciones no diferenciables en conjuntos convexos.

Objetivos Específicos

- Estudiar condición de calificación.
- Estudiar convexidad generalizada y cuasiconvexidad generalizada.
- Definir los problemas duales correspondientes a los problemas primales multiobjetivos en optimización no diferenciable.
- Demostrar algunos resultados de dualidad sin la necesidad de una condición de calificación y ni diferenciabilidad.

CAPÍTULO 1

PRELIMINARES

Los problemas de optimización tienen una larga historia dentro del campo de la investigación de operaciones, por lo que la matemática busca desarrollar cada día importantes base teórica que permita dar respuesta a cualquier situación que pueda presentarse. La intención de este capitulo es realizar una breve introducción a la terminología y caracterizaciones de la optimización para el desarrollo del presente trabajo.

§1.1. OPTIMIZACIÓN

En matemáticas se tiene que la optimización o programación matemática intenta dar respuesta a un tipo general de problemas donde se desea elegir el mejor entre un conjunto de elementos. En su forma más simple, el problema equivale a resolver una ecuación de este tipo:

$$\begin{aligned} max(min) \quad & f(x) \\ & x \in \Omega \subseteq \mathbb{R}^n \end{aligned}$$

donde $x = (x_1, x_2, ..., x_n)^t$ es un vector y representa variables de decisión, $f(x)$ es llamada función objetivo y representa o mide la calidad de las decisiones (usualmente números enteros o reales) y Ω es el conjunto de decisiones factibles o restricciones del problema.

Algunas veces es posible expresar el conjunto de restricciones (Ω) como solución de un sistema de igualdades o desigualdades:

$$h(x_1, x_2, ..., x_n) = 0$$

$$g((x_1, x_2, ..., x_n) \leqslant 0$$

En un problema de optimización se tratará de encontrar una solución que represente el valor óptimo para una función objetivo o un conjunto de funciones objetivos. En el caso más sencillo se tendrá un único objetivo, que estará representado por una función del tipo $f : M \to N$, donde $M \subset \mathbb{R}$ y $N \subset \mathbb{R}$. Tanto el dominio como la imagen de la función serán números reales (escalares), y el valor óptimo corresponderá a un mínimo o a un máximo.

Un problema de optimización trata entonces de tomar una decisión óptima para maximizar (ganancias, velocidad, eficiencia, etc.) o minimizar un criterio determinado (costos, tiempo, riesgo, error, etc). Las restricciones significan que no cualquier decisión es posible.

§1.2. CONVEXIDAD

El estudio de la convexidad es uno de los temas de gran relevancia que se utiliza en el análisis de problemas de programación matemática, la caracterización de sus solución óptima y en el desarrollo de programas de computación; pues la mayoría de lo que hasta ahora se conoce en optimización gira entorno a ello, incluyendo el desarrollo del presente trabajo de investigación. Es por ende que nos dimos a la tarea de revisar y entender las siguientes definiciones extraídas de (3), (5), (4) y (10):

§1.2.1. Conjuntos Convexos

Definición 1.2.1. (3): Un conjunto C es *Convexo* si, y sólo si, se cumple que:

$$\text{Para todo} \quad x^1, x^2 \in C, \quad \lambda x^1 + (1-\lambda)x^2 \in C, \quad \lambda \in [0,1]$$

es decir, que dados dos puntos cualquiera del conjunto, el segmento lineal cerrado que une los dos puntos está totalmente contenido en el conjunto.

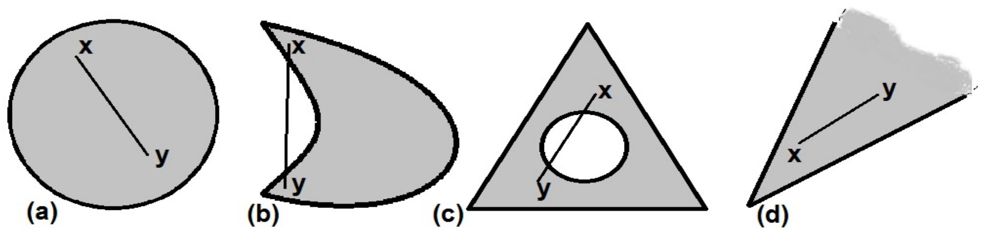

FIGURA 1.2.1: Conjuntos Convexos y No Convexos en $\mathbb{R}^2$

De la definición dada podemos deducir que en la Figura 1.2.1 se tiene que los conjuntos (a) y (d) son convexos y los no convexos son (b) y (c) en $\mathbb{R}^2$.

Definición 1.2.2. ((3), (4)): Sea C un conjunto arbitrario en $\mathbb{R}^n$. La Cápsula Convexa de C, denotado por $co(C)$, es la colección de todas las combinaciones convexas de C. En otras palabras, $x \in co(C)$ si y solamente si puede ser expresado como

$$x = \sum_{j=1}^{k} \lambda_j x^j \quad \text{con} \quad \sum_{j=1}^{k} \lambda_j = 1 \quad \text{para} \quad \lambda_j \geqslant 0 \quad \text{y} \quad j = 1, 2, .., k$$

donde k es un entero positivo y $x^1, x^2, ..., x^k \in C$.

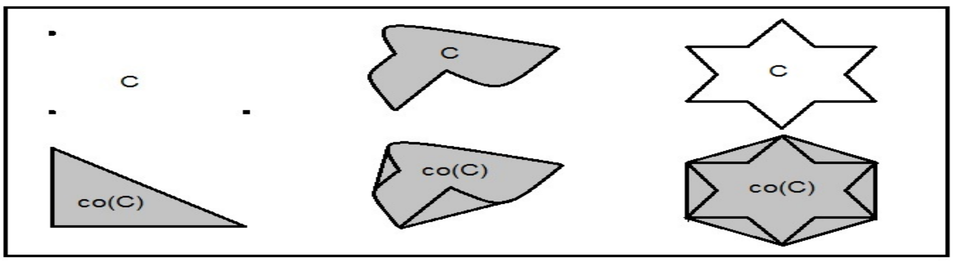

FIGURA 1.2.2: Cápsula convexa en $\mathbb{R}^2$

En tal sentido, podemos ver en la Figura 1.2.2 la representación de la cápsula convexa de tres conjuntos no vacío C dados de $\mathbb{R}^2$.

La siguiente definición nos ayudará en gran medida a comprender las caracterizaciones e ideas geométricas que deseamos desarrollar en la investigación debido a que las condiciones necesarias para el problema de optimización que plantearemos pueden ser expresados en términos del cono tangente, el cono normal y sus polares. De (3), (4) y (10) se tiene que:

Definición 1.2.3. : Sea C un conjunto no vacío en $\mathbb{R}^n$ y $\alpha, \lambda \in \mathbb{R}$, diremos que:

- C es un Cono con vértice cero, si $\lambda d \in C$ para todo $d \in C$ y para todo $\lambda > 0$.(ver Figura 1.2.3(c))
- El Cono Generado por C, denotado $Cono(C)$, es dado por $Cono(C) = \{p \in \mathbb{R}^n : \lambda p,\ \forall p \in C\ y\ \forall \lambda > 0\}$. (ver Figura 1.2.3(b))
- El Cono Polar de C, denotado C^*, es dado por $C^* = \{p \in \mathbb{R}^n : p^t x \geqslant 0, para\ todo\ x \in C\}$. Si C es vacío, C^* será interpretado como $\mathbb{R}^n$.(ver Figura 1.2.3(a))
- El Cono Normal de C en x^*, denotado $N_c(x^*)$ es dado por: $N_c(x^*) = \{d \in \mathbb{R}^n : \langle d, x - x^* \rangle \leqslant 0, para\ todo\ x \in C\}$. Ciertamente: $(C - x^*)^* = -N_c(x^*)$.(ver Figura 1.2.3(a))
- El Cono Tangente de C en $\bar{x}$ (con $\bar{x} \in cl(C)$),denotado por $T_C(\bar{x})$, es el conjunto de todas las direcciones d tal que $d = \lim\limits_{k \to \infty} \lambda_k (x^k - \bar{x})$, donde $\lambda_k > 0$, $x^k \in C$ para cada k, y $x^k \to \bar{x}$.(ver Figura 1.2.3(d))

De la definición de cono tangente, es claro que d pertenece al cono tangente si hay una sucesión factible $\{x^k\}$ convergiendo a x tal que las direcciones $x^k - x$ convergen a d.

En $\mathbb{R}^2$ podemos denotar el cono $C = \{d \in \mathbb{R}^2 : d_k \geqslant 0, k = 1, 2\} = \mathbb{R}^2_{\geqslant}$. Este es el Cono de elementos no negativos de las componentes de débil orden.(ver Figura 1.2.3(c))

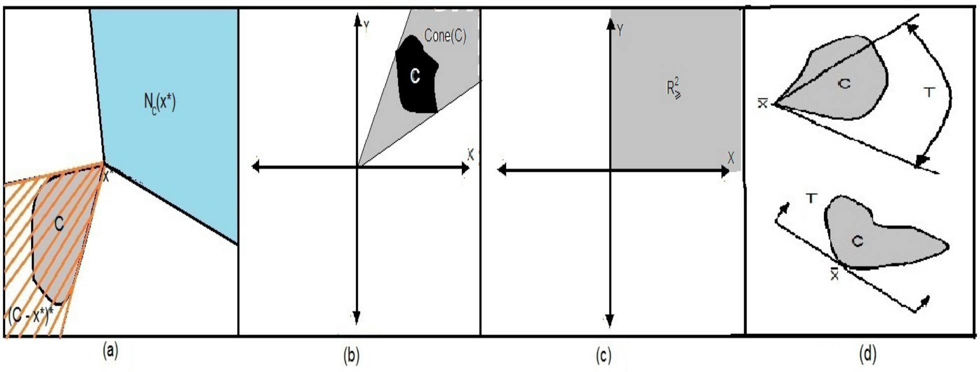

FIGURA 1.2.3: Tipos de conos en $\mathbb{R}^2$

§1.2.2. Funciones Convexas y Cóncavas

En esta sección definiremos dos tipos especiales de funciones: la función convexa y la función cóncava.

Definición 1.2.4. ((5) y (3)) : Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo no vacío; y sea la función $f : C \longrightarrow \mathbb{R}$. Entonces diremos que:

1. f es una función convexa si y solo si $f(\lambda x^1 + (1-\lambda)x^2) \leqslant \lambda f(x^1) + (1-\lambda)f(x^2)$ Para todo $\lambda \in [0,1]$ y Para todo $x^1, x^2 \in C$. (ver Figura 1.2.4 (a)).
2. f es una función estrictamente convexa si y solo si $f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda)f(x^2)$ para todo $\lambda \in (0,1)$ y para todo $x^1, x^2 \in C$, con $x^1 \neq x^2$.(ver Figura 1.2.4 (a)).
3. f es una función cóncava si y solo si $-f$ es convexa.(ver Figura 1.2.4 (b)).

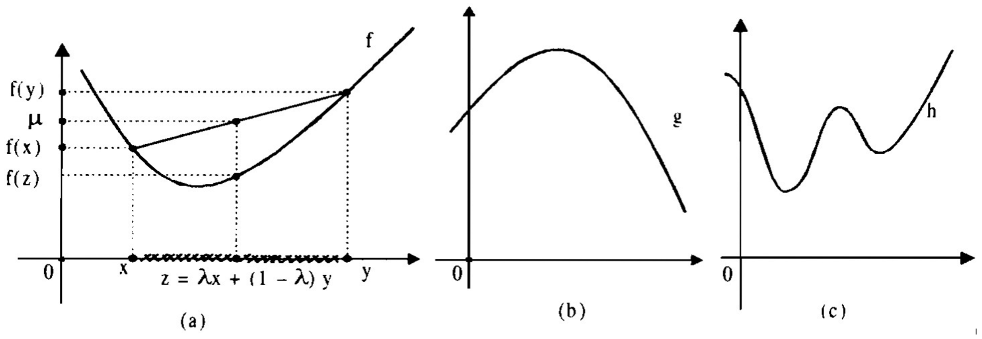

FIGURA 1.2.4: Función Convexa, Cóncava; y ni Convexa ni Cóncava en $\mathbb{R}^2$

No es difícil demostrar que las únicas funciones que son cóncavas y convexas simultáneamente son las funciones lineales en $\mathbb{R}^n$ con $n \in \mathbb{N}$.

La siguiente sección nos ayudará a entender el estudio de un caso particular de las derivadas direccionales bajo la condición de Lipschitz como lo es la derivada direccional generalizada de Clarke, el cual es utilizado para el caso de funciones no diferenciable, dado que es uno de los objetivos de esta investigación.

§1.3. DERIVADA DIRECCIONAL

Definición 1.3.1. ((2)): Dado un campo escalar $f : C \subset \mathbb{R}^n \to \mathbb{R}$, $a \in int(C)$, y sea $\vec{d} \neq \oslash$ un vector unitario fijo de $\mathbb{R}^n$, se llama Derivada Direccional de f a lo largo del vector $\vec{d}$ en **a** y se denota por $\frac{\partial f(a)}{\partial \vec{d}}$ al siguiente límite, si existe y es finito:

$$\frac{\partial f(a)}{\partial \vec{d}} := \lim_{t \to 0} \frac{f(a + t\vec{d}) - f(a)}{t}.$$

El cociente $\frac{f(a+t\vec{d})-f(a)}{t}$ representa el promedio de variación de f por unidad de distancia según la dirección del vector $\vec{d}$.

Debido a la existencia de funciones que poseen derivadas direccionales en un punto para cualquier dirección pero no son continuas en dicho punto, el concepto de derivada direccional no es el adecuado para generalizar el de derivada de funciones de una variable.

§1.3.1. La Condición de Lipschitz

Sea C un subconjunto de $\mathbb{R}^n$. Diremos que una función $f : C \to \mathbb{R}$ satisface la condición de Lipschitz (en C), siempre que, para algún K no-negativo, y para todo par de puntos $y^1, y^2 \in C$ satisface:

$$| f(y^1) - f(y^2) | \leqslant K \parallel y^1 - y^2 \parallel .$$

esto también es equivalente a la condición de Lipschitz de constante K.

Observación 1. Esta condición implica en particular que f es Uniformemente Continua en C.

Definición 1.3.2. Diremos que una función $f : \mathbb{R}^n \to \mathbb{R}$ es Localmente Lipschitziana con constante K en $a \in \mathbb{R}^n$ si existe un $r > 0$ tal que:

$$| f(y) - f(x) | \leqslant K \parallel y - x \parallel, \ \forall x, y \in B(a, r).$$

§1.3.2. Subdiferencial convexo

Definición 1.3.3. ((5), (3)):(Subgradiente de una Función). Sea $f : C \to \mathbb{R} \bigcup \{+ \propto\}$ una función convexa y $x^* \in C^*$ es un subgradiente de f en x si: $f(x) + \langle y - x, x^* \rangle \leqslant f(y), \forall y \in C$.(Ver Figura 1.3.1)

Denotaremos por $\partial f(x) \subset C^*$ el conjunto de todos los subgradientes de f en $x \in domf$ dado por $\partial f(x) = \{x^* \in C^* : f(x) + \langle y - x, x^* \rangle \leqslant f(y), \forall y \in C\}$, y el cual es llamado el subdiferencial de f en $x \in domf$.

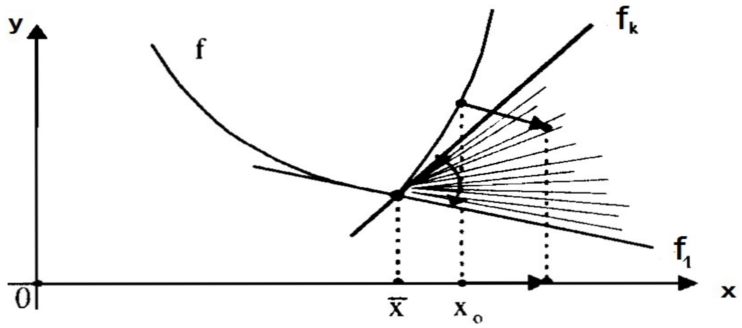

FIGURA 1.3.1: Función Convexa f y un punto $\bar{x}$ de no diferenciabilidad

§1.3.3. Derivada Direccional Generalizada Clarke

Definición 1.3.4. (2): Sean $C \subset \mathbb{R}^n$, $f : C \to \mathbb{R}$, $y \in C$ y $t \in \mathbb{R}^+$. La Derivada Direccional Generalizada Clarke de la función f localmente Lipschitz en x en dirección del vector $d \in C$ se define por:

$$f^c(x, d) = \limsup_{y \to x,\ t \downarrow 0} \frac{f(y + td) - f(y)}{t} \tag{1.1}$$

El subgradiente generalizado Clarke de una función f localmente Lipschitz en x se define por:

$$\partial_c f(x) := \{\xi \in R^n : f^c(x; d) \geqslant \langle \xi, d \rangle, \ Para\ todo\ d \in \mathbb{R}^n\} \tag{1.2}$$

En la definición nos podemos dar cuenta que no se supone la existencia del límite (sólo del límite superior). Esto implica solo el comportamiento de f arbitrariamente muy próximo a x, y se diferencia de las definiciones tradicionales de derivada direccional en que el punto base y del cociente de la diferencia varía.

