UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

ESCUELA DE POSGRADO

FACULTAD DE CIENCIAS MATEMÁTICAS

UNIDAD DE POSGRADO

"OPTIMIZACIÓN EN ESPACIOS DE BANACH Y APLICACIONES"

PARA OBTENER EL GRADO ACADÉMICO DE MAGISTER EN MATEMÁTICA APLICADA

MENCIÓN MATEMÁTICA COMPUTACIONAL

AUTOR

Milton Angelino Aycho Flores

Lima – Perú

2015

Optimización en espacios de Banach y aplicaciones

por

Milton Angelino Aycho Flores

Tesis presentada a consideración del Cuerpo Docente de la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para obtener el Grado Académico de Magíster en Matemática Aplicada.

Aprobada por el siguiente jurado:

Dr. Alfonso Pérez Salvatierra
Presidente

Dr. Luis Enrique Carrillo Díaz
Miembro

Dr. Renato Mario Benazic Tome
Miembro

Dra. Roxana López Cruz
Miembro

Mg. Tomás Alberto Núñez Lay
Miembro Asesor

Ficha Catalográfica

AYCHO FLORES, Milton Angelino

Título: *Optimización en espacios de Banach y aplicaciones*, (Lima) 2015 `ix.,91p.,29.7cm (UNMSM, Magíster, Matemática Aplicada, 2015)`

Tesis, Universidad Nacional Mayor de San Marcos. Escuela de Posgrado. Facultad de Ciencias Matemáticas. Unidad de Posgrado.

`1. Matemática Aplicada 2. Optimización`
`I. UNMSM/EPG/FCM/UPG II. Magíster (Serie)`

Dedicatoria

A mi madre Brígida, por su
cariño, ejemplo y motivación.
A mi familia, por su
inmenso apoyo, fé
y confianza.

A Jossita, mi futura compañera,
por que eres mi gran amor,
mi hermosa y bella inspiración.

(In memoriam)
A Don Aurelio y Fredy Aycho,
mi padre y hermano, que descansan
en el infinito cielo.
Les dedico con profundo
cariño este trabajo.

Agradecimientos

- A Dios, por haber logrado mi objetivo y protegerme con su infinita bendición. A ti Virgen de Guadalupe, Madre bendita y milagrosa, gracias por tu protección.

- Al miembro asesor, Mg. Tomás Alberto Núñez Lay, por su paciencia y dedicación en las rutinas de exposición, por la orientación y observaciones en la elaboración y redacción de este trabajo, para él mi profundo agradecimiento.

- A los señores miembros del jurado calificador: Dra. Roxana López Cruz, Dr. Luis Enrique Carrillo Díaz, Dr. Renato Benazic Tome y Dr. Alfonso Pérez Salvatierra, por sus sugerencias y recomendaciones en la redacción final de este trabajo.

- A mi hermana Susana, mi sobrina Samantha (*Achuzi*) y Robert (*Bobby*), por su valioso y gran apoyo durante la redacción digital de esta tesis. A mi hermano Wilman, mis hermanos y hermanas, por su apoyo y confianza durante estos años de estudios.

- A mi estimado amigo y compañero de estudios Mg. Frank Navarrro Rojas, por su acertado consejo para iniciar los estudios de esta maestría. Al ingeniero Percy Triveño Aucahuasi, alumno de la Unidad de Posgrado - FCM, por su valiosa recomendación en la edición digital de este trabajo.

- A mis amigos y colegas Lic. Jesús Luque Rivera, Luis Huacausi, Rosario Taipe y los señores licenciados: Julio Quesada, Alex Cruz Huallpara, Joel Rojas, César Rayme, Adrián Allauca, Milagros Carhuapoma López y Rocio Caja Rivera, con quienes compartí una agradable experiencia en las aulas y aportaron su granito de arena para el desarrollo de esta Tesis de Maestría.

- A cada uno de mis profesores: Dr. Luis Enrique Carrillo Díaz, Dra. Roxana López Cruz, Dr. Efraín Carbajal Peña, Dr. Renato Benazic Tome, Dr. Erik Papa Quiroz, Mg. Tomás Núñez Lay, Mg. Jorge Condado, Mg. Johnny Avendaño y Mg. Augusto Cortez, por haber hecho de estos estudios de maestría, una experiencia gratificante, única e invalorable.

Resumen

Optimización en espacios de Banach y aplicaciones

Milton Angelino Aycho Flores
Octubre, 2015

Asesor	:	Mg. Tomás Alberto Núñez Lay
Grado Obtenido	:	Magíster en Matemática Aplicada

En este trabajo se estudia el problema de optimización

$$\min_{x \in S} f(x)$$

donde S es un subconjunto convexo en un espacio normado X y $f : X \longrightarrow \mathbb{R}$ es una funcional. Asimismo, se presenta una extensión del teorema de Kuhn-Tucker que resuelve el problema de minimización sobre el conjunto $\hat{S} = \{x \in S / g(x) \in -C \ \wedge \ h(x) = 0_Z\}$ donde C es un cono de orden y h, g dos funcionales Fréchet diferenciables.

Palabras Clave:

Espacios de Banach

Funciones Convexas

Problema de Optimización

Teorema de Kuhn Tucker Generalizado

Subgradiente

Abstract

Optimization in Banach Spaces and Applications

Milton Angelino Aycho Flores
October, 2015

Adviser	:	Mg. Tomás Alberto Núñez Lay
Obtained Degree	:	Master in Applied Mathematics

In this work, we study the optimization problem

$$\min_{x \in S} f(x)$$

where S is a convex subset in a normed space X and $f : X \longrightarrow \mathbb{R}$ is a functional. It is also presented an extension of Kuhn Tucker theorem which solves the minimization problem in set $\hat{S} = \{x \in S / g(x) \in -C \ \wedge \ h(x) = 0_Z\}$, where C is a order cone and h, g are two Fréchet differentiable functionals.

Keywords:

Banach Spaces

Convex Functions

Optimization Problem

Generalized Kuhn Tucker Theorem

Subgradient

Índice general

Introducción

En la teoría de optimización clásica, dada una función diferenciable $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, la condición necesaria para encontrar el punto óptimo de f en D es dar solución a la ecuación

$$f'(x) = 0$$

donde $f'(x)$ es la derivada de la función, en el sentido habitual. En el caso de una función real de varias variables reales, se emplean los métodos del gradiente y la matriz Hessiana para resolver los problemas de optimización sin restricciones; además el uso de la función Lagrangiana y los multiplicadores de Lagrange para el caso de la optimización con restricciones de igualdad.

En 1951, los matemáticos H. W. Kuhn y A. W. Tucker estudian el problema de optimización sobre conjuntos convexos en el espacio $\mathbb{R}^n$, estableciendo condiciones necesarias y suficientes para el problema de optimización

$$\min_{x \in S} f(x)$$

donde $S = \{x \in \mathbb{R}^n / g(x) \leq 0\}$ siendo $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ y $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dos funciones suficientemente diferenciables. Desde entonces dicha técnica recibe el nombre de método de Kuhn-Tucker.

A partir de ese momento, muchos investigadores han estudiado la existencia de solución del problema de minimización

$$\min_{x \in X} f(x)$$

donde X es un espacio normado y f es una funcional real.

La necesidad de esta extensión, se debe a que la mayoría de modelos matemáticos de la vida real, en especial los modelos de aplicación hacia otras ramas de las ciencias exactas y aplicadas, trabajan en espacios que no son necesariamente el conjunto $\mathbb{R}$ o en general como los que aparecen en la investigación de operaciones como es el caso del espacio $\mathbb{R}^n$.

P. Varaiya [16], en 1967 trabaja el problema de maximización

$$\max_{x \in S} f(x) \tag{1}$$

donde $S = \{x \in A / g(x) \in A_Y\} \subseteq X$, siendo X, Y espacios de Banach, A un conjunto no vacío en X, A_Y un subconjunto convexo no vacío de Y; $g : X \longrightarrow Y$ una función Fréchet diferenciable y $f : X \longrightarrow \mathbb{R}$ una función diferenciable. En su trabajo, encuentra condiciones necesarias y suficientes para el punto solución de (1).