Recordamos ahora el siguiente resultado de (2).

Lema 1.3.1. Sea f una función localmente Lipschitz, y $x \in domf$. Entonces para todo d en $\mathbb{R}^n$,

$$f^c(x,d) = max\{\langle \xi, d\rangle,\ Para\ todo\ \xi \in \partial_c f(x)\} \tag{1.3}$$

y $\partial_c f(x)$ es un conjunto no vacío, convexo, y compacto.

Prueba: Probemos que $f^c(x,d) = max\{\langle \xi, d\rangle,\ \xi \in \partial_c f(x)\}$ para todo d

(i) Por definición de $\partial f(x)$, tenemos que $f^c(x,d) \geqslant max\{\langle \xi, d\rangle,\ \xi \in \partial_c f(x)\}$ para todo d.

(ii) Probemos ahora que $f^c(x,d) \leqslant max\{\langle \xi, d\rangle,\ \xi \in \partial_c f(x)\}$ para todo d.
Sabemos que $d \to f^c(x,d)$ es convexa. Fijemos d arbitrariamente y escojamos $\xi \in \partial_c f(x,d)$ (en sentido de análisis conexo). Entonces:

$$\langle \xi, w-d\rangle \leqslant f^c(x,w) - f^c(x,d), \quad para\ todo\ w \in \mathbb{R}^n$$

substituyendo $w = \alpha\varpi$, donde $\varpi \in \mathbb{R}^n$ arbitrario y $\alpha > 0$, tenemos:

$$\langle \xi, \alpha\varpi - d\rangle \leqslant f^c(x,\alpha\varpi) - f^c(x,d), \quad para\ todo\ \alpha\varpi \in \mathbb{R}^n$$

Luego por propiedad de $\langle , \rangle$ y de prop.1.4.1 (2) se obtiene:

$$\alpha\langle \xi, \varpi - \frac{1}{\alpha}d\rangle \leqslant \alpha f^c(x,\varpi) - f^c(x,d), \quad para\ todo\ \alpha\varpi \in \mathbb{R}^n$$

así, dividiendo por α y tomando límite cuando $\alpha \to \infty$ tenemos:

$$\langle \xi, \varpi\rangle \leqslant f^c(x,\varpi), \quad para\ todo\ \varpi \in \mathbb{R}^n$$

por lo que, $\xi \in \partial f(x)$. Por otro lado, tomando $w = 0$, obtenemos directamente $\langle \xi, d\rangle \geqslant f^c(x,d)$; por tanto $f^c(x,d) \leqslant max\{\langle \xi, d\rangle,\ \xi \in \partial_c f(x)\}$ y así

$$\therefore\ De\ (i)\ y\ (ii):\ f^c(x,d) = max\{\langle \xi, d\rangle,\ \xi \in \partial_c f(x)\}$$

Proposicion 1.3.1. Un elemento $\bar{x} \in C$ es un mínimo global de f (es decir, $f(\bar{x}) \leqslant f(x), \forall x \in C$) si y solamente si $\oslash \in \partial f(\bar{x})$.

Prueba: Dado que $\bar{x} \in X$ es un mínimo global entonces $\bar{x}$ es un punto óptimo para f; así:

$$\begin{aligned} \bar{x} \in X \text{ es un mínimo global} \quad &\Leftrightarrow f(\bar{x}) \leqslant f(x),\ \forall x \in C \\ &\Leftrightarrow f(\bar{x}) + 0 \leqslant f(x),\ \forall x \in C \\ &\Leftrightarrow f(\bar{x}) + \langle x - \bar{x}, \oslash\rangle \leqslant f(x),\ \forall x \in C \\ &\Leftrightarrow \oslash \in \partial f(\bar{x}) \end{aligned}$$

Proposicion 1.3.2. Si la función convexa $f : C \to \mathbb{R} \bigcup \{+\infty\}$ es diferenciable en $x \in C$ entonces $\partial f(x) = \{\nabla f(x)\}$.

Prueba: Sabemos que $\partial f(x) = \{\nabla f(x)\} \equiv \{\nabla f(x)\} \subset \partial f(x) \wedge \partial f(x) \subset \{\nabla f(x)\}$

(i) Probemos que $\{\nabla f(x)\} \subset \partial f(x)$.
Es fácil mostrar que si f es diferenciable entonces $\nabla f(x) \in \partial f(x)$.
De hecho, una propiedad de las funciones convexas diferenciables es

$$f(x) + \langle y - x \ , \ \nabla f(x) \rangle \leqslant f(y), \ \forall y \in C \tag{1.4}$$

pues

$$\begin{aligned} \text{f convexa} \quad & \Rightarrow f(t.y + (1-t).x) \leqslant t.f(y) + (1-t).f(x) \\ & \Rightarrow f(x + t.(y-x)) \leqslant t.f(y) + f(x) - t.f(x) \\ & \Rightarrow t.f(x) - f(x) + f(x + t.(y-x)) \leqslant t.f(y) \\ & \Rightarrow f(x) + \tfrac{f(x+t.(y-x)) - f(x)}{t} \leqslant f(y) \qquad \text{(para t pequeño)} \end{aligned}$$

y tomando $t \to 0$ se obtiene la expresión (1.4).

(ii) Probemos ahora $\partial f(x) \subset \nabla f(x)$.

$$\begin{aligned} x^* \in \partial f(x) \quad & \Rightarrow f(x) + \langle x - y \ , \ x^* \rangle \leqslant f(y), \ \forall y \in C && (\text{def. } \partial f(x)) \\ & \Rightarrow f(x) + \langle x \ , \ x^* \rangle - \langle y \ , \ x^* \rangle \leqslant f(y), \ \forall y \in C && (\text{propiedad } \langle , \rangle) \\ & \Rightarrow f(x) + \langle x \ , \ x^* \rangle \leqslant f(y) + \langle y \ , \ x^* \rangle, \ \forall y \in C \\ & \Rightarrow g(x) := f(x) + \langle x \ , \ x^* \rangle \leqslant f(y) + \langle y \ , \ x^* \rangle = g(y), \ \forall y \in C \end{aligned}$$

es decir, la función diferenciable g se minimiza en x, por lo tanto

$$\nabla g(x) = \nabla f(x) - x^* = 0$$

$$\therefore \ De \ (i) \ y \ (ii): \ \partial f(x) = \{\nabla f(x)\}$$

Ejemplo 1.3.1. : Dada la siguiente gráfica en $\mathbb{R}^3$

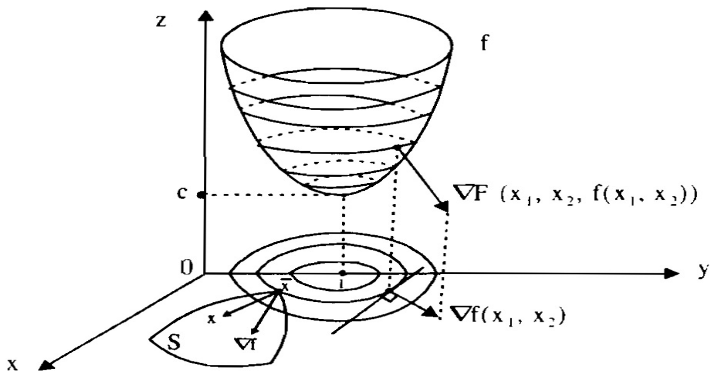

FIGURA 1.3.2: gráfica en que $\nabla f(x_1, x_2) = (0, 1) = \partial f(x_1, x_2)$

De la figura 1.3.2 se tiene que el gradiente $\nabla f(x_1, x_2)$ se anula en el punto $(0, 1)$ y en cualquier otro punto es diferente de cero. Por otro lado, vemos que ∇f indica la dirección de crecimiento de la función y es ortogonal a la tangente trazada a la curva de nivel en el punto de tangencia.

Definición 1.3.5. (10): Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función localmente Lipschitz. Entonces diremos que:

(i) f es Convexa Generalizada en x si para cualquier y, se cumple

$$f(y) - f(x) \geqslant \langle \xi \ , \ y - x \rangle, \quad \forall \xi \in \partial_c f(x) \tag{1.5}$$

(ii) f es Cuasiconvexa Generalizada en x si para cualquier y, se cumple

$$f(y) \leqslant f(x), \quad \langle \xi \ , \ y - x \rangle \leqslant 0, \quad \forall \xi \in \partial_c f(x) \tag{1.6}$$

(iii) f es Cuasiconvexa Estrictamente Generalizada en x si para cualquier y, se cumple

$$f(y) \leqslant f(x), \quad y \neq x, \quad \langle \xi \ , \ y - x \rangle < 0, \quad \forall \xi \in \partial_c f(x) \tag{1.7}$$

§1.4. Multiplicadores de Lagrange

A lo largo del tratamiento de los problemas mono-objetivos y multiobjetivo, es de vital importancia la definición de la llamada *función de Lagrange* asociada, que es una función que de alguna forma engloba todo el problema, al añadir a la función objetivo las restricciones, cada una de ellas acompañada por un multiplicador (multiplicador de Lagrange).

Considere el siguiente problema de programación no lineal:

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.a:} & g_i(x) \leqslant 0, \quad i \in P, \\ & h_j(x) = 0, \quad j \in Q, \\ & x \in C \end{array} \tag{SP}$$

Definición 1.4.1. Se llama Función de Lagrange asociada al problema (SP) a la función: $\mathcal{L} : D \times \mathbb{R}^n \to \mathbb{R}$ dada por:

$$\mathcal{L}(x, \lambda, \mu) = f_0(x) + \sum_{i \in I(x)} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x)$$

donde $I(x) = \{i \in P / g_i(x) = 0\}$ es el conjunto de las restricciones activas, $\lambda \in \mathbb{R}^p$ y $\mu \in \mathbb{R}^q$ se denominan Multiplicadores de Lagrange, también denominados multiplicadores de Kuhn-Tucker.

$\mathcal{L}(x, \lambda, \mu)$ proporciona una cota inferior de $f(x)$ para $\lambda \geqslant \oslash$ y μ

§1.5. Condición de Optimalidad y Dualidad

§1.5.1. Condiciones de Optimalidad Fritz John y Karush-Kuhn-Tucker

En matemáticas, Las condiciones de Karush-Kuhn-Tucker (también conocidas como las condiciones KKT o Kuhn-Tucker) son condiciones necesarias y suficientes para que la solución de una programación no lineal sea óptima. Es una generalización del método de los Multiplicadores de Lagrange.
Consideremos el problema siguiente de optimización no lineal:

$$\begin{array}{ll} \min & f(x) \\ \text{s.a:} & g_i(x) \leqslant 0, \quad i \in P, \\ & h_j(x) = 0, \quad j \in Q, \\ & x \in C \end{array} \tag{SP}$$

donde $f(x)$ es la función a ser minimizada, $g_i(x)$ son las restricciones de desigualdad y $h_j(x)$ son las restricciones de igualdad, además P y Q denotan el número de restricciones de desigualdad e igualdad, respectivamente.

§1.5.2. Condiciones Necesarias de KKT

(De (5)) Dado el problema (SP) y supongamos que f, g_i y h_j son continuamente diferenciables en el punto $x^* \in C$. Si x^* es un mínimo local, entonces existe constantes $\lambda > 0$, $\mu_i \geqslant 0$ $(i = 1, 2, ..., p)$ y ν_j $(j = 1, 2, ..., q)$ tales que:

- $\lambda + \sum_{i=1}^{p} \mu_i + \sum_{j=1}^{q} \mid \nu_j \mid > 0,$
- $\lambda \nabla f(x^*) + \sum_{i=1}^{p} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{q} \nu_j \nabla h_j(x^*) = 0,$
- $\mu_i g_i(x^*) = 0$, para todo $i = 1, 2, ..., p$

§1.5.3. Condiciones de regularidad (o restricciones de calificación)

En esta sección, veremos que en algunos casos no se verifican ciertas propiedades deseables debido a particularidades del conjunto eficiente. Es por ello que en ocasiones hará falta exigir ciertas condiciones a las restricciones del problema para evitar estos casos extremos. Estas condiciones reciben el nombre genérico de restricciones de calificaciones. En esta sección enunciaremos algunas de las más empleadas en la literatura; pero para el desarrollo de la presente investigación sólo nos interesaremos por la calificación de Kuhn-Tucker, calificaciones de Slater (5), y calificación de Mangasarian-Fromovitz (8).

En la condición necesaria anterior, el multiplicador dual λ puede ser igual a cero. Este caso se denomina degenerado o anormal. La condición necesaria no tiene en cuenta las propiedades de la función sino la geometría de las restricciones.

Existe una serie de condiciones de regularidad que aseguran que la solución no es degenerada (es decir $\lambda \neq 0$). Estas incluyen:

- La restricción de calificación de independencia lineal (CRIL): los gradientes de las restricciones activas de desigualdad y los gradientes de las restricciones de igualdad son linealmente independientes en x^*.
- La restricción de calificación de Mangasarian-Fromowitz (CRMF): los gradientes de las restricciones activas de desigualdad y los gradientes de las restricciones de igualdad para C abierto son linealmente independientes positivos en x^* y que $G_0 \bigcap H_0 \neq \emptyset$. Donde $G_0 = \{d : \nabla g_i(x^*)^t d < 0, \ \forall \ i \in I\}$, $I = \{i : g_i(x^*) = 0\}$ y $H_0 = \{d : \nabla h_j(x^*)^t d = 0, \ para \ j = 1, 2, ...q \ \}$.
- La restricción de calificación de rango constante (CRRC): para cada subconjunto de las restricciones activas de desigualdad y los gradientes de las restricciones de igualdad, el rango en el entorno de x^* es constante.

- La restricción de calificación de dependencia lineal constante positiva (DLCP): para cada subconjunto de restricciones activas de desigualdad y de gradientes de las restricciones de igualdad, si es linealmente dependiente positivo en x^* entonces es linealmente dependiente positivo en el entorno de x^*. ($\{v_1, v_2, ..., v_n\}$ es linealmente dependiente positivo si existe $\alpha_1 \geqslant 0, \alpha_2 \geqslant 0, ..., \alpha_n \geqslant 0$ no todos nulos tal que $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \emptyset$

- Condición de Slater: Sea C abierto, cada g_i con $i \in I$ es pseudoconvexa en x^*, cada g_i con $i \notin I$ es continua en x^*, cada h_j para $j = 1, 2, ..., q$ es cuasiconvexa, cuasiconcava, y continuamente diferenciable en x^*, y $\nabla h_j(x^*)$ para $j = 1, 2, ..., q$ son linealmente independientes. Además, existe un punto $x \in C$ tal que $g_i(x) \leqslant 0$ para todo $i \in I$ y $h_j(x) = 0$ para todo $j = 1, 2, ..., q$.

Puede verse que CRIL => CRMF => DLCP, CRIL => CRRC => DLCP, aunque CRMF no es equivalente a CRRC. En la práctica, se prefiere la restricción de Calificación más débiles ya que proporcionan condiciones de optimalidad más fuertes.

§1.5.4. Condiciones Suficientes de KKT

Dado el problema (SP) y supongamos que f, g_i sean funciones convexas, h_j sean funciones afín, y sea un punto $x^* \in C$. Si existen constantes $\mu_i \geqslant 0$ $(i = 1, 2, ..., p)$ y ν_j $(j = 1, 2, ..., q)$ tales que:

- $\nabla f(x^*) + \sum_{i=1}^{p} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{q} \nu_j \nabla h_j(x^*) = 0,$

- $\mu_i g_i(x^*) = 0$, para todo $i = 1, 2, ..., p$

entonces el punto $x^* \in C$ es un mínimo global.

§1.6. PUNTOS MINIMAX

El punto minimax de la función lagrangiana es otro concepto relacionado con la solución de un problema de optimización. Si bien su definición no le hace útil a la hora de la resolución directa del problema, sí constituye un paso intermedio muy importante en la obtención del problema dual, que estudiaremos más adelante. En esta sección definimos dicho punto y estudiamos su relación con otro concepto, el punto silla del lagrangiano.