Años más tarde, M. Guignard [5] estudia el problema de optimización (1) sobre un espacio de Banach, para una función dos veces Fréchet diferenciable bajo condiciones de pseudoconvexidad y pseudoconcavidad sobre las condiciones calificadas, logrando una generalización de las condiciones de Kuhn-Tucker en un espacio abstracto.

En 1970, los trabajos de K. Ritter [14], estudian y presentan el fundamento teórico que permite la generalización de la teoría de optimización en espacios normados para el caso de funcionales con valores vectoriales, es decir $f : X \longrightarrow Y$ donde Y no es necesariamente el conjunto $\mathbb{R}$. En estos trabajos se demuestran resultados para las condiciones de optimalidad y las de calificación, además hace un estudio minucioso del Análisis Convexo desde el punto de vista topológico para un espacio normado.

Uno de los primeros investigadores en estudiar el teorema generalizado de los multiplicadores de Lagrange fue R. Wheatherwax [18], quien en 1974 enfoca el problema de programación no lineal en un espacio de Banach, bajo condiciones calificadas definidas sobre conos tangentes, obteniendo condiciones suficientes para problemas de minimización restringidos, con la hipótesis de Fréchet diferenciabilidad en la función objetivo.

En la década de los 80, muchos investigadores estudian el problema de minimización de funcionales reales Fréchet diferenciables sobre un espacio de Banach, para el caso irrestricto y bajo restricciones sobre conos convexos. En particular, Jahn [7] (1986), considera los casos en el cual la funcional objetivo es cuasidiferenciable y Fréchet diferenciable, además estudia con mayor detalle el problema del funcional Lagrangiano en la optimización con restricciones en espacios de Banach.

En los últimos veinte años, los estudios sobre problemas de optimización en espacios normados, tienen un gran auge, gracias a los resultados e investigaciones dentro del campo del Análisis Funcional Convexo, permitiendo así dar nuevas generalizaciones a muchos conceptos del análisis convexo clásico en $\mathbb{R}^n$, que a su vez concluyeron en resultados que resuelven los problemas de optimización abstracta y en particular sobre los problemas de optimización con restricciones, logrando extender el teorema de Kuhn-Tucker en espacios topológicos localmente convexos. Los trabajos de Kazmi [10] y Yu-Lui [19] son algunas referencias de la investigación en la Optimización sobre espacios normados.

Para abordar el problema de la optimización en un espacio de Banach X, se emplean conceptos de diferenciabilidad más generales que la definición de derivada clásica, así como también, resultados específicos del Análisis Convexo y del Análisis Funcional en dimensión infinita, para lograr la formulación de los problemas de optimización en el espacio abstracto X.

Con estos resultados, se pueden tener las herramientas matemáticas que permiten resolver los problemas de optimización global y con restricciones planteados en espacios normados y de Banach que no son necesariamente los euclidianos ($\mathbb{R}$ o $\mathbb{R}^n$), como por ejemplo, tratar el problema de optimización de funcionales definidos sobre el espacio de Banach $C^1([a,b])$, el cual es un caso especial de la Programación Dinámica, conocido como Cálculo Variacional o Cálculo de Variaciones.

En este trabajo, se estudia la existencia de solución para el problema de optimización

$$\min_{x\in S} f(x) \tag{2}$$

donde S es un conjunto convexo en un espacio de Banach X y $f : X \longrightarrow \mathbb{R}$ es una funcional Fréchet diferenciable. Asimismo se demostraran diversos resultados de existencia de solución para el problema (2) para el caso global (cuando $S = X$), el caso restringido y bajo condiciones calificadas (conjunto de restricciones), además presentaremos resultados sobre condiciones necesarias y suficientes para el problema (2) cuando f es una funcional no necesariamente diferenciable.

También, se demuestra una generalización del teorema de multiplicadores de Lagrange dentro de un espacio de Banach. Este problema de optimización con condiciones generalizadas, viene siendo utilizado en problemas relacionados con la física, mecánica, ingeniería, economía, logística, administración de operaciones, etc. Para el caso de un espacio de Banach, este resultado es conocido como el teorema de Karush-Kuhn-Tucker.