Definición 1.6.1. Dado el problema (SP), diremos que $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{D}$ es un punto **minimax** de la función lagrangiana $\mathcal{L}(x, \lambda, \mu)$ en el conjunto $\mathbb{R}^n \times \mathbb{D}$ si

$$\mathcal{L}(x^0, \lambda^0, \mu^0) = \min_{\mathbb{R}^n} \max_{\mathbb{D}} \mathcal{L}(x, \lambda, \mu) = \max_{\mathbb{D}} \min_{\mathbb{R}^n} \mathcal{L}(x, \lambda, \mu)$$

Observación 2. La relación del punto minimax con la solución del problema de programación no lineal se obtiene de forma inmediata sin mas que tener en cuenta que:

$$\min_{\mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) = f(x) + \max_{\mathbb{R}^n} \lambda^t g(x) + \max_{\mathbb{R}^n} \mu^t h(x)$$

Pasamos ahora a dar los teoremas que relacionan los conceptos de punto de silla de $\mathcal{L}$ y punto **minimax**. Veremos que dicha relación es casi una equivalencia, en el sentido de que todo punto **minimax** es punto de silla, y todo punto de silla es un punto **minimax** considerado sobre conjuntos mas restringidos.

Teorema 1.6.1. Si existe un punto **minimax** $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{D}$ de la función de Lagrange, es decir, si

$$\mathcal{L}(x^0, \lambda^0, \mu^0) = \min_{\mathbb{R}^n} \max_{\mathbb{D}} \mathcal{L}(x, \lambda, \mu) = \max_{\mathbb{D}} \min_{\mathbb{R}^n} \mathcal{L}(x, \lambda), \mu$$

entonces $\mathcal{L}(x, \lambda, \mu)$ posee puntos de sillas en $\mathbb{R}^n \times \mathbb{D}$, y (x^0, λ^0, μ^0) es uno de ellos.

Teorema 1.6.2. Si $\mathcal{L}(x, \lambda)$ posee un punto silla (x^0, λ^0) con respecto a $\mathbb{D} \times \mathbb{R}^{m+}$, entonces (x^0, λ^0) es un punto **minimax** de la forma

$$\mathcal{L}(x^0, \lambda^0) = \max_{\mathbb{X}} \min_{\mathbb{R}^{m+}} \mathcal{L}(x, \lambda) = \min_{\mathbb{N}} \max_{\mathbb{D}} \mathcal{L}(x, \lambda)$$

entonces $\mathcal{L}(x, \lambda)$ posee puntos de sillas en $D \times \mathbb{R}^{m+}$, y (x^0, λ^0) es uno de ellos.

§1.7. Programas matemáticos con restricciones: sin derivabilidad

Consideremos problemas de optimización descritos de la siguiente forma:

$$\begin{gathered} minf(x) \\ sujeto\ a\ x \in \Omega \subseteq \mathbb{R}^n \end{gathered}$$

donde f es una función definida en Ω, el cual es un subconjunto de R^n definido como

$$\Omega = \{x \in \mathbb{R}^n / g_i(x) \leqslant 0, \forall i = 1, 2, ..., n\}$$

Emplearemos elementos de análisis convexo ya presentados para estudiar programas matemáticos no lineales descritos como en (1). Supongamos que f,g_1,g_2, ..., g_m son funciones definidas en un conjunto abierto $A \subseteq \mathbb{R}^n$ y planteamos el programa no lineal:

$$\begin{gathered} minf(x) \\ sujeto\ a\ g_i(x), \forall i = 1, 2, ..., m \end{gathered}$$

Con este programa matemático podemos asociar dos tipos particulares de problemas conocidos como el problema de punto de silla de Fritz-John y el problema de Kuhn-Tucker.

§1.7.1. Problema de Fritz-John

Este problema consiste en encontrar un punto $\bar{x} \in A$, un valor $\bar{\gamma}_0 \geqslant 0$ y un vector $\bar{\lambda} \in \mathbb{R}_+{}^n$ tales que:

$$\varphi(\bar{x}, \bar{\gamma}_0, \lambda) \leqslant \varphi(\bar{x}, \bar{\gamma}_0, \bar{\lambda}) \leqslant \varphi(x, \bar{\gamma}_0, \bar{\lambda})$$

para cada $x \in A$ y cada $\lambda \in \mathbb{R}_+{}^n$, donde φ es una función Lagrangiana definida como:

$$\varphi(x, \gamma_0, \lambda) = \gamma_0.f(x) + \lambda.g(x)$$

siendo $g(x) = (g_1(x), g_2(x), ..., g_m(x))$.

§1.7.2. Problema de Kuhn-Tucker

Este problema consiste en encontrar un punto $\bar{x} \in A$, un vector $\bar{\lambda} \in \mathbb{R}_+{}^n$ tales que:

$$\psi(\bar{x}, \lambda) \leqslant \psi(\bar{x}, \bar{\lambda}) \leqslant \psi(x, \bar{\lambda})$$

para cada $x \in A$ y cada $\lambda \in \mathbb{R}_+{}^n$, donde ψ es una función Lagrangiana definida como:

$$\psi(x, \lambda) = f(x) + \lambda.g(x)$$

siendo $g(x) = (g_1(x), g_2(x), ..., g_m(x))$. Las componentes del vector λ se denominan multiplicadores de Lagrange. Vemos fácilmente que una solución al problema de Fritz-John con $\bar{\gamma}_0 > 1$ nos brinda una solución para el problema de Kuhn-Tucker tomando $(\bar{x}, \frac{\bar{\lambda}}{\gamma_0})$. Basta ver que $\varphi(x, \gamma_0, \lambda) = \gamma_0(f(x) + \frac{\lambda}{\gamma_0}.g(x))$ $= \gamma_0.\psi(x, \frac{\lambda}{\gamma_0})$ cuando $\gamma_0 > 0$.

De igual modo una solución al problema de Kuhn-Tucker nos brinda una solución al problema de Fritz-John tomando $(\bar{x}, 1, \bar{\lambda})$. Basta ver que $\psi(x, \lambda) = \varphi(x, 1, \lambda)$.

CAPÍTULO 2

CONDICIONES NECESARIAS Y SUFICIENTES DE OPTIMALIDAD MONO-OBJETIVO

Considere el siguiente problema de programación no lineal:

$$
\begin{array}{ll}
\min & f_0(x) \\
\text{s.a:} & g_i(x) \leqslant 0, \quad i \in P, \\
 & x \in C
\end{array}
\tag{SP}
$$

donde $i \in P$ son escalares; f_0 y g_i son funciones localmente Lipschitz, y C es un conjunto convexo. Se establece para el conjunto de índice $P = \{1, 2, ..., p\}$ lo siguiente:

- Denotamos el conjunto factible: $F = \{x \in C / g_i(x) \leqslant 0, \ i = 1, ..., p\}$.
- Sea $I(x^*) = \{i \in P / g_i(x^*) = 0\}$ denota el conjunto de índice de las restricciones activas en x^*.
- El conjunto de índice minimal de las restricciones activas de F se denota por:
$$I^{=} = \{i \in P : x \in F \rightarrow g_i(x) = 0\}$$
- Denotamos:
$$I^{<} = I(x^*) \setminus I^{=} = \{i \in I(x^*) : \textit{Existe } x^i \in F \textit{ s.a. } g_i(x^i) < 0\}.$$
Observe que el conjunto $I^{<}$ de índices de las restricciones activas que tienen un punto interior satisfacen la condición de calificación de Slater.

Para probar las caracterizaciones importantes de este capítulo necesitamos el siguiente lema para el programa no lineal (SP) tomados de (2).

Lema 2.0.1. : Si un punto factible x^* de (SP) es óptimal, entonces la siguiente intersección es vacía:

$$Cono(C - x^*) \bigcap F_0 \bigcap G_0 \tag{2.1}$$

donde $F_0 = \{d : f^c{}_0(x^*, d) < 0\}$ y $G_0 = \{d : g^c{}_j(x^*, d) < 0, j \in I(x^*)\}$.

Prueba: (Procedamos por reducción al absurdo.)

Supongamos que existe $d \in Cono(C - x^*)$ tal que $d \in F_0 \bigcap G_0$. Entonces existen números positivos suficientemente pequeño λ_1, λ_2, λ_3, tal que $(x^* + \alpha.d) \in C$, siempre que $0 < \alpha \leqslant \lambda_1$ y

$$
\begin{array}{ll}
f_0(x^* + \alpha.d) < f(x^*), & \forall \alpha \in (0, \lambda_2), \\
g_j(x^* + \alpha.d) < g_j(x^*), & \forall \alpha \in (0, \lambda_3)
\end{array}
\tag{2.2}
$$

Luego $g_j(x^*) < 0$, para $j \notin I(x^*)$, y por la continuidad de g_j, existe $\lambda_4 > 0$ tal que

$$g_j(x^* + \alpha.d) < 0, \quad \forall \alpha \in (0, \lambda_4) \tag{2.3}$$

Ahora, Tomando $\lambda = min\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$; tenemos que: existe $d \in [Cono(C - x^*) \bigcap F_0 \bigcap G_0]$ tal que

$(x^* + \alpha.d) \in C$ y $g_j(x^* + \alpha.d) < 0$, $\forall\alpha \in (0, \lambda)$, es decir $x^* + \alpha.d$ es un mínimo local de (SP). Esto contradice la solución óptimal de (SP) en x^*.

$$\therefore Cono(C - x^*) \bigcap F_0 \bigcap G_0 = \emptyset$$

Observación 3. el Lema 2.0.1 implica que la intersección siguiente también es vacía:

$$Cono(C - x^*) \bigcap \{d : {f^c}_0(x^*, d) < 0\} \bigcap \{d : {g^c}_j(x^*, d) < 0, j \in I(x^*) \backslash I^=\} \tag{2.4}$$

el cual podemos apreciar una idea geométrica del lema 2.0.1 en la Figura 2.0.1.

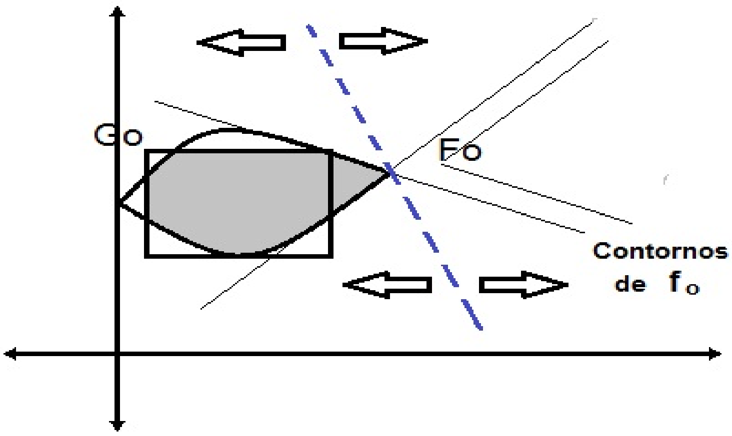

FIGURA 2.0.1: Ilustración de la Condición Necesaria

Teorema 2.0.1. Si un punto factible x^* de (SP) es óptimal y g_i con $i \in I^<$ sean cuasiconvexa estrictamente generalizadas en x^*, entonces existe un vector $d \in -N_c(x^*)$ y escalares no negativos λ_i, $i \in I^<(x^*)$, tal que

$$d \in \{\partial_c f_0(x^*) + \sum_{i \in I^<(x^*)} \lambda_i \partial_c g_i(x^*)\} \tag{I}$$

Prueba: Al aplicar el lema 2.0.1, obtenemos que el sistema

$$d \in Cono(C - x^*), \quad {f^c}_0(x^*, d) < 0, \quad {g^c}_j(x^*, d) < 0, \quad \forall j \in I^=(x^*), \tag{2.5}$$

no tiene solución (No se estudia cuando $j \in I$ ya que es evidente que $G_0 = \emptyset$). Entonces por el lema 1.3.1 el sistema

$$d \in Cono(C - x^*), \quad \max_{\xi \in A} \langle \xi, d \rangle < 0 \tag{2.6}$$

no tiene solución, donde $A = \partial_c f_0(x^*) \cup \bigcup_{j \in I^<(x^*)} \partial_c g_j(x^*)$.

Como A es un conjunto compacto no vacío y el cono $(C - x^*)$ es convexo, por el teorema alternativo [(13), Teorema 3.1],

$$\oslash \in (-(C - x^*)^* + co(A)) \tag{2.7}$$

Es decir, existen $\lambda_0 \geqslant 0$, $\lambda_j \geqslant 0$, $j \in I^<(x^*)$, y

$$\begin{aligned} &\xi_j \in \partial_c g_j(x^*), \quad j \in I^<(x^*) \\ &\xi_0 \in \partial_c f_0(x^*), \quad d \in -N_C(x^*) \end{aligned} \tag{2.8}$$

tal que

$$\lambda_0 + \sum_{j\in I^<(x^*)} \lambda_j = 1 \quad y \quad d = \lambda_0\xi_0 + \sum_{j\in I^<(x^*)} \lambda_j\xi_j \tag{2.9}$$

Observe que cuando $I^= = \emptyset$, tenemos $\lambda_0 \neq 0$.
De lo contrario, si $\lambda_0 = 0$, entonces

$$\sum_{j\in I^<(x^*)} \lambda_j > 0 \quad y \quad d = \sum_{j\in I^<(x^*)} \lambda_j\xi_j$$

lo que implica que el sistema:

$$d \in Cono(C - x^*), \quad \max_{\xi\in B}\langle \xi \ , \ d\rangle < 0 \tag{2.10}$$

no tiene solución, donde $B = \bigcup_{j\in I^<(x^*)} \partial_c g_j(x^*)$

Por otro lado, ya que $I^= \neq \emptyset$, para cada $j \in I^<(x^*)$, existe $x^i \in F$ tal que $g_i(x^i) < 0 = g_i(x^*)$.

Por cuasiconvexidad estrictamente generalizada de g_i, $i \in P$, existe un $d_i \in cono(C - x^*)$, tal que

$$g^c{}_i(x^*, d_i) < 0 \tag{2.11}$$

Por lo tanto, para $\bar{d} = \frac{1}{card(I^<(x^*))} \sum_{i\in I^<(x^*)} d_i$ tenemos:

$$g^c{}_i(x^*, \bar{d}) < 0, \quad \forall i \in I^<(x^*) \tag{2.12}$$

Esto contradice el teorema alternativo [(13), Teorema 3.1] y por lo tanto se puede suponer $\lambda_0 = 1$, y esto completa la prueba. □

Observación 4. Si $I^= = \emptyset$, entonces las restricciones del problema satisfacen a una restricciones de calificación. En este caso, la condición de optimalidad se simplifica.

Teorema 2.0.2. Sea x^* un punto factible de (SP). Suponga que se satisface la condición (I) en x^*. Si las funciones f_0, g_i con $i \in P$, son convexas generalizadas en x^*, entonces x^* es un mínimo de (SP).

Prueba: Sea $x \in F$ un punto factible cualquiera de (SP).

Entonces por la condición (I), para cada $d \in (C - x^*)$ tenemos:

$$\begin{aligned} \langle \xi_0 + \sum_{i\in I^<(x^*)} \lambda_i\xi_i \ , \ d\rangle \geqslant 0 \quad &\Rightarrow \ \langle \xi_0 \ , \ d\rangle + \langle \sum_{i\in I^<(x^*)} \lambda_i\xi_i \ , \ d\rangle \geqslant 0, \\ &\Rightarrow \ \langle \xi_0 \ , \ d\rangle \geqslant -\langle \sum_{i\in I^<(x^*)} \lambda_i\xi_i \ , \ d\rangle \end{aligned} \tag{2.13}$$

para algún $\xi_0 \in \partial_c f_0(x^*)$, $\xi_i \in \partial_c g_i(x^*)$, $i \in I^<(x^*)$.