Capítulo 1

Preliminares

En este capítulo desarrollaremos los conceptos, teoremas y resultados preliminares que serán utilizados a lo largo de este trabajo. Comenzaremos con los conceptos de espacio vectorial normado, el espacio de funciones lineales continuas, espacios de Hilbert y Banach.

1.1. Espacios vectoriales normados

Definición 1.1. Sean X un espacio vectorial real y $\| \, \| : X \longrightarrow \mathbb{R}$ una aplicación. Diremos que $\| \, \|$ es una norma en X si:

(i) $\|x\| \geq 0$ para todo $x \in X$ y $\|x\| = 0$ si y sólo si $x = 0$.

(ii) $\|\alpha x\| = |\alpha| \, \|x\|$ para todo $x \in X$ y $\alpha \in \mathbb{R}$.

(iii) $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in X$.

Si X es un espacio vectorial con una norma $\| \, \|$, entonces el par $(X, \| \, \|)$ se denomina *espacio vectorial normado*, el cual abreviaremos con las siglas e.v.n.

Definición 1.2. Sean $(X, \| \, \|)$ un e.v.n, $x \in X$ y $r > 0$. Definimos como bola abierta (bola cerrada, respectivamente) a los conjuntos:

$$B(x, r) = \{y \in X / \ \|y - x\| < r\}$$
$$\bar{B}(x, r) = \{y \in X / \ \|y - x\| \leq r\}$$

Definición 1.3. Sea A un subconjunto de X.

(i) Decimos que el punto $a \in A$ es punto interior de A, si existe $\delta > 0$ (que depende de a) tal que $B(a, \delta) \subset A$. Al conjunto de todos los puntos interiores de A se denomina interior de A y será denotado por $\overset{\circ}{A}$.

(ii) Decimos que el punto $b \in X$ es punto adherente de A si dado $\epsilon > 0$ se cumple que $B(b,\epsilon) \cap A \neq \emptyset$. Al conjunto de todos los puntos adherentes de A se denomina adherencia o cerradura de A y se denota por $\bar{A}$.

(iii) El conjunto frontera de A será denotado por ∂A y es definido como $\partial A = \bar{A} \setminus \overset{\circ}{A}$.

Definición 1.4. Sean $(X, \| \ \|)$ un e.v.n, $\{x_n\}_{n\in\mathbb{N}}$ una sucesión en X y un punto $x \in X$. Decimos que la sucesión $\{x_n\}$ *converge fuertemente o converge en norma* $\| \ \|$ a x si

$$\lim_{n\to\infty} \|x_n - x\| = 0$$

En tal caso, denotamos dicha convergencia por $x_n \longrightarrow x$. Diremos que una sucesión $\{x_n\}_{n\in\mathbb{N}}$ es de Cauchy en X si $\lim\limits_{n,m\to\infty} \|x_n - x_m\| = 0$.

El siguiente lema, caracteriza los puntos interiores y adherentes en términos de convergencia de sucesiones.

Lema 1.5. Sean $(X, \| \ \|)$ un e.v.n y A un subconjunto de X, se cumplen:

(i) $a \in A$ es un punto interior de A si y sólo si para cada sucesión $\{x_n\}_{n\in\mathbb{N}}$ convergente hacia el punto a, existe $n_0 \in \mathbb{N}$ tal que $x_n \in A$ para todo $n \geq n_0$.

(ii) $a \in X$ es punto adherente de A si y sólo si existe una sucesión $\{x_n\}_{n\in\mathbb{N}} \subseteq A$ tal que $x_n \longrightarrow a$.

Demostración. Ver Jost [9] (*pag. 86, Teorema 7.28*).

□

Definición 1.6. Sean $(X, \| \ \|)$ un e.v.n y A un subconjunto de X.

(i) Decimos que el conjunto A es abierto si $A = \overset{\circ}{A}$, es decir todos los puntos de A son puntos interiores.

(ii) Decimos que el conjunto A es cerrado si $A = \bar{A}$, es decir todos los puntos de A son puntos de adherencia.