Dado que cada función es convexa generalizada en x^*, x es factible para (SP), y $\sum_{i\in I^<(x^*)} \lambda_i g_i(x^*) = 0$

entonces,

$$
\begin{aligned}
f_0(x) - f_0(x^*) &\geqslant \langle \xi_0 \ , \ x - x^* \rangle && \forall\ x \in C \\
&\geqslant \langle - \sum_{i \in I^<(x^*)} \lambda_i \xi_i \ , \ x - x^* \rangle && \text{Considerando } d = x - x^* \text{ y } 2.13 \\
&= - \sum_{i \in I^<(x^*)} \lambda_i \xi_i x + \sum_{i \in I^<(x^*)} \lambda_i \xi_i x^* \\
&\geqslant - \sum_{i \in I^<(x^*)} \lambda_i g_i(x) + \sum_{i \in I^<(x^*)} \lambda_i g_i(x^*) \\
&= - \sum_{i \in I^<(x^*)} \lambda_i g_i(x) && \text{ya que } \sum_{i \in I^<(x^*)} \lambda_i g_i(x^*) = 0 \\
&\geqslant 0 && \text{ya que } -g_i(x) > 0 \ y \ \lambda_i \geqslant 0
\end{aligned}
$$

Por lo tanto, $f_0(x) \geqslant f_0(x^*)$ y así x^* es una solución optimal para (SP).□

CAPÍTULO 3

OPTIMIZACIÓN MULTIOBJETIVO

En las ciencias como en la ingeniería se dan, en muchas ocasiones, problemas que requieran encontrar la optimización simultánea de dos o más función objetivo (optimización multiobjetivo). Habrá que optimizar por tanto una función de la forma $f : S \to T$, donde $S \subset \mathbb{R}^n$ y $T \subset \mathbb{R}^m$. Pero el problema está en que normalmente no existe un elemento de S que produzca un óptimo de forma simultánea para cada uno de los m objetivos que componen f. Esto se deberá a la existencia de conflictos entre objetivos, que harán que la mejora de uno de ellos dé lugar a un empeoramiento de algún otro. Pensemos por ejemplo:

- En el caso de un avión con dos objetivos que fueran su velocidad y el ahorro de combustible: un aumento de la velocidad traería consigo una disminución del ahorro. Habrá que llegar por tanto a una situación de compromiso en la que todos los objetivos sean satisfechos en un grado aceptable, desde el punto de vista de diseño.
- En la fabricación de cierto producto que se requiera minimizar el costo de su producción, reducir la cantidad de horas de sobretiempos y el de reducir al mínimo la merma de fabricación de dicho producto; el cual evidencia que el problema posee más de un objetivo.

A diferencia de los problemas de optimización con un único objetivo, el concepto de óptimo es ahora relativo y será necesario decidir de alguna forma cuál es la mejor solución (o cuáles son las mejores soluciones) al problema presente.

En términos matemáticos, el problema de optimización multiobjetivo restringido puede establecerse como: encontrar un vector $x = (x_1, x_2, ..., x_n)^t$, que satisfaga:

$$\begin{array}{ll} \min & f(x) = (f_1(x), ..., f_m(x)) \\ \text{s.a:} & g(x) = (g_1(x), ..., g_p(x)) \leqslant 0, \\ & h(x) = (h_1(x), ..., h_q(x)) = 0, \\ & x \in C \end{array} \qquad (MP)$$

donde C es un conjunto convexo, con $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^p$ y $h : \mathbb{R}^n \to \mathbb{R}^q$ funciones localmente Lipschitz.

Si una solución x satisface todas las restricciones se dice que es factible y en caso contrario se dice que no es factible; Además el conjunto de todas las soluciones factibles es llamado *Región Factible* $\mathcal{X}$ (Espacio de Decisión).

La teoría que a continuación presentamos en las siguientes secciones fue tomada de (4), el cual nos ayudo a entender los problemas de optimización multiobjetivo.

§3.1. Óptimo de Pareto

En contraste con la optimización mono-objetivo, la solución a un problema multiobjetivo es más un concepto que una definición. Por regla general, no hay una única solución global y es necesario determinar un conjunto de puntos que cumplan una definición predeterminada de óptimo. El concepto predominante en la definición de este óptimo es el óptimo en el sentido de Pareto, que se define de la siguiente manera:

Definición 3.1.1. (Óptimo de Pareto) Un punto $\bar{x} \in \mathbb{X}$, es un óptimo en el sentido de Pareto de (MP) si y sólo si no existe otro punto $x \in \mathbb{X}$, de manera que $f_i(x) < f_i(\bar{x})$ para algún $i \in M$ y $f_j(x) \leqslant f_j(\bar{x})$ para todo $j \in M$.

Con frecuencia, los algoritmos proporcionan soluciones que puede que no sean óptimos de Pareto, pero que pueden satisfacer otros criterios, haciéndolos apropiados para aplicaciones prácticas. Es por ello que se define el óptimo débil de Pareto como sigue:

Definición 3.1.2. (Óptimo Débil de Pareto) Un punto $\bar{x} \in \mathbb{X}$, es un óptimo en el sentido de Pareto si y sólo si no existe otro punto $x \in \mathbb{X}$, de manera que $f(x) \leqslant f(\bar{x})$. Es decir, un punto es un óptimo débil de Pareto si no hay otro punto que mejore todas las funciones objetivo simultáneamente.

Por el contrario, un punto es un óptimo de Pareto si no existe otro punto que mejore al menos una de las funciones objetivo sin empeorar cualquier otra de ellas. Los óptimos de Pareto son también óptimos débiles de Pareto, pero el recíproco no es cierto.

§3.2. Eficiencia y Dominación

La eficiencia, admisibilidad o no inferioridad, es otro concepto básico en el campo de la optimización multiobjetivo y suele definirse como sigue:

Definición 3.2.1. (Eficiencia e ineficiencia) Un punto $\bar{x} \in \mathbb{X}$ es eficiente del problema (MP) si, y sólo si, no existe otro punto $x \in \mathbb{X}$ tal que $f(x) \leqslant f(\bar{x})$ con al menos una $f_i(x) \leqslant f_i(\bar{x})$. En caso contrario, $\bar{x}$ es ineficiente. El conjunto de todos los puntos eficientes de (MP) se denomina Conjunto Eficiente y se denota por $\mathcal{X}$.

Se presentará la definición de una solución propiamente eficiente y de un problema escalarizado; con el cual podemos obtener un par de caracterizaciones para este tipo de problema y una solución optimal.

Definición 3.2.2. x^0 es llamado *Solución Propiamente Eficiente* de (MP), si se cumple que:

1. x^0 es eficiente;

2. Existe un escalar $K > 0$ tal que, para cada $i \in M$ tenemos,

$$\frac{f_i(x)-f_i(x^0)}{f_j(x^0)-f_j(x)} \leqslant K$$

Para algún $j \in M$ tal que $f_j(x) < f_j(x^0)$ siempre que $x \in \mathcal{X}$ y $f_i > f_i(x^0)$.

Así, podemos decir que un punto eficiente que no es propiamente eficiente es llamado *Impropiamente Eficiente*.

Definición 3.2.3. Definiremos un problema escalarizado como:

$$Max \sum_{i=1}^{p} \lambda_i f_i(x) \qquad \text{sujeto a} \qquad x \in \mathcal{X} \qquad\qquad (P_\lambda)$$

donde los λ_i son parámetros no negativos frecuentemente normalizado, ajustado $\sum_{i=1}^{p} \lambda_i = 1$.

Teorema 3.2.1. Sea $\lambda_i \geqslant 0$ (i=1,...,p) fijo. Si x^0 es optimal en (P_λ), entonces x^0 es propiamente eficiente en (MP).

Teorema 3.2.2. Sea $\mathcal{X}$ un conjunto convexo, y sea f_i cóncavo sobre $\mathcal{X}$. Entonces, x^0 es propiamente eficiente en (MP), si y sólo si, x^0 es optimal en (P_λ) para algún λ con componentes estrictamente positivas.

En optimización multiobjetivo las funciones constituyen un espacio multidimensional en adición al espacio de decisión. Éste espacio adicional es llamado *Espacio objetivo* $\mathcal{Y}$. Para cada solución x en el espacio de variables de decisión existe un punto en el espacio objetivo denotado por $f(x) = Z = (z_1, z_2, ..., z_m)^t$; en el que muchas veces nos referiremos a $f(x)$ como el vector de optimización porque es un vector de objetivos optimizado en lugar de un único objetivo. De esta manera presentaremos las situaciones en la que podemos decir cuando una solución es mejor que otra.

Definición 3.2.4. (Puntos dominados y no dominados) Un vector de funciones objetivo, $f(x*) \in Z$, es no dominado si, y sólo si, no existe otro vector, $f(x) \in Z$ tal que $f(x) \leqslant f(x*)$ con al menos una $f_i(x) < f_i(x*)$ e $i \in \{1, 2, ..., m\}$. En otro caso, $f(x*)$ es dominado.

Observación 5. : En la práctica, podemos decir que:

- las definiciones 3.2.1. y 3.2.4. resultan equivalentes. Sin embargo, la eficiencia se refiere generalmente a un vector de variables de decisión en el espacio de las decisiones mientras que la dominación lo hace a un vector de funciones en el espacio de los criterios.
- La definición de óptimo de Pareto es parecida también a la de eficiencia, de manera que un punto que es óptimo de Pareto en el espacio de los criterios se suele considerar con frecuencia como no dominado. El óptimo de Pareto es, en realidad, un sutil caso particular de eficiencia, pero su distinción es irrelevante en términos de aplicación práctica.

En problemas de optimización multiobjetivo, las funciones objetivos pueden ser categorizadas como sin conflicto, parcialmente en conflicto y totalmente en conflicto. Por lo que se tiene:

Definición 3.2.5. Las funciones objetivos se dicen que:

- Están *Sin Conflicto* entre si, cuando cualquier par de funciones $\vec{x_a}$ y $\vec{x_b}$ en un conjunto S satisfacen $f(\vec{x_a}) \lhd f(\vec{x_b}) \vee f(\vec{x_b}) \lhd f(\vec{x_a})$.

  Ejemplo 3.2.1. : Al determinar un número real x que sea mínimo en el problema:

  $$Min\ f(x) = \begin{cases} f_1(x) = x^2 \\ f_2(x) = x^2 + 2 \end{cases} \qquad con - 2 \leqslant x \leqslant 2$$

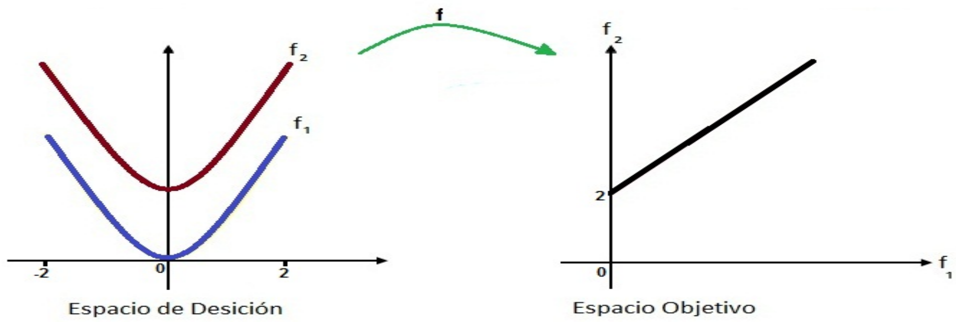

FIGURA 3.2.1: f_1 y f_2 Sin Conflicto. Punto mínimo $x = 0$

- Están *Parcialmente en Conflicto* si existen dos conjuntos de soluciones P_a y P_b, los cuales cuentan con al menos un elemento cada uno $\vec{x_a} \in P_a$, $\vec{x_b} \in P_b$ que cumplen con $f_(\vec{x_a}) \lhd f(\vec{x_b}) \vee f(\vec{x_b}) \lhd f(\vec{x_a})$.

  Ejemplo 3.2.2. : Al determinar un número real x que sea mínimo en el problema planteado por Shaffer (1984):

  $$f(x) = \begin{cases} f_1(x) = x^2 \\ f_2(x) = (x-2)^2 \end{cases} \qquad con - 2 \leqslant x \leqslant 4$$

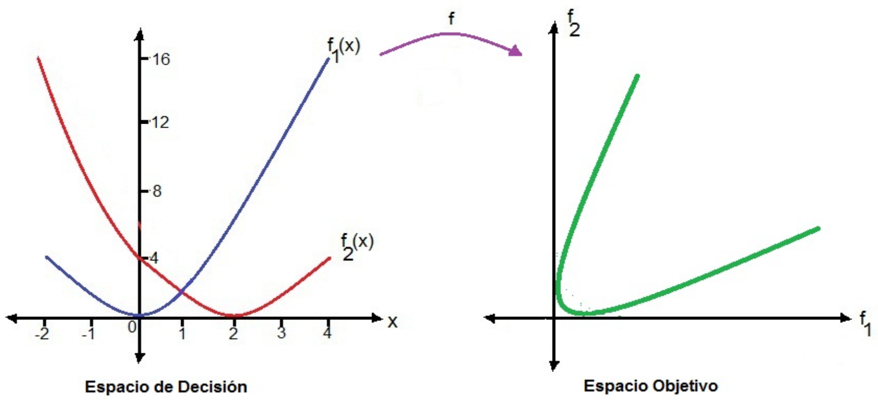

FIGURA 3.2.2: Gráfica del problema de Shaffer

:

- Están *Totalmente en Conflicto* si no existen dos soluciones $\vec{x}_a$ y $\vec{x}_b$ en un conjunto de soluciones S tal que $f(\vec{x}_a) \lhd f(\vec{x}_b) \vee f(\vec{x}_b) \lhd f(\vec{x}_a)$.

  Ejemplo 3.2.3. : Al determinar un número real x que sea mínimo en el problema de optimización multiobjetivo:
  $f(x) = \begin{cases} f_1(x) = x \\ f_2(x) = -x \end{cases} \qquad con -2 \leqslant x \leqslant 2$

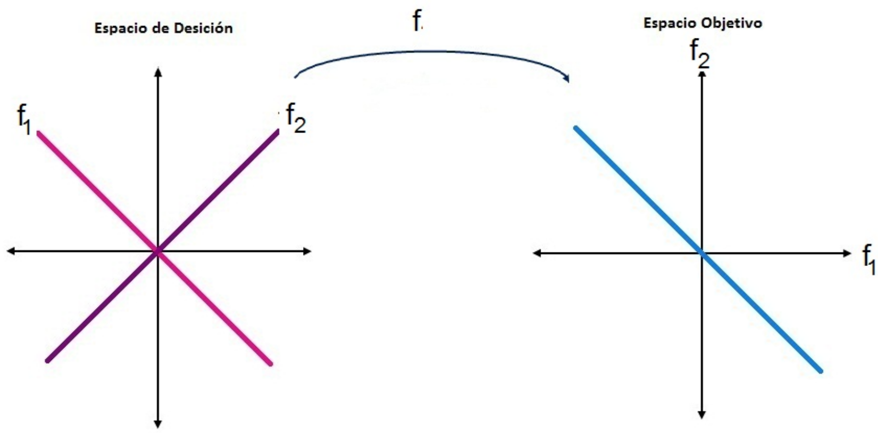

FIGURA 3.2.3: Totalmente en Conflicto

En optimización multiobjetivo también se puede usar teoría del Análisis Convexo para definir lo que es una solución eficiente. Para ello se presentará unos conceptos previos.

Definición 3.2.6. : $C \subset \mathbb{R}^p$, es un *Cono* si, y sólo si, para todo $\vec{y} \in C, \quad \vec{y} \geqslant 0$. (Ver Figura 3.2.4(a))

Definición 3.2.7. : $\bar{C} \subset \mathbb{R}^p$, es un *Cono Negativo* si, y sólo si, para todo $\vec{y} \in \bar{C}, \vec{y} \leqslant 0$.(Ver Figura 3.2.4(b))

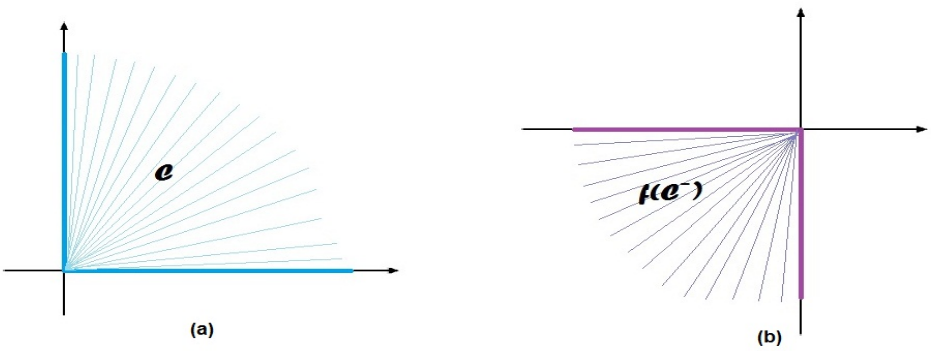

FIGURA 3.2.4: Cono (a) y Cono Negativo (b)

En este orden de ideas mostraremos una definición equivalente de solución eficiente.

Definición 3.2.8. Si $\hat{x} \in \mathcal{X}$ diremos que, $\hat{x}$ es *Eficiente* si $f(\mathcal{X}) \bigcap (f(\hat{x}) - \mathbb{R}^m_{\geqslant}) = \{f(\hat{x})\}$.

En la figura 3.2.5 se muestran los conos correspondiente al espacio de decisión y al espacio objetivo respectivamente.

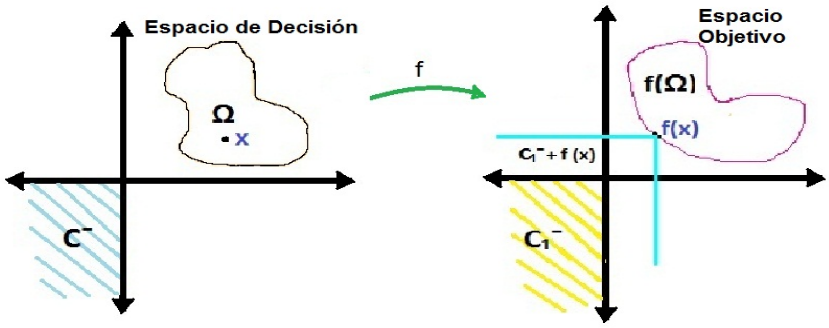

FIGURA 3.2.5: Espacio de decisión y Espacio objetivo

Observación 6. En relación a las soluciones eficientes y los elementos no dominados, tenemos que:

- El conjunto de todas las soluciones eficientes $\hat{x} \in \mathcal{X}$ es denotado por $\mathcal{X}_E$ y es llamado *Conjunto Eficiente*.
- El conjunto de todos los puntos no dominados $\hat{y} = f(\hat{x}) \in \mathcal{Y}$, donde $\hat{x} \in \mathcal{X}_E$, se denota $\mathcal{Y}_N$ y es llamado *Conjunto No Dominado*.

A continuación presentaremos algunas propiedades fundamentales de puntos no dominados y soluciones eficientes que nos ayuda a mostrar las condiciones para la existencia de soluciones eficientes y puntos no dominados ya que es nuestra primera preocupación en el estudio de optimización multicriterio.

Proposicion 3.2.1. $\mathcal{Y}_N = (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N$

Prueba: Este resultado es trivial si, $\mathcal{Y} = \emptyset$, pues $\mathcal{Y} + \mathbb{R}^p_{\geqq} = \mathbb{R}^p_{\geqq}$

y por tanto los subconjuntos no dominados de ambos es vacío.
Dado que $\mathcal{Y}_N = (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N \equiv \mathcal{Y}_N \subset (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N \wedge (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N \subset \mathcal{Y}_N$. Entonces:

1. Probemos $\mathcal{Y}_N \subset (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N$
   Para $\mathcal{Y} \neq \emptyset$. Y procedamos por reducción al absurdo:

   a) Supongamos $y \in (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N$, pero $y \in \mathcal{Y}_N$. Hay dos posibilidades: Si $y \notin \mathcal{Y}_N$, existe $y' \in Y$ y $0 \neq d \in \mathbb{R}^p_{\geqq}$ tal que $y = y' + d$. Por lo que $y' = y' + 0 \in \mathcal{Y} + \mathbb{R}^p_{\geqq}$ y así $y \notin (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N$, contradicción!!.

   b) Si $y \in \mathcal{Y}$ y existe $y' \in \mathcal{Y}$ tal que $y' \leqslant y$.
   Sea $d = y - y'$, el cual esta en $\mathbb{R}^p_{\geqq} \setminus \{0\}$. Por lo que, $y = y' + d$ y $y \notin (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N$, contradicción con lo supuesto.

$$\therefore \text{ en ambos casos } y \in \mathcal{Y}_N.$$

2. Probemos $(\mathcal{Y} + \mathbb{R}^p_{\geqq})_N \subset \mathcal{Y}_N$
   Supongamos que $y \in \mathcal{Y}_N$, pero $y \notin (Y + \mathbb{R}^p_{\geqq})_N$.
   Entonces, existe algún $y' \in \mathcal{Y} + \mathbb{R}^p_{\geq}$ con $y - y' = d' \in \mathbb{R}^p_{\geq} \backslash \{0\}$.
   Sea $y' = y'' + d$ con $y'' \in Y, d'' \in \mathbb{R}^p_{\geqq}$. Así $y = y' + d' = (y'' + d'') + d' = y'' + (d'' + d') = y'' + d$, con $d = d'' + d' \in \mathbb{R}^p_{\geq} \backslash \{0\}$. Por lo tanto $y \in \mathcal{Y}_N$, el cual contradice lo supuesto.

$$\therefore y \in (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N.$$

$$\therefore \text{ De (1) y (2) se tiene que } \mathcal{Y}_N = (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N.$$

La siguiente proposición nos garantiza que el conjunto de los puntos no dominados se encuentran en la frontera del conjunto eficiente.

Proposicion 3.2.2. $\mathcal{Y}_N \subset bd(\mathcal{Y})$.

Prueba: Sea $y \in \mathcal{Y}_N$ y supongamos por reducción al absurdo que $y \notin bd(\mathcal{Y})$.

Así, $y \in int\mathcal{Y}$ y existe una $\epsilon - vencidad B(y, \epsilon)$ de y (con $B(y, \epsilon) := y + B(0, \epsilon) \subset \mathcal{Y}$, el cual $B(0, \epsilon)$ es una bola abierta de radio ϵ y centrada en el origen).
Sea $d \neq 0,\ tal\ que\ d \in \mathbb{R}^p_{\geqq}$.
Entonces podemos escoger algún $\alpha \in \mathbb{R}, 0 < \alpha < \epsilon$ tal que $\alpha d \in B(0, \epsilon)$.
Ahora, $y - \alpha d \in \mathcal{Y}$ con $\alpha d \in \mathbb{R}^p_{\geq} \backslash \{0\}$, es decir, $y \notin \mathcal{Y}_N$, el cual contradice lo supuesto.

$$\therefore \mathcal{Y}_N \subset bd(\mathcal{Y}).$$

Proposicion 3.2.3. $(\mathcal{Y}_1 + \mathcal{Y}_2)_N \subset (\mathcal{Y}_1)_N + (\mathcal{Y}_2)_N$.

Prueba: Sea $Y \in (\mathcal{Y}_1 + \mathcal{Y}_2)_N$ para $Y = Y^1 + Y^2$, para algún $Y^1 \in \mathcal{Y}_1$ y $Y^2 \in \mathcal{Y}_2$.

Debemos probar que $Y^1 \in (\mathcal{Y}_1)_N$ y $Y^2 \in (\mathcal{Y}_2)_N$.
Supongamos por reducción al absurdo que $Y^1 \notin (\mathcal{Y}_1)_N$, esto es, existe algún $Y' \in Y^1$ y $d \in \mathbb{R}^p_{\geq}$ tal que $Y^1 = Y' + d$, y así:

$$Y = Y^1 + Y^2 = Y' + d + Y^2 \notin (\mathcal{Y}_1 + \mathcal{Y}_2)_N (\rightarrow\leftarrow)$$

Ya que $Y' \in Y^1$ y $Y^2 \in Y_2$. Luego $Y^1 \in (\mathcal{Y}_1)_N$.
Análogamente se prueba que $Y^2 \in (\mathcal{Y}_2)_N$.

$$\therefore (\mathcal{Y}_1 + \mathcal{Y}_2)_N \subset (\mathcal{Y}_1)_N + (\mathcal{Y}_2)_N$$

Observación 7. La inclusión $(\mathcal{Y}_1)_N + (\mathcal{Y}_2)_N \subset (\mathcal{Y}_1 + \mathcal{Y}_2)_N$ no es verdadera en forma general.

Proposicion 3.2.4. $(\alpha\mathcal{Y})_N = \alpha(\mathcal{Y})_N$ para $\alpha \in \mathbb{R}$ y $\alpha > 0$.

Prueba: Este resultado es trivial si $\alpha = 1$. Para $\alpha > 0$.

Dado que $(\alpha\mathcal{Y})_N = \alpha(\mathcal{Y})_N \equiv (\alpha\mathcal{Y})_N \subset \alpha(\mathcal{Y})_N \wedge \alpha(\mathcal{Y})_N \subset (\alpha\mathcal{Y})_N$. Entonces:

1. Probemos $(\alpha\mathcal{Y})_N \subset \alpha(\mathcal{Y})_N$
   Sea $y \in (\alpha\mathcal{Y})_N$; esto es $y \in \alpha\mathcal{Y}$ y así existe $y' \in \mathcal{Y}$ tal que $y = \alpha y'$. Procedamos por reducción al absurdo suponiendo que $y' \notin \mathcal{Y}_N$.
$$\begin{aligned} y' \notin \mathcal{Y}_N \quad & \Rightarrow \; \textit{existe } y'' \in \mathcal{Y} \textit{ y } d \in \mathbb{R}^p_{\geqq} \textit{ con } d \neq 0 \textit{ tal que } y' = y'' + d \\ & \Rightarrow \; y = \alpha y' = \alpha(y'' + d) = \alpha y'' + \alpha d \textit{ con } \alpha d \in \mathbb{R}^p_{\geqq} \\ & \Rightarrow \; y \notin (\alpha\mathcal{Y})_N \text{ Contradicción} \end{aligned}$$
$$\therefore y' \in \mathcal{Y}_N$$
$$\therefore y = \alpha y' \in \alpha\mathcal{Y}_N$$

2. Probemos $\alpha(\mathcal{Y})_N \subset (\alpha\mathcal{Y})_N$
   Sea $y \in \alpha\mathcal{Y}_N$; esto es, existe $y' \in \mathcal{Y}_N$ tal que $y = \alpha y'$. Puesto que $y' \in \mathcal{Y}'$ entonces $y \in \alpha y$. Procedamos por reducción al absurdo suponiendo que $y \notin (\alpha\mathcal{Y})_N$.
$$\begin{aligned} y \notin (\alpha\mathcal{Y})_N \quad & \Rightarrow \; \textit{existe } y'' \in \mathcal{Y} \textit{ y } d \in \mathbb{R}^p_{\geqq} \textit{ con } d \neq 0 \textit{ tal que } \alpha y' = y = \alpha y'' + d \\ & \Rightarrow \; \alpha y' - \alpha y'' = d \\ & \Rightarrow \; \alpha(y' - y'') = d \\ & \Rightarrow \; y' - y'' \in \mathbb{R}^p_{\geqq} \\ & \Rightarrow \; y' \geqslant y'' \textit{ ya que } \mathbb{R}^p_{\geqq} > 0 \\ & \Rightarrow \; y' \notin \mathcal{Y}_N \text{ Contradicción} \end{aligned}$$
$$\therefore y' \in (\alpha\mathcal{Y})_N$$
$$\therefore y \in (\alpha\mathcal{Y})_N$$

$$\therefore \; \text{De (1) y (2) se tiene que } (\alpha\mathcal{Y})_N = \alpha(\mathcal{Y})_N$$

Ahora veamos cuando una solución factible es débilmente y estrictamente eficiente.

Definición 3.2.9. Una solución factible $\hat{x} \in \mathcal{X}$ es:

- *Débilmente Eficiente* (Débilmente Pareto optimal), si no hay otro $x \in \mathcal{X}$ tal que $f(x) < f(\hat{x})$, es decir, $f_k(x) < f_k(\hat{x})$, para todo $k = 1, ..., p$. El punto $\hat{y} = f(\hat{x})$ es llamado *Débilmente No Dominado*.(Ver Figura 3.2.6)

- *Estrictamente Eficiente* (Estrictamente Pareto optimal) si no hay $x \in \mathcal{X}, x \neq \hat{x}$ tal que $f(x) \leqq f(\hat{x})$. Los conjuntos débilmente (estrictamente) eficiente y no dominados son denotados por $\mathcal{X}_{wE}(\mathcal{X}_{sE})$ y $\mathcal{Y}_{wE}$.

De la definición anterior, se obtiene que:

$$\mathcal{Y} \subset \mathcal{Y}_{wN} \quad y \quad \mathcal{X}_{sE} \subset \mathcal{X}_E \subset \mathcal{X}_{wE}$$

Observación 8. Una solución factible $\hat{x} \in \mathcal{X}$ es débilmente eficiente si y sólo si $(f(\hat{x}) - \mathbb{R}^p_{>}) \bigcap \mathcal{Y} = \emptyset$.

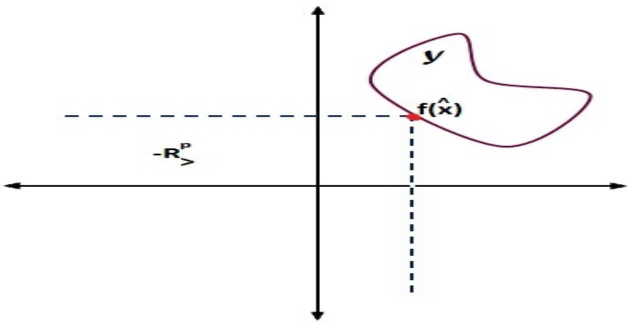

FIGURA 3.2.6: Solución débilmente eficiente

Ejemplo 3.2.4. Consideremos el conjunto

$$\mathcal{Y} = \{(y_1, y_2) \in \mathbb{R}^2 : 0 < y_1 < 1, 0 \leqslant y_2 \leqslant 1\}$$

Cuya gráfica viene dada por la Figura 3.2.7(a)

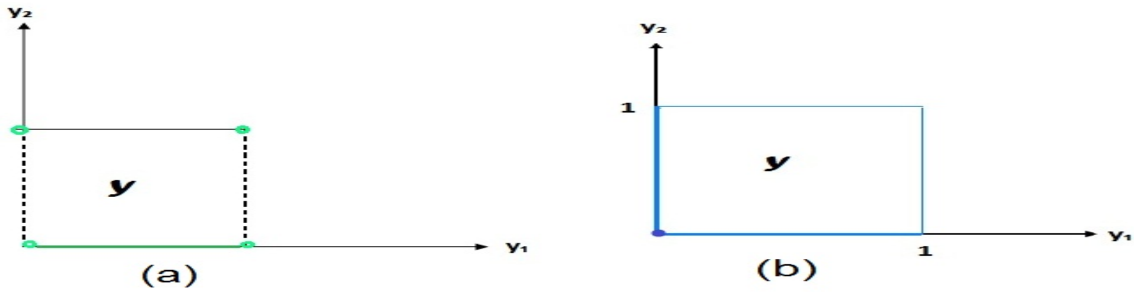

FIGURA 3.2.7: $(a)\mathcal{Y}_N$ es vacío,$\mathcal{Y}_{wN}$ y (b)Puntos dominados y débilmente no dominados

- $\mathcal{Y}_N = \emptyset$, ya que al intersectar el cono $(-\mathbb{R}^p_{\geqq}) + (0,0)$ con el conjunto $\mathcal{Y}$, el resultado es vacío.
- $\mathcal{Y}_{wN} = (0,1) \times \{0\}$.

Ahora veamos el siguiente conjunto

$$\mathcal{Y} = \{(y_1, y_2) \in \mathbb{R}^2 : 0 \leqslant y_i \leqslant 1\}$$

Cuya gráfica viene dada por la Figura 3.2.7(b).
notemos que, $\mathcal{Y}_N = \{0\}$. $\mathcal{Y}_{wN} = \{(y_1, y_2) \in Y : y_1 = 0 \quad \text{o} \quad y_2 = 0\}$.

CAPÍTULO 4

CONDICIONES NECESARIAS Y SUFICIENTES DE OPTIMALIDAD MULTIOBJETIVO

En el presente capitulo mostraremos los objetivos principales de la investigación, el cual se centra en las caracterizaciones de la optimalidad del problema multiobjetivo.

§4.1. PROBLEMA MULTIOBJETIVO

Consideremos el problema multiobjetivo siguiente:

$$\begin{aligned} \min \quad & f(x) = (f_1(x), ..., f_m(x)) \\ \text{s.a:} \quad & g(x) = (g_1(x), ..., g_p(x)) \leqslant 0, \\ & x \in C \end{aligned} \tag{MP}$$

donde C es un conjunto convexo y $f : R^n \to R^m$ y $g : R^n \to R^p$ son funciones localmente Lipschitz.
El Índice se establece $M = \{1, 2, ..., m\}$ y $P = \{1, 2, ..., p\}$.
Denotamos el conjunto factible: $F = \{x \in C / g_i(x) \leqslant 0, \quad i = 1, ..., p\}$.
Sea $I(x^*) = \{i \in P / g_i(x^*) = 0\}$ denota el conjunto de índice de las restricciones activas en x^*.
El conjunto de índice minimal de las restricciones activas de F se denota por:

$$I^{=} = \{i \in P : x \in F \to g_i(x) = 0\} \tag{4.1}$$

También denotamos:

$$I^{<} = I(x^*) \setminus I^{=} = \{i \in I(x^*) : \mathit{Existe}\ x_i \in F\ \mathit{s.a.}\ g_i(x_i) < 0\} \tag{4.2}$$

Para un $r \in M$ fijo y $x^* \in R^n$, denota

$$\begin{aligned} & M^r = \mathrm{M} \setminus \mathrm{r} \\ & F^r(x^*) = \{x : f_i(x) \leqslant f_i(x^*)\} \\ & M^{r=}(x^*) = \{i \in M^r : f_i(x) = f_i(x^*), \forall x \in F^r(x^*)\} \end{aligned} \tag{4.3}$$

Consideramos el problema Multiobjetivo (MP). El siguiente resultado (de (12)) es una caracterización bien conocida de las soluciones eficientes de (MP).