Definición 1.7. Sea $(X, \| \ \|)$ un e.v.n. Decimos que X es un *espacio de Banach* si toda sucesión de Cauchy en X admite un punto de convergencia x en X. Es decir si $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ es una sucesión de Cauchy, entonces existe $x \in X$ tal que $x_n \longrightarrow x$.

Ejemplos

- Consideremos $X = \mathbb{R}^n$ dotado de las normas:

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|$$
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$
$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

para cada $x \in \mathbb{R}^n$. Se tiene que los espacios $(X, \|\ \|_\infty)$, $(X, \|\ \|_1)$ y $(X, \|\ \|_2)$ son espacios de Banach.

- Sea $X = C[a,b] = \{f : [a,b] \longrightarrow \mathbb{R} / f$ es continua en $[a,b]\}$, dotado con la norma

$$\|f\|_{C[a,b]} = \max_{x\in[a,b]} |f(x)|$$

para cada $f \in C[a,b]$. $(X, \|\ \|_{C[a,b]})$ es un espacio de Banach.

- Como una extensión del ejemplo anterior, consideremos $X = C^1[a,b]$ al espacio vectorial de las funciones reales continuas definidas en $[a,b]$ tales que f' es continua en $[a,b]$. Este espacio, equipado con la norma

$$\|f\|_{C^1[a,b]} = \max_{x\in[a,b]} \left\{|f(x)| + \left|f'(x)\right|\right\}$$

es un espacio de Banach.

Definición 1.8. Sea X un espacio vectorial real. Consideremos $(\cdot,\cdot) : X \times X \longrightarrow \mathbb{R}$ una aplicación. Decimos que $(\cdot,\cdot)$ es un *producto interno* sobre X si se cumplen:

(*i*) $(x,x) \ge 0$ para todo $x \in X$ y $(x,x) = 0$ si y sólo si $x = 0$.

(*ii*) $(x, \alpha y + \beta z) = \alpha (x,y) + \beta (x,z)$ para todo $x,y,z \in X$ y $\alpha, \beta \in \mathbb{R}$.

(*iii*) $(x,y) = (y,x)$ para todo $x,y \in X$.

Un espacio vectorial X con producto interno $(\cdot,\cdot)$ se denomina espacio *pre-Hilbertiano* y será denotado por $(X, (\ ,\ ))$, donde $(\ ,\ )$ es el producto interno definido en X.

Definición 1.9. Sea $(X, (\ ,\ ))$ un espacio Pre-Hilbertiano. Definimos la norma asociada al producto interno $\|\ \|_{(\ ,\ )} : X \longrightarrow \mathbb{R}$ por

$$\|x\|_{(\ ,\ )} = \sqrt{(x,x)}$$

para todo $x \in X$. Con esto $(X, \|\ \|_{(\ ,\ )})$ es un espacio vectorial normado.

Decimos que $(X, \| \|_{(\,,\,)})$ es un espacio de Hilbert, si el espacio normado $(X, \| \|_{(\,,\,)})$ es un espacio de Banach.

Ejemplos.

- Sea $X = \mathbb{R}^n$ dotado del producto interno $(x, y) = \sum_{i=1}^{n} x_i y_i$ y la norma inducida por dicho producto, $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$, constituyen un espacio de Hilbert.

- Consideremos el espacio vectorial $X = \left\{ (x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R} / \sum_{n=1}^{\infty} x_n^2 < \infty \right\}$, denotado por $l_2(\mathbb{R})$ con el producto interno $(x, y) = \sum_{n=1}^{\infty} x_n y_n$ y la norma asociada a este producto interno $\|x\|_{l_2} = \sqrt{\sum_{i=1}^{\infty} x_i^2}$, resulta ser un espacio de Hilbert.

- Sea $X = L^2(a, b)$ donde $L^2(a, b) = \left\{ f : [a, b] \longrightarrow \mathbb{R} / f \text{ es medible y } \int_a^b |f(x)|^2\, dx < \infty \right\}$ el espacio vectorial de las funciones medibles cuadrado integrables. Este espacio, dotado del producto interno $(f, g)_{L^2(a,b)} = \int_a^b f(x)g(x)dx$ y la norma asociada a dicho producto $\|f\|_{L^2(a,b)} = \sqrt{\int_a^b |f(x)|^2\, dx}$, es un espacio de Hilbert.

1.2. Aplicaciones continuas en espacios de Banach

En esta sección presentamos los resultados de continuidad de aplicaciones en espacios normados, además daremos los conceptos de espacio dual algebraico y topológico, exponiendo algunas de sus propiedades.

Definición 1.10. Sean $(X, \| \|_X)$, $(Y, \| \|_Y)$ dos espacios vectoriales normados y $f : X \longrightarrow Y$ una aplicación.

(i) Decimos que f *es lineal* si $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ para todo $x, y \in X$ y $\alpha, \beta \in \mathbb{R}$.

(ii) Decimos que f *es continua en* $a \in X$, si dado $\epsilon > 0$ existe $\delta > 0$ tal que si $0 < \|x - a\|_X < \delta$ entonces $\|f(x) - f(a)\|_Y < \epsilon$.

(iii) Decimos que f *es continua en* X si es continua en todo punto $a \in X$.

Los siguientes teoremas caracterizan la continuidad de una aplicación $f : X \longrightarrow Y$.

Teorema 1.11. Sean $(X, \| \|_X)$, $(Y, \| \|_Y)$ dos espacios vectoriales normados, $f : X \longrightarrow Y$ una aplicación y $a \in X$. Son equivalentes:

(i) f es continua en a.

(*ii*) Para toda sucesión $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ convergente a a, la sucesión $\{f(x_n)\}_{n\in\mathbb{N}} \subseteq Y$ converge a $f(a)$.

Demostración. Ver Luenberger [13] (*pag. 28, Proposición 1*).

$\square$

Teorema 1.12. Sean $(X, \|\ \|_X)$, $(Y, \|\ \|_Y)$ dos espacios vectoriales normados y $f : X \longrightarrow Y$ una aplicación lineal. Son equivalentes:

(*i*) f es continua en 0.

(*ii*) Existe $k \geq 0$ tal que $\|f(x)\|_Y \leq k\,\|x\|_X$, para todo $x \in X$.

(*iii*) f es continua en X.

Demostración. Ver Kreiszig [11] (*pag. 97, Teorema 2.7.9*).

$\square$

Denotaremos por $L(X,Y)$ al espacio vectorial de todas las aplicaciones lineales y continuas de X en Y. En $L(X,Y)$ definimos la aplicación $\|\ \|_{L(X,Y)} : L(X,Y) \longrightarrow \mathbb{R}$ dada por

$$\|f\|_{L(X,Y)} = \sup_{x\neq 0} \frac{\|f(x)\|_Y}{\|x\|_X}$$

esta aplicación, constituye una norma en $L(X,Y)$. El siguiente resultado nos da una condición para que el par $(L(X,Y), \|\ \|_{L(X,Y)})$ sea un espacio de Banach.

Teorema 1.13. Sean $(X, \|\ \|_X)$, $(Y, \|\ \|_Y)$ dos espacios vectoriales normados. Si $(Y, \|\ \|_Y)$ es un espacio de Banach, entonces $(L(X,Y), \|\ \|_{L(X,Y)})$ es un espacio de Banach.

Demostración. Ver Eidelman, Vitalis, Tsolomitis [4] (*pag. 55, Teorema 4.1.1*).

$\square$

Como caso particular del Teorema 1.13, si $Y = \mathbb{R}$, el espacio $L(X,\mathbb{R})$ será un espacio de Banach. Este espacio recibe el nombre de *espacio dual topológico* y será denotado por $X^\star$. La norma del espacio $X^\star$ viene dada por

$$\|f\|_{X^\star} = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|_X}$$

Además, el espacio vectorial de todas las aplicaciones lineales $f : X \longrightarrow \mathbb{R}$ será llamado *espacio dual algebraico*, simbolizado por X'. Observemos que $X^\star \subseteq X'$.