Lema 4.1.1. : Una solución factible x^* en (MP) es una solución eficiente si y sólo si x^* resuelve el problema de optimización escalar

$$\begin{aligned} \min \quad & f_r(x) \\ \text{s.a:} \quad & f_i(x) \leqslant f_i(x^*), \quad i \in M^r, \\ & g_j(x) \leqslant 0, \qquad j \in P, \\ & x \in C, \qquad\quad \mathit{para\ cada}\ r = 1, 2, ..., m \end{aligned} \tag{$P_r(x^*)$}$$

El Teorema 2.0.3 y el Lema 4.1.1 pueden ser utilizados para obtener las condiciones de optimalidad necesaria y suficiente, sin calificación de las restricciones para el problema de programación multiobjetivo (MP).

Teorema 4.1.1. Si x^* es una solución eficiente para (MP) y g_i con $i \in P$, son cuasiconvexa estrictamente generalizadas en x^*, entonces existen escalares $\lambda_i{}^* > 0$, $i \in M$, con

$$\sum_{i=1}^{m} \lambda_i{}^* = 1 \quad y \quad \mu_j{}^* \geqslant 0, \quad j \in I^{<}(x^*)$$

tal que

$$\oslash \in \Big(\sum_{i \in M} \lambda_i{}^* \partial_c f_i(x^*) + \sum_{j \in I^{<}(x^*)} \mu_j{}^* \partial_c g_j(x^*) + N_c(x^*) \Big) \tag{4.5}$$

Recíprocamente, si (4.5) se cumple y para todo f_i, g_j son convexa generalizadas en x^*, entonces x^* es una solución eficiente para (MP).

Prueba: $(Necesaria)$ $(\Rightarrow)$ Puesto que x^* es una solución eficiente para (MP),

entonces por el lema 4.1.1, x^* es una solución óptimal para cada uno $(P_r(x*))$, $r \in M$.
Por lo tanto, por el Teorema 2.0.3, existe $\lambda_{ri}{}^* \geqslant 0, \quad i \in M^r, \quad \mu_{rj}{}^* \geqslant 0, \quad j \in I^{<}(x^*)$
tal que

$$\oslash \in \Big(\partial_c f_r(x^*) + \sum_{i \in M^r(M_c^r(x^*))} \lambda_i{}^* \partial_c f_i(x^*) + \sum_{j \in I^{<}(x^*)} \mu_{rj}{}^* \partial_c g_j(x^*) + N_c(x^*) \Big) \tag{4.6}$$

para cada $r \in M$. Sumando ahora el anterior encima de $r \in M$, con la escala apropiada, y usando las propiedades de $N_c(x^*)$ entonces se determinan la condición necesaria.

$(Suficiencia)$ $(\Leftarrow)$ (Procedamos por Reducción al absurdo)
Supongamos que x^* no es solución eficiente para (MP) y se cumple (4.5). Entonces existen

$$\begin{array}{llll} \xi_i \in \partial_c f_i(x^*), & i = 1, 2, ..., m, & \eta_i \in \partial_c g_i(x^*), & i \in I^{<}(x^*) \\ d \in -N_c(x^*), & \lambda_i{}^* > 0, \quad i \in M, & \mu_i{}^* \geqslant 0, & i \in I^{<}(x^*) \end{array} \tag{4.7}$$

tal que

$$\sum_{i \in M} \lambda_i{}^* \xi_i + \sum_{i \in I^{<}(x^*)} \mu_i{}^* \eta_i \quad = d \tag{4.8}$$

y existe $u \in F$ tal que

$$\begin{array}{ll} f_i(u) < f_i(x^*), & \text{para algún i} \\ f_j(u) \leqslant f_j(x^*), & para\ todo\ j \in M \end{array} \tag{4.9}$$

Por lo tanto

$$\sum_{i \in M} \lambda_i{}^* f_i(u) < \sum_{i \in M} \lambda_i{}^* f_i(x^*) \tag{4.10}$$

Luego como las funciones f_i, $i \in M$ y g_i, $i \in P$ son convexas generalizadas, se deduce que

$$\begin{aligned} \sum_{i \in M} \lambda_i{}^* f_i(u) - \sum_{i \in M} \lambda_i{}^* f_i(x^*) \quad &\geqslant (u - x^*)^t \sum_{i \in M} \lambda_i{}^* \xi_i = (u - x^*)^t (d - \sum_{i \in I^{<}(x^*)} \mu_i{}^* \eta_i) \\ &\geqslant -(u - x^*)^t \sum_{i \in I^{<}(x^*)} \mu_i{}^* \eta_i \\ &= - \sum_{i \in I^{<}(x^*)} \mu_i{}^* u^t \eta_i + \sum_{i \in I^{<}(x^*)} \mu_i{}^* x^{*t} \eta_i \\ &\geqslant - \sum_{i \in I^{<}(x^*)} \eta_i g_i(u) + \sum_{i \in I^{<}(x^*)} \eta_i g_i(x^*) \end{aligned} \tag{4.11}$$

$$\begin{aligned}\sum_{i\in M}\lambda_i{}^* f_i(u) - \sum_{i\in M}\lambda_i{}^* f_i(x^*) &= -\sum_{i\in I^<(x^*)}\eta_i g_i(u)\\ &\geqslant 0\end{aligned}$$

$$\therefore \sum_{i\in M}\lambda_i{}^* f_i(u) - \sum_{i\in M}\lambda_i{}^* f_i(x^*) \geqslant 0$$

Obteniéndose que $\sum_{i\in M}\lambda_i{}^* f_i(u) \geqslant \sum_{i\in M}\lambda_i{}^* f_i(x^*)$ lo cual contradice (4.10), por lo tanto, x^* es una solución eficiente.

§4.2. Dualidad del Problema Multiobjetivo

Recordemos el Problema Dual de Wolfe para un objetivo:
Cuando las funciones son continuas y diferenciables el Dual de un Problema Convexo puede ser representado convenientemente. Para cualquier λ fijo, el punto $\bar{x}$ minimiza $L(x,\lambda)$ si y solo si $\nabla_x L(x,\lambda)\mid_{x=\bar{x}}=\emptyset$.
el problema dual puede escribirse como:

$$\begin{aligned}&\max_{x,\lambda\geqslant 0} && L(x,\lambda)\\ &\text{s.a:} && \nabla_x L(x,\lambda)=\emptyset\end{aligned} \tag{4.12}$$

el cual se conoce como Dual de Wolfe donde $L(x,\lambda)=f(x)+\lambda^t g(x)$.

Ahora asociamos el vector tipo Dual de Wolfe (8) al problema primal multiobjetivo (MP):

$$\begin{aligned}&\max && \bar{f}(u)=(f_1(u)+y^t g(u),...,f_m(u)+y^t g(u))\\ &\text{s.a:} && 0\in(\sum_{i\in M}\lambda_i\partial_c f_i(u)+\sum_{i\in P}y_i\partial_c g_i(u)+N_c(u))\\ &&& y_i\geqslant 0,\quad i=1,2,...,p\\ &&& \lambda_i\geqslant 0,\quad i=1,2,...,m\\ &&& \sum_{i\in M}\lambda_i=1,\quad g_I{}^{=}(u)=0\end{aligned} \tag{DM}$$

Aquí F_D denota el conjunto de soluciones factibles para (DM) y $g_I{}^{=}(\cdot)$ para $g_i(\cdot)$, $i\in I^{=}$.
Ahora probemos el Teorema Débil Dualidad, el cual es similar a [(6), Teorema5].

Teorema 4.2.1. (Dualidad Débil). Suponga que $x\in F$ y $(u,\lambda,y)\in F_D$ y f_i, $i\in M$, g_j, $j\in P$ son funciones convexas generalizadas en u. Entonces lo siguiente no puede cumplirse:

$$\begin{aligned}&f_j(x)\leqslant f_j(u)+y^t g(u) && \forall j\in M\\ &f_i(x)< f_i(u)+y^t g(u) && \text{Para algún } i\in M\end{aligned} \tag{4.13}$$

Prueba: (procedemos por reducción al absurdo)

Sea x solución factible para (MP) y sea (u,λ,y) solución factible para (DM).
Supongamos que el resultado de (4.13) se cumple. Entonces:

$$\sum_{i\in M}\lambda_i f_i(x) < \sum_{i\in M}\lambda_i f_i(u)+y^t g(u) \tag{4.14}$$

Por la factibilidad de (u,λ,y), se satisfacen las restricciones de (DM), así existen $\xi_i\in\partial_c f_i(u)$, $i\in M$, $\eta_i\in\partial_c g_i(u)$, $i\in P$, y $d\in -N_c(u)$, tal que

$$\sum_{i\in M}\lambda_i\xi_i+\sum_{i\in P}y_i\eta_i \quad = d \tag{4.15}$$

Entonces

$$
\begin{aligned}
\sum_{i\in M}\lambda_i\{f_i(x)-f_i(u)\}+y^t g(u) &= \sum_{i\in M}\lambda_i f_i(x)-\sum_{i\in M}\lambda_i f_i(u)-y^t g(u)\\
&\geqslant \sum_{i\in M}\lambda_i x^t-\sum_{i\in M}\lambda_i u^t-y^t g(u) \qquad \text{(f convexa)}\\
&= (x-u)^t\sum_{i\in M}\lambda_i\xi_i-y^t g(u)\\
&= (x-u)^t(d-\sum_{i\in P}y_i\eta_i)-y^t g(u) \qquad \text{(por 4.15)}\\
&= (x-u)^t d-(x-u)^t\sum_{i\in P}y_i\eta_i-y^t g(u)\\
&= (u-x)^t\sum_{i\in P}y_i\eta_i-y^t g(u)+(x-u)^t d\\
&= \sum_{i\in P}(u-x)^t y_i\eta_i-y^t g(u)+(x-u)^t d\\
&\geqslant \sum_{i\in P}(y_i(g_i(u)-g_i(x))-y^t g(u) \qquad \text{(g convexa)}\\
&= -y^t g(x)\\
&\geqslant 0
\end{aligned}
$$

El cual es una contradicción con (4.13).

Teorema 4.2.2. (Dualidad Fuerte). Si x^* es una solución eficiente para (MP) y el teorema de dualidad débil (Teorema 4.2.1) se cumple entre (MP) y (DM) y las g_i, $i\in P$, son estrictamente cuasiconvexa generalizada en x^*, entonces existen $y_i{}^*$, $i\in P$, $\lambda_i{}^*>0$, $i\in M$, tal que (x^*,λ^*,y^*) es eficiente para (DM) y los valores objetivo de (MP) y (DM) son iguales.

Prueba: Puesto que x^* es una solución eficiente para (MP).

por el teorema 4.1.1., existen escalares $\lambda_i{}^*>0$, $i\in M$ con $\sum_{i=1}^{m}\lambda_i=1$ y $y_i{}^*\geqslant 0$, $i\in I^<(x^*)$ y $d\in -N_c(x^*)$ tal que (3.5) se satisface.
Al tomar $y_i{}^*=0$ para $i\notin I^<(x^*)$, entonces $(x^*,\lambda_i{}^*,y^*)$ es factible para (DM). Supongamos que (x^*,λ^*,y^*) no es eficiente para (DM), entonces existe (u,λ,y) factible para (DM) de tal manera que

$$
\begin{aligned}
&f_i(u)+y^t g(u)>f_i(x^*)+y^{*t}g(x^*) \quad \text{para algún i}\\
&f_j(u)+y^t g(u)\geqslant f_j(x^*)+y^{*t}g(x^*) \quad \text{para todo } j\in M
\end{aligned}
\tag{4.16}
$$

Sin embargo, dado que $y^{*t}g(x^*)=0$, esto estaría en contradicción con la dualidad débil. Los valores de los objetivos de (MP) y (DM) son claramente iguales en sus respectivos puntos eficientes.□

§4.3. Dualidad Sin Calificación de las Restricciones

Consideramos el siguiente problema dual para el problema (MP) y desarrollar teoremas de dualidad sin calificación de las restricciones de (MP), donde las funciones de (MP) no se supone que son convexa y diferenciable. Considere el problema dual:

$$\begin{aligned}
\max \quad & \bar{f}(u) = (f_1(u), f_2(u), ..., f_m(u)) \\
\text{s.a:} \quad & 0 \in (\sum_{i \in M} \lambda_i \partial_c f_i(u) + \sum_{i \in P} y_i \partial_c g_i(u) + N_c(u)) \\
& g_I{}^{=}(u) = 0 \\
& y_i g_i \geqslant 0, \quad i \in P \\
& y_i \geqslant 0, \quad i = 1, 2, ..., p \\
& \lambda_i > 0, \quad i = 1, 2, ..., m \\
& \sum_{i \in M} \lambda_i = 1
\end{aligned} \tag{D2M}$$

Aquí F_{D2} denota el conjunto de soluciones factibles para (D2M).

Teorema 4.3.1. (la dualidad débil). Supongamos que $x \in F$ y $(u, \lambda, y) \in F_{D2}$. Si f_i, $i \in M$, son funciones cuasiconvexa estrictamente generalizadas y g_i, $i \in P$, son cuasiconvexa generalizadas en u, la siguiente no puede contener:

$$\begin{aligned}
f_j(x) \leqslant f_j(u) \quad & \text{Para todo } j \in M \\
f_i(x) < f_i(u) \quad & \text{Para algún } i \in M
\end{aligned} \tag{4.17}$$

Prueba: (procediendo por reducción al absurdo)

Supongamos lo contrario de los resultados que se espera en (4.6). Entonces:

$$\begin{aligned}
f_j(x) \leqslant f_j(u) \quad & \text{Para todo } j \in M \\
f_i(x) < f_i(u) \quad & \text{Para algún } i \in M
\end{aligned} \tag{4.18}$$

Por supuesto en f_i, $i \in M$, y g_i, $i \in P$, tenemos

$$\begin{aligned}
\langle \sum_{i \in M} \lambda_i \xi_i \ , \ x - u \rangle < 0 \quad & \text{Para todo } \xi_i \in \partial_c f_i(u) \\
\langle \sum_{j \in P} y_j \eta_i \ , \ x - u \rangle \leqslant 0 \quad & \text{Para todo } \eta_i \in \partial_c g_i(u)
\end{aligned} \tag{4.19}$$

Esto implica que (por propiedad del producto escalar)

$$\langle \sum_{i \in M} \lambda_i \xi_i + \sum_{j \in P} y_j \eta_i \ , \ x - u \rangle < 0 \tag{4.20}$$

para todo $\xi_i \in \partial_c f_i(u)$ y $\eta_i \in \partial_c g_i(u)$.

De las restricciones de (D2M), se deduce que para algún $d \in -N_c(u)$, $\xi_i \in \partial_c f_i(u)$ y $\eta_i \in \partial_c g_i(u)$.

$$\langle \sum_{i \in M} \lambda_i \xi_i + \sum_{j \in P} y_j \eta_i \ , \ x - u \rangle \geqslant \quad (x - u)^t d \quad \geqslant 0 \tag{4.21}$$

Esta es una contradicción con (4.20).□

Teorema 4.3.2. (dualidad fuerte). Si x^* es una solución eficiente para (MP) y el teorema de dualidad débil (Teorema 4.2.1) se mantiene entre (MP) y (D2M) y g_i, $i \in P$, son cuasiconvexa estrictamente generalizadas en x^*, entonces existen $y_i{}^*$, $i \in P$, $\lambda_i{}^* > 0$, $i \in M$, tal que (x^*, λ^*, y^*) es eficiente para (D2M) y los valores objetivo de (MP) y (D2M) son iguales.

Prueba: Puesto que x^* es eficiente para (MP), luego por el teorema 4.1.1, existen $\lambda_i{}^* > 0$,

$i \in M$,y $y_i{}^* \geqslant 0$, $i \in I^<(x^*)$, y $d \in -N_c(x^*)$ tal que se satisface (4.5). Al tomar $y_i{}^* = 0$ para $i \in I^<(x^*)$, entonces (x^*, λ^*, y^*) es factible para (D2M). Supongamos que (x^*, λ^*, y^*) no es eficiente para (D2M), entonces existe $(x, \lambda, y$ factible para (D2M) tal que

$$\begin{aligned}
f_j(x) \leqslant f_j(u) \quad & \text{Para todo } j \in M \\
f_i(x) < f_i(u) \quad & \text{Para algún } i \in M
\end{aligned} \tag{4.22}$$

Sin embargo, dado que x^* es eficiente para (MP), esto estaría en contradicción con la dualidad débil. Los objetivos de (MP) y (D2M) son iguales en sus respectivos puntos eficientes. $\square$

Conclusión

La Optimización multiobjetivo permite dar respuesta a problemas donde se necesite decidir entre varios criterios a la vez, obteniendo así una solución que satisface las necesidades del usuario. De esta forma, debido a las aplicaciones económicas de los valores duales es necesario el estudio de la teoría de dualidad multiobjetivo, mas aun sabiendo que en las mayoría de los casos reales no tenemos convexidad de las funciones y mucho menos diferenciabilidad, es por ello que los resultados asociados a los teoremas débil y fuerte de dualidad de la optimización clásica mono-objetivo debe ser generalizada.