Definición 1.14. Sean $(X, \|\ \|_X)$, $(Y, \|\ \|_Y)$ dos espacios vectoriales y $f : X \longrightarrow Y$ una aplicación lineal, definimos los conjuntos:

(*i*) $Nu(f) = \{x \in X / f(x) = 0_Y\}$ denominado *núcleo de f*, donde 0_Y es el vector nulo del espacio Y.

(ii) $Im(f) = \{f(x)/x \in X\}$ es llamado *imagen de* f.

Definición 1.15. Sean $(X, \| \|_X)$, $(Y, \| \|_Y)$ dos espacios vectoriales normados y $f : X \longrightarrow Y$ una aplicación lineal:

(i) Decimos que f es una *aplicación inyectiva* si y sólo si $Nu(f) = \{0_X\}$ donde 0_X es el vector nulo del espacio X.

(ii) Decimos que f es una *aplicación sobreyectiva* si y sólo si $Im(f) = Y$.

(iii) Dado $A \subseteq X$, denominamos al conjunto $f(A) = \{f(x)/x \in A\}$ como la *imagen de* A *vía* f.

Definición 1.16. Sean $(X, \| \|_X)$, $(Y, \| \|_Y)$ dos espacios vectoriales normados y $f \in L(X, Y)$. Decimos que f es una *aplicación abierta* si dado cualquier conjunto abierto $U \subseteq X$, su imagen $f(U)$ es un conjunto abierto en Y.

El siguiente resultado es conocido como el teorema de la aplicación abierta.

Teorema 1.17. Sean $(X, \| \|_X)$, $(Y, \| \|_Y)$ dos espacios espacios de Banach y $T \in L(X, Y)$ una aplicación sobreyectiva. Entonces T es una aplicación abierta, es decir para todo conjunto abierto $U \subseteq X$, $f(U)$ es abierto en Y.

Demostración. Ver Eidelman, Vitalis, Tsolomitis [4] (*pag. 131, Teorema 9.2.1*).

□

1.3. Convergencias débil y débil - $\star$. Espacios reflexivos

Sean $(X, \| \|_X)$ un espacio normado y $X^\star$ su espacio dual.

Definición 1.18. Decimos que una sucesión $(x_n)_{n\in\mathbb{N}} \subseteq X$ *converge débilmente* al punto $x \in X$, denotado por $x_n \rightharpoonup x$, si para cada $l \in X^\star$ se cumple

$$\lim_{n\to\infty} l(x_n) = l(x)$$

Definición 1.19. Decimos que una sucesión $(l_n)_{n\in\mathbb{N}} \subseteq X^\star$ *converge débil-*$\star$ (débil estrella) al punto $l \in X^\star$, denotado por $l_n \overset{\star}{\rightharpoonup} l$, si para cada $x \in X$ se cumple

$$\lim_{n\to\infty} l_n(x) = l(x)$$

Ejemplos.

- Consideremos el espacio de Hilbert $X = l_2(\mathbb{R})$ y la sucesión $(x_n)_{n\in\mathbb{N}} \subseteq X$ definida por

$$x_n = (x_n(i)) = \begin{cases} 0, & \text{si } i \neq n \\ 1, & \text{si } i = n \end{cases}$$

Afirmamos que $x_n \rightharpoonup 0_X$. En efecto, por el teorema de representación de Riesz (ver Jost [9], pag. 274, Teorema 21.6), para cada $l \in (l_2(\mathbb{R}))^\star$ existe un único $y = (y_n)$ en $l_2(\mathbb{R})$ tal que $l(x) = (y, x)$ para todo $x \in l_2(\mathbb{R})$. Además tenemos $y_n \longrightarrow 0$ en $\mathbb{R}$, debido a la convergencia de la serie $\sum_{n=1}^{\infty} y_n^2$. Así

$$\begin{aligned} \lim_{n\to\infty} l(x_n) &= \lim_{n\to\infty} (y, x_n) \\ &= \lim_{n\to\infty} y_n \\ &= 0 \end{aligned}$$

Por lo tanto, de esta igualdad tenemos la afirmación.

- Sea $X = c_0(\mathbb{R})$ el espacio de las sucesiones de números reales convergentes a 0, dotado con la norma $\|(a_n)\| = \sup_{n\in\mathbb{N}} |a_n|$ para cada elemento en X. Este espacio vectorial, está contenido en el espacio $l_1(\mathbb{R}) = \left\{(a_n)_{n\in\mathbb{N}} / \sum_{n=1}^{\infty} |a_n| < \infty\right\}$. Se tiene que $X^* = l_1(\mathbb{R})$. Definamos la sucesión $(f_n)_{n\in\mathbb{N}}$ en $X^\star$ dada por

$$f_n(i) = \begin{cases} 0, & \text{si } i \neq n \\ 1, & \text{si } i = n \end{cases}$$

Veamos que $f_n \overset{\star}{\rightharpoonup} 0_{X^\star}$. En efecto, sea $x = (a_n)$ en $c_0(\mathbb{R})$. Por la definición de (f_n) se tiene $(f_n, x) = a_n$ para cada $n \in \mathbb{N}$. Como $a_n \longrightarrow 0$, resulta que $(f_n, x) \longrightarrow 0$ cuando $n \to \infty$. Por tanto $f_n \overset{\star}{\rightharpoonup} 0_{X^\star}$.

Teorema 1.20. Sea $(X, \| \|_X)$ un espacio normado. Se cumplen:

(i) Los límites débil y débil-$\star$ son únicos.

(ii) Si $(x_n) \subseteq X$ y $x_n \longrightarrow x$ entonces $x_n \rightharpoonup x$.

(iii) Si $\dim(X)$ es finita, las convergencias fuerte, débil y débil-$\star$ son equivalentes.

(iv) Si $x_n \rightharpoonup x$ entonces toda subsucesión (x_{n_k}) de (x_n) converge débilmente a x.

(v) Si (x_n) y (w_n) son dos sucesiones débilmente convergentes a x y w respectivamente, entonces $\alpha x_n + \beta w_n \rightharpoonup \alpha x + \beta w$ para todo $\alpha, \beta \in \mathbb{R}$.

Demostración. Ver Kreyszig [11] (*pag. 258, Lema 4.8.3 y Teorema 4.8.4*).

□

Definición 1.21. Sean X un e.v.n y $A \subseteq X$ un subconjunto no vacío.

(i) Decimos que A es un conjunto *débilmente cerrado por sucesiones* (débil cerrado por sucesiones), si para toda sucesión $(x_n) \subseteq A$ tal que $x_n \rightharpoonup x$ se tiene que $x \in A$.

(*ii*) Diremos que A es un conjunto *débilmente compacto por sucesiones* (débil compacto por sucesiones), si para toda sucesión $(x_n) \subseteq A$, existe una subsucesión (x_{n_k}) de (x_n) y $x \in A$ tales que $x_{n_k} \rightharpoonup x$ si $k \to \infty$.

Observación. Si $A \subseteq X$ es un subconjunto débilmente cerrado, por la Definición 1.21, el Lema 1.5 y el Teorema 1.20 tenemos que A es cerrado. El recíproco no es cierto, sin embargo bajo ciertas hipótesis, se cumple que un cerrado (fuerte), es cerrado débil por sucesiones.

Definición 1.22. Sea $(X, \| \|)$ un espacio de Banach. Decimos que X es un espacio *reflexivo* si la bola unitaria $\bar{B}(0,1) = \{x \in X / \|x\| \leq 1\}$ es débilmente compacta por sucesiones.

Ejemplos.

(*i*) El espacio $\mathbb{R}^n$ con la norma usual es un espacio reflexivo. En general todo espacio de dimensión finita es reflexivo.

(*ii*) Los espacios $l^p(\mathbb{R}) = \left\{ (x_n) \subseteq \mathbb{R} / \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}$ con $1 < p < \infty$ son reflexivos.

(*iii*) Todo espacio de Hilbert es reflexivo.

(*iv*) El espacio $C[a,b]$ no es reflexivo.

El siguiente Teorema caracteriza a los espacios de Banach reflexivos.

Teorema 1.23. (Eberlein-Shmulyan) Sea $(X, \| \|)$ un espacio de Banach. Entonces son equivalentes