En este trabajo se observó que bajo la hipótesis de cuasiconvexidad y no diferenciabilidad se obtuvo los equivalentes resultados de dualidad para optimización multiobjetivo.

Referencias

[1] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 1995.

[2] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory, Graduate Texts in Mathematics*, vol. 178, Springer, New York, 1998.

[3] J. B. Hiriart, C. Lemaréchal, *Convex Analysis and Minimization Algorithms I*, segunda edición, Springer, New York, 1996.

[4] M. Ehrgott, *Multicriteria Optimization*, Berlin, 2005.

[5] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms, John Wiley & Sons*, New York, 1993.

[6] , *Multiple objective programming duality without a constraint qualification*, Utilitas Mathematica 39 (1991), 41-55.

[7] O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, Journal of Mathematical Analysis and Applications 17 (1967), 37-47.

[8] P. Wolfe, *A duality theorem for non-linear programming, Quarterly of Applied Mathematics 19*, (1961), 239-244.

[9] R. R. Egudo, T. Weir, and B. Mond, *Duality without constraint qualification for multiobjectivo programming*, Journal of the Australian Mathematical Society. Series B: Applied Mathematics 33 (1992), no. 4, 531-544.

[10] S. Nobakhtian, *Duality without constraint qualification in nonsmooth optimization*, Vol 1, (2006), 1-11.

[11] T.Weir and B. Mond, *Duality for generalized convex programming without a constraint qualification*,Utilitas Mathematica 31 (1987), 233-242.

[12] V. Chankong and Y. Y. Haimes, *Multiobjective Decision Making. [Theory and Methodology]*,North-Holland Series in System Science and Engineering, vol. 8, North-Holland, New York, 1983.

[13] V. Jeyakumar, *On optimality conditions in nonsmooth inequality constrained minimization*, Numerical Functional Analysis and Optimization 9, (1987), no. 5-6, 535-546.

APÉNDICE A

El objetivo principal de esta sección es definir un conjunto de términos utilizados en el desarrollo de la tesis:

En el desarrollo del trabajo de investigación se consideró $\mathbb{R}^n$ al producto cartesiano $\overbrace{\mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R}}^{n}$, siendo $\mathbb{R} = (-\infty, +\infty)$ el conjunto de los números reales. Las diferentes coordenadas de $x \in \mathbb{R}^n$ cuando sea $n > 1$, se distinguirán mediante subíndices en el que se escribirán como $x = (x_1, x_2, ..., x_n)$. En general, se distinguirán los elementos de $\mathbb{R}^n$ mediante superíndices, así por ejemplo, $x^1, x^2, ..., x^k \in \mathbb{R}^n$, *para* $k \in \mathbb{N}$. Por defecto, se entenderá que cualquier vector $x \in \mathbb{R}^n$ está expresado como vector columna y x^t representará su traspuesta, esto es, $x^t = (x_1, x_2, ..., x_n)$. Por simplicidad, cuando $x = (x_1, x_2, ..., x_n)^t$ se utilice para $f : \mathbb{R}^n \to \mathbb{R}$ se escribirá como $f(x_1, x_2, ..., x_n)$ para denotar a $f(x) = f((x_1, x_2, ..., x_n)^t)$.

§A.1. ALGUNOS CONCEPTOS TOPOLÓGICOS

Es bien sabido que $\mathbb{R}^n$ es un espacio vectorial con las operaciones suma y producto por escalar dadas por:

- $x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n),\ con\ x, y \in \mathbb{R}^n,$
- $\alpha x = (\alpha x_1, \alpha x_2, ..., \alpha x_n),\ con\ x \in \mathbb{R}^n\ y\ \alpha \in \mathbb{R}.$

Definición A.1.1. Una Distancia definida en $\mathbb{R}^n$ es una aplicación $dist : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$ que verifica las propiedades:

1. $dist(x, y) = 0$ si, y sólo si $x = y$, para $x, y \in \mathbb{R}^n$.
2. $dist(x, y) = dist(y, x)$, para cualquier $x, y \in \mathbb{R}^n$.
3. $dist(x, y) \leqslant dist(x, z) + dist(z, y)$, para cualquier $x, y, z \in \mathbb{R}^n$.

Definición A.1.2. (Distanca Euclideana). Sean x e y puntos en $\mathbb{R}^n$, la distancia entre x e y es: $d(x, y) = \|x - y\| = \sqrt{(x - y)(x - y)}$. Si denotamos $x = (x_1, x_2, ..., x_n)^t$ y $y = (y_1, y_2, ..., y_n)^t$ entonces $\|x - y\| = [\sum_{i=1}^{n} (x_i - y_i)^2]^{1/2}$.

Definición A.1.3. Dado $x \epsilon \mathbb{R}^n$ y $\varepsilon > 0$, al conjunto $\beta(x, \varepsilon) = \{y \epsilon \mathbb{R}^n : \|y - x\| < \varepsilon\}$ se le llama una <u>Bola Abierta</u> con centro en x y radio ε. En general, una vecindad de x es cualquier conjunto $\bigvee$ tal que existe $\varepsilon > 0$ con $\beta(x, \varepsilon) \subset \bigvee$.

Definición A.1.4. Se dice que un conjunto $S \subset \mathbb{R}^n$ es un <u>Conjunto Abierto</u> si $\forall x \in S, \exists \varepsilon_x > 0$ tal que $\beta(x, \varepsilon_x) \subset S$.

Ejemplo A.1.1. 1. $\mathcal{C} = \{x \in \mathbb{R}^2 / \|x\| < 4\}$ es un conjunto abierto.

2. $\mathcal{H} = \{x \in \mathbb{R}^2/2x_1 + 3x_2 > 6\}$ es un conjunto abierto.

3. $\mathcal{K} = \{x \in \mathbb{R}^2/4 < \|x\| < 16\}$ es un conjunto abierto.

4. $\mathcal{Q} = \{(x, y) \in \mathbb{R}^2/0 < x < 1, 0 < y < 1\}$ es un conjunto abierto.

Definición A.1.5. Se dice que un conjunto $S \subset \mathbb{R}^n$ es un Conjunto Acotado si existe $M \in \mathbb{R}_+$ tal que $\|x\| \leqslant M, \forall x \in S$.

Teorema A.1.1. (Heine-Borel) Los Conjuntos Compactos de $\mathbb{R}^n$ son los Conjuntos Cerrados y Acotados.

Definición A.1.6. (Funciones Continuas). Una función $f : \mathbb{R}^n \to \mathbb{R}^m$ es continua si para cualquier abierto $A \subset \mathbb{R}^m$ se tiene que:

$$f^{-1}(A) := \{x \in \mathbb{R}^n \backslash f(x) \in A\} \textit{ es abierto}.$$

La definición anterior también puede ser enunciada en términos de conjuntos cerrados $f : \mathbb{R}^n \to \mathbb{R}^m$ es continua si para cualquier cerrado $C \subset \mathbb{R}^m$ se tiene que $f^{-1}(C)$ es también cerrado.

Teorema A.1.2. Dada una función continua $f : \mathbb{R}^n \to \mathbb{R}^m$ y un conjunto compacto $S \subset D(f)$ se tiene que $f(S) \subset \mathbb{R}^m$ es un compacto.

Teorema A.1.3. (Teorema de Weierstrass) Si $S \subset \mathbb{R}^n$ es compacto y la función $f : S \subset \mathbb{R}^n \to \mathbb{R}$ es continua en S, entonces f está acotada en S y alcanza en puntos de S los valores máximo $M = \sup_{\bar{x} \in S} f(S)$ y mínimo $m = \inf_{\bar{x} \in S} f(S)$.

Observación 9. El Teorema de Weierstrass nos garantiza, bajo sus condiciones establecidas que el problema:

$$\begin{array}{ll} \text{min (o max)} & f(x) \\ \text{s.a:} & x \in S \end{array}$$

tiene una solución óptima $\bar{x} \in S$.

Definición A.1.7. (Convergencia). Sea $(x_n) \subset \mathbb{R}$ una sucesión. Se dice que (x_n) converge a $\bar{x}$ si dado $\varepsilon > 0, \exists K_{(\varepsilon)}$ tal que para toda $n \geqslant K_{(\varepsilon)} \Rightarrow x_n \in \bigvee_{(\bar{x};\varepsilon)}$.

§A.2. ALGUNOS CONCEPTOS ANÁLISIS MATEMÁTICO

Definición A.2.1. (Función Diferenciable) Una función $f : \mathbb{R}^n \to \mathbb{R}$ es Diferenciable en un punto $a \in \mathbb{R}^n$ cuando existen las derivadas parciales $\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), ..., \frac{\partial f}{\partial x_n}(a)$ y además, para todo vector $v = (v_1, v_2, ..., v_n)^t \in \mathbb{R}^n$ tenemos que $f(a+v) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) v_i + r(v)$ con $\lim_{v \to \emptyset} \frac{r(v)}{\|v\|} = 0$.

Teorema A.2.1. Sea $d \in \mathbb{R}^n$ y $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable en $a \in \mathbb{R}^n$. Si $d.\nabla f(a) < 0$ entonces d es una dirección de descenso para f en a.

§A.3. ALGUNOS CONCEPTOS DEL ANÁLISIS CONVEXO

Definición A.3.1. Sea $x \in \mathbb{R}^n$ con $x \neq 0_n$ y $\alpha \in \mathbb{R}$, definiremos:

- Hiperplano: $H = \{y \in \mathbb{R}^n \backslash \langle x, y \rangle = \alpha\}$
- Semiespacio Abierto : $S_\alpha{}^- = \{y \in \mathbb{R}^n \backslash \langle x, y \rangle < \alpha\}$
- Semiespacio Abierto : $S_\alpha{}^+ = \{y \in \mathbb{R}^n \backslash \langle x, y \rangle > \alpha\}$

- Semiespacio Cerrado : ${S_\alpha}^- = \{y \in \mathbb{R}^n \backslash \langle x, y\rangle \leqslant \alpha\}$
- Semiespacio Cerrado : ${S_\alpha}^+ = \{y \in \mathbb{R}^n \backslash \langle x, y\rangle \geqslant \alpha\}$

Proposicion A.3.1. Los Hiperplanos y Semiespacios son Conjuntos Convexos.

Definición A.3.2. : Se define una *Combinación Afin* a toda expresión de la forma

$$\sum_{i=1}^{n} \lambda_i x_i, \quad \sum_{i=1}^{n} \lambda_i = 0$$

El conjunto de todas las combinaciones afines se llama *Cápsula Affin*.

Definición A.3.3. : Sea $C \subset \mathbb{R}^n$ un conjunto convexo. El *Interior Relativo* de C, (riC) se define por:

$$x \in riC \quad \text{si, y sólo si,} \quad x \in affC \wedge \exists\delta > 0 \quad \text{tal que} \quad (affC) \cap B(x, \delta) \subset C$$

Proposicion A.3.2. Sean $X_1, X_2, ..., X_k \subset \mathbb{R}^n$ convexos y $\alpha \in \mathbb{R}$. Entonces:

1. $X_1 \bigcap X_2 \bigcap ... \bigcap X_k$ es convexo.
2. $X^1 + X^2 + ... + X^k = \{\sum_{i=1}^{k} x^i \in \mathbb{R}^n : \ x^i \in \mathbb{R}^n, \ i = 1, 2, ..., k\}$ es convexo.
3. $\alpha X_1 = \{\alpha x \in \mathbb{R}^n, \ \forall x \in X_1\}$ es convexo.

Definición A.3.4. Un conjunto C no vacío en $\mathbb{R}^n$ y $\alpha, \lambda \in \mathbb{R}$, diremos que:

- C es un Cono No Trivial o Propio si C es un cono con $C \neq \emptyset$ y $C \neq \mathbb{R}^n$.
- C es un Cono Convexo si C es un cono y $\alpha d^1 + (1-\alpha)d^2 \in C, \ \forall d^1, d^2 \in C$ y $\forall \ 0 < \alpha < 1$.
- C es un Cono Punteado si C es un cono y para $d \in C$, $d \neq 0$, $-d \notin C$, es decir $C \bigcap \{-C\} \subseteq \{0\}$.

Definición A.3.5. Se denomina epigrafo de una función f en un conjunto $X \subset \mathbb{R}^n$, al conjunto:

$$epi(f) = \{(\bar{x}, z) \in X \times \mathbb{R} : f(\bar{x}) \leqslant z\}.$$

Geométricamente, esto se interpreta como el conjunto de los puntos $(\bar{x}, z) \in X \times \mathbb{R}$ que están sobre el gráfico de f (recordemos que el gráfico de una función f es el conjunto $\{(\bar{x}, z) \in X \times \mathbb{R} : f(\bar{x}) = z\}$).

Proposicion A.3.3. Una función $f(\cdot) : C \subset \mathbb{R}^n \to \mathbb{R}$ definida sobre C convexo es convexa si, y solamente si el $epi(f)$ es un conjunto convexo.

Como todo conjunto convexo, el epígrafo de una función convexa admite hiperplanos soporte en cualquier punto de su frontera.

Proposicion A.3.4. Sea $\emptyset \neq C \subset \mathbb{R}^n$ un abierto convexo y $\alpha \in \mathbb{R}$. Se tiene:

1. Si $f, g : C \to \mathbb{R}$ son convexas en C, entonces $f + g$ es convexa en C.
2. Si $f : C \to \mathbb{R}$ es convexas en C y $\alpha \geqslant 0$, entonces αf es convexa en C.

Proposicion A.3.5 (Proposición 4.13). Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función Lipschitz en una vecindad de x con constante de Lipschitz K, con $x \in dom(f)$. Entonces, la función $v \to f_c(x; v)$ con dominio $\mathbb{R}^n$ tiene las siguientes propiedades:

1. $f_c(x, v)$ es limitada por la constante de Lipschitz K;
2. La derivada direccional generalizada es homogénea positiva, esto es, $f_c(x, \lambda v) = \lambda f_c(x; v)\ y\ \lambda > 0$;
3. $f_c(x, v)$ es convexa;
4. $f_c(x, -v) = (-f)_c(x, v)$.

Teorema A.3.1. Sea A, B $\subset \mathbb{R}^n$ conjuntos convexos y compactos. Entonces:

$$min_{x \in A} max_{y \in B} x.y = max_{y \in B} min_{x \in A} x.y$$

§A.4. PROBLEMAS IRRESTRICTOS

Definición A.4.1. Considere el Problema:

$$Min\ f(x)$$

$$x \in \mathbb{R}^n$$

y sea $\bar{x} \in \mathbb{R}^n$. Si $f(\bar{x}) \leqslant f(x),\ \forall x \in \mathbb{R}^n$ entonces $\bar{x}$ es un mínimo global.
Si existe $V_\varepsilon(\bar{x}) = \{x \in \mathbb{R}^n : \|x - \bar{x}\| < \varepsilon\}$ tal que $f(\bar{x}) \leqslant f(x),\ \forall x \in V_\varepsilon(\bar{x})$ entonces $\bar{x}$ es un mínimo local.
Claramente todo mínimo global es mínimo local.

§A.4.1. Condición Necesaria de Optimalidad

Dado un punto $\bar{x} \in \mathbb{R}^n$, se requiere determinar en esta sección la caracterización que nos permita deducir que este punto es o no un mínimo local o global de una función f diferenciable.

Condición Necesaria de Primer Orden

Teorema A.4.1. Supongamos que f es diferenciable en $\bar{x}$. Si existe d tal que $\nabla f(\bar{x})^t d < 0$, entonces existe $\delta > 0$ tal que $f(\bar{x} + \lambda d) < f(\bar{x})$ para cada $\lambda \in (0, \delta)$, así d es una dirección de descenso de f en $\bar{x}$.

Prueba 1. Como f es diferenciable en $\bar{x}$ tenemos por expansión de Taylor:

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})^t d + \lambda \|d\| \alpha(\bar{x}, \lambda d)$$

Reorganizando los términos

$$f(\bar{x} + \lambda d) - f(\bar{x}) = \lambda \nabla f(\bar{x})^t d + \lambda \|d\| \alpha(\bar{x}, \lambda d)$$

Dividiendo entre $\lambda > 0$ se tiene:

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})^t d + \|d\| \alpha(\bar{x}, \lambda d)$$

Como $\nabla f(\bar{x})^t < 0$ y $\alpha(\bar{x}, \lambda d) \to 0$ cuando $\lambda \to 0$, existe $\delta > 0$ tal que

$$\nabla f(\bar{x})^t d + \|d\| \alpha(\bar{x}, \lambda d) < 0 \quad \forall \lambda \in (0, \delta)$$

$$\forall \varepsilon > 0\ \exists \delta > 0\ \mid \lambda \mid < \delta \Rightarrow \mid \alpha(\bar{x}, \lambda d) \mid < \varepsilon$$

Así

$$\frac{f(\bar{x}+\lambda d)-f(\bar{x})}{\lambda}<0$$

y por tanto $f(\bar{x}+\lambda d)<f(\bar{x}), \forall\lambda \in (0,\delta)$

Geométricamente tenemos:

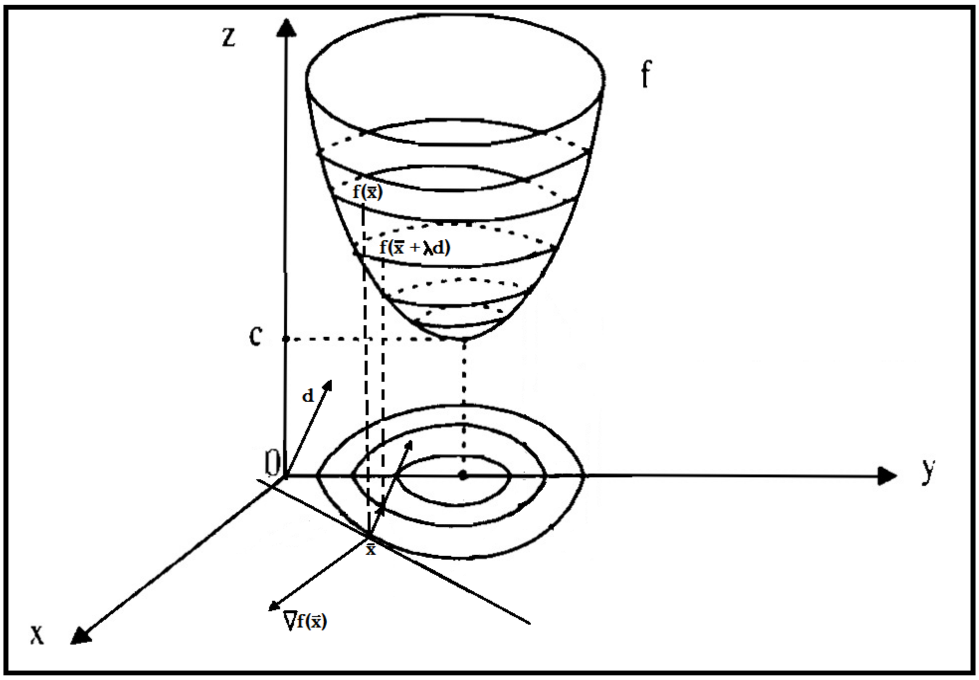

FIGURA A.4.1: Condición Necesaria de Primer Orden

El siguiente corolario del teorema anterior da una 1^{era} condición necesaria para que $\bar{x}$ sea un óptimo local

Corolario A.4.1. Supongamos que $f:\mathbb{R}^n \to R$ es diferenciable en $\bar{x}$. Si $\bar{x}$ es un mínimo local, entonces $\nabla f(\bar{x})=0$.

Prueba 2. Procediendo por reducción al absurdo
Supongamos que $\nabla f(\bar{x}) \neq 0$, considerando $\nabla f(\bar{x})>0$ y sea $d=-\nabla f(\bar{x})$ entonces $\nabla f(\bar{x})^t.d<0$ y así por el teorema anterior existe $\delta>0$ tal que $f(\bar{x}+\lambda d)<f(\bar{x}), \forall\lambda \in (0,\delta)$
Contradiciendo el hecho de que $\bar{x}$ es mínimo local.
Por lo tanto, $\nabla f(\bar{x})=0$.

Condición Necesaria de Segundo Orden

Definición A.4.2. Una matriz A de orden $N\times N$ es:

i) Semidefinida Positiva si $x^tAx \geqslant 0\ \forall x \in \mathbb{R}^n$.

ii) Definida Positiva si $x^tAx>0\ \forall x \in \mathbb{R}^n,\ x \neq 0$.

iii) Indefinida si tiene el eigenvalor positivo y negativo.

iv) SDP si es Simétrico y Definido Positivo.

Teorema A.4.2. Supongamos que $f : \mathbb{R}^n \to \mathbb{R}$ es 2-veces continuamente diferenciable en $\bar{x}$. Si $\bar{x}$ es mínimo local entonces $\nabla f(\bar{x}) = 0$ y $f(\bar{x})$ es semidefinida positiva.

Prueba 3. Por el corolario anterior se cumple que $\nabla f(\bar{x}) = 0$, sólo falta probar que $H(\bar{x})$ es semidefinida positiva.

Sea d una dirección arbitraria y dado que $f \in C^2$ entonces por la expansión de Taylor se tiene:

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})^t d + \frac{1}{2}\lambda^2 d^t H(\bar{x}) d + \lambda^2 \|d\|^2 \alpha(\bar{x}, \lambda d)$$

Reorganizando los términos

$$f(\bar{x} + \lambda d) - f(\bar{x}) = \lambda \nabla f(\bar{x})^t d + \frac{1}{2}\lambda^2 d^t H(\bar{x}) d + \lambda^2 \|d\|^2 \alpha(\bar{x}, \lambda d)$$

Dividiendo por $\lambda^2 > 0$ en ambos lado de la igualdad:

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda^2} = \frac{\nabla f(\bar{x})^t d}{\lambda} + \frac{1}{2} d^t H(\bar{x}) d + \lambda^2 \|d\|^2 \alpha(\bar{x}, \lambda d)$$

Dado que $\nabla f(\bar{x}) = 0$ entonces $\frac{\nabla f(\bar{x})^t d}{\lambda} = 0$; y como $\alpha(\bar{x}, \lambda d) \to 0$ cuando $\lambda \to 0$ entonces

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda^2} = \frac{1}{2} d^t H(\bar{x}) d$$

Como $\bar{x}$ es un mínimo local (por hipótesis) entonces $f(\bar{x} + \lambda d) \geqslant f(\bar{x})$ para $\lambda > 0$ suficientemente pequeño. Y así $f(\bar{x} + \lambda d) - f(\bar{x}) \geqslant 0$. obteniéndose que $\frac{1}{2} d^t H(\bar{x}) d \geqslant 0$ con d arbitrario. Luego, $d^t H(\bar{x}) d \geqslant 0$ para cualquier d.

Por lo tanto, $H(\bar{x})$ es semidefinida positiva.

§A.4.2. Condición Suficiente de Optimalidad

Teorema A.4.3. Suponga que $f : \mathbb{R}^n \to \mathbb{R}$ es 2-veces diferenciable en $\bar{x}$. Si $\nabla f(\bar{x}) = 0$ y $f(\bar{x})$ es definida positiva entonces $\bar{x}$ es un mínimo local estricto.

Prueba 4. Como f es dos veces diferenciable en $\bar{x}$, tenemos que para cada $\bar{x} \in \mathbb{R}^n$ la expansión de Taylor:

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^t (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^t H(\bar{x})(x - \bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}, x - \bar{x})$$

donde $\alpha(\bar{x}, x - \bar{x}) \to 0$ cuando $x \to \bar{x}$.

Supongamos por reducción al absurdo que $\bar{x}$ no es un mínimo local estricto; esto es, suponga que existe una sucesión $\{x_k\}$ convergiendo a $\bar{x}$ tal que $f(x_k) \leqslant f(\bar{x})$, $x_k \neq \bar{x}$ para cada k.

Considerando esta secuencia, notando que $\nabla f(\bar{x} = 0$ y que $f(x_k) \leqslant f(\bar{x})$; y haciendo $d_k = \frac{x_k - \bar{x}}{\|x_k - \bar{x}\|}$ una sucesión de vectores directores unitarios, entonces implica que

$$\frac{1}{2} {d_k}^t H(\bar{x}) d_k + \alpha(\bar{x}, x_k - \bar{x}) \ para\ cada\ k$$

Pero $\|d_k\| = 1$ para cada k, y por lo tanto existe un conjunto de índice $\mathcal{W}$ tal que $\{d_k\}_{\mathcal{W}}$ convergen a d, donde $\|d\| = 1$.

Considerando esta subsecuencia y el hecho de que $\alpha(\bar{x}, x_k - \bar{x}) \to 0$ cuando $k \in \mathcal{W}$ acercándose a ∞, implica que $d^t H(\bar{x}) d \leqslant 0$.

Esto contradice lo supuesto que $H(\bar{x})$ es definida positiva ya que $\|d\| = 1$.

Por lo tanto, $\bar{x}$ es un mínimo local estricto.

Ejemplo A.4.1. *Min* $f(x) = (x^2 - 1)^3$
Calculo del gradiente de f: $\nabla f(x) = 6x(x^2 - 1)^2$
Calculo de puntos críticos: $\nabla f(x) = 0 \Rightarrow x = 0,\ x = -1,\ x = 1$
Calculo del Hessiano: $H(x) = 24x^2(x^2 - 1) + 6(x^2 - 1)^2 = 6(x^2 - 1)(5x^2 - 1)$
De done se obtiene que: $H(1) = 0 = H(-1)$ semidefinida positiva y $H(0) = 6$.

§A.5. PROBLEMAS CON RESTRICCIONES DE DESIGUALDAD

Considere el Problema:

$$Min\ f(x)$$
$$s.a:\ x \in S$$

Definición A.5.1. Sea S un conjunto no vacío en $\mathbb{R}^n$, y sea $\bar{x} \in clS$. El cono de direcciones factibles de S en $\bar{x}$ denotado por $\mathbb{D}$, está dado por

$$\mathbb{D} = \{d : d \neq 0 \quad y \quad \bar{x} + \lambda d \in S, \quad \forall \lambda \in (0, \delta)\ para\ algun\ \delta > 0\}$$

Cada vector $d \in \mathbb{D}$ es llamado dirección factible. Además, dada una función $f : \mathbb{R}^n \to \mathbb{R}$, el cono de direcciones de mejora en $\bar{x}$, denotado por $\mathcal{F}$, es dado por

$$\mathcal{F} = \{d : f(\bar{x} + \lambda d) < f(\bar{x}) \quad \forall \lambda \in (0, \delta)\ para\ algun\ \delta > 0\}.$$

Cada dirección $d \in \mathcal{F}$ es llamada una dirección de mejora, o una dirección descendente, de f en $\bar{x}$.

El siguiente teorema nos dice que toda dirección de mejora de una función no es factible.

Teorema A.5.1. Considere el Problema:

$$Min\ f(x)$$
$$s.a:\ x \in S$$

Siendo $f : \mathbb{R}^n \to \mathbb{R}$ y S un conjunto no vacío en $\mathbb{R}^n$. Supongamos que f es diferenciable en un punto $\bar{x} \in S$. Si $\bar{x}$ es una solución óptima local, entonces $\mathcal{F}_0 \bigcap \mathbb{D} = \emptyset$ donde $\mathcal{F}_0 = \{d : \nabla f(\bar{x})^t d < 0\}$ y $\mathbb{D}$ es el cono de direcciones factibles de S en $\bar{x}$.

Geométricamente se tiene (Ver FIGURA 1.2.1)

Prueba 5. Procedemos por reducción al absurdo.
Supongamos que existe un vector $d \in \mathcal{F}_0 \bigcap \mathbb{D}$, entonces $d \in \mathcal{F}_0$ y $d \in \mathbb{D}$.
Esto es,
existe $\delta_1 > 0$ tal que $f(\bar{x} + \lambda d) < f(\bar{x}) \quad \forall \lambda \in (0, \delta_1)$
y existe $\delta_2 > 0$ tal que $\bar{x} + \lambda d \in S \quad \forall \lambda \in (0, \delta_2)$
tomando $\delta = min\{\delta_1, \delta_2\}$ se tiene
$f(\bar{x} + \lambda d) < f(\bar{x}) \quad y \quad \bar{x} + \lambda d \in S \quad \forall \lambda \in (0, \delta)$,
Lo cual contradice el hecho de que $\bar{x}$ es un mínimo local ya que $\forall N_\lambda(\bar{x})$ se tiene $f(\bar{x} + \lambda d) < f(\bar{x})$
Por lo tanto, $\mathcal{F}_0 \bigcap \mathbb{D} = \emptyset$

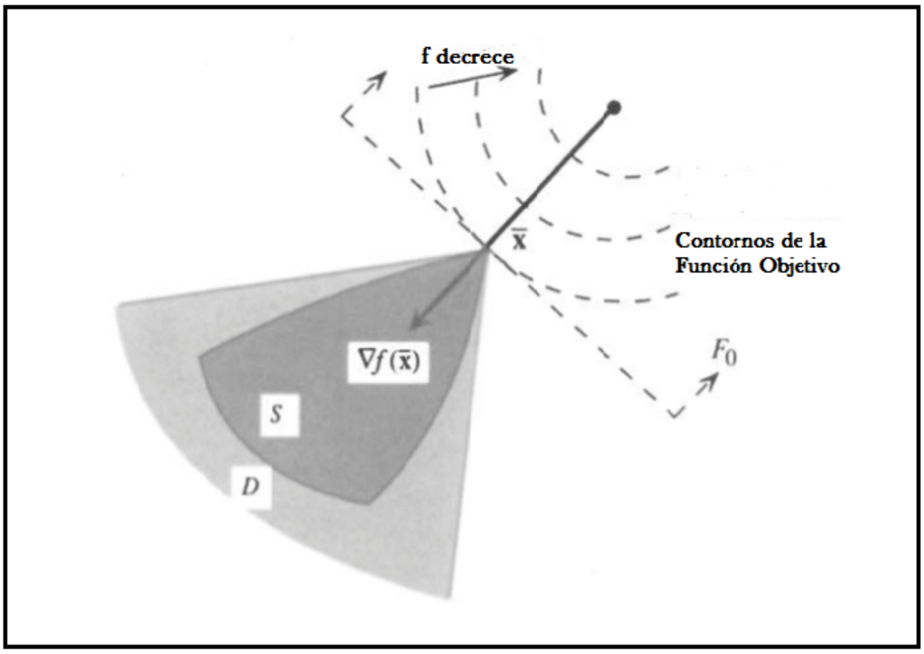

FIGURA A.5.1: Condición Necesaria $\mathcal{F}_0 \bigcap \mathbb{D} = \emptyset$

Considere el siguiente problema:

$$\begin{array}{ll} \min & f(x) \\ \text{s.a:} & g_i(x) \leqslant 0 \quad \forall i = 1,2,...,m \\ & x \in X \end{array}$$

$S = \{x \in X : g_i(x) \leqslant 0 \quad \forall i = 1,2,...,m\}$

Sabemos que las condiciones de optimalidad se cumple cuando $\mathcal{F}_0 \bigcap \mathbb{D} = \emptyset$.

Ahora $\mathcal{G}_0 \subset \mathbb{D} \Rightarrow \mathcal{G}_0 \bigcap \mathcal{F}_0 = \emptyset$.

Donde $\mathcal{G}_0$ es el conjunto de direcciones de descenso de los vectores activos.

Teorema A.5.2. Sea $g_i : \mathbb{R}^n \to \mathbb{R}$ para $i = 1,2,...,m$, y sea X un conjunto abierto no vacío en $\mathbb{R}^n$

Considere el siguiente problema (P):

$$\begin{array}{ll} \min & f(x) \\ \text{s.a:} & g_i(x) \leqslant 0 \quad \forall i = 1,2,...,m \\ & x \in X \end{array}$$

Sea $\bar{x}$ un punto factible y sea $I = \{i : g_i(\bar{x}) = 0\}$. Más aún, suponga que f y g_i $\forall i \in I$ son diferenciables en $\bar{x}$ y que para $i- \in I$ es continua en $\bar{x}$.

Si $\bar{x}$ es una solución óptima local entonces $\mathcal{F}_0 \bigcap \mathcal{G}_0 = \emptyset$ donde

$\mathcal{F}_0 = \{d : \nabla\mathcal{F}(\bar{x})^t d < 0\}$

$\mathcal{G}_0 = \{d : \nabla\}_i(\bar{x})^t d < 0, \quad \forall i \in I\}$

"Dado que la forma del Universo entero es la más perfecta y, de hecho, la más sabiamente creada, absolutamente nada en el mundo ocurriría sin que una minimización o maximización esté actuando".

Leonhard Euler (1744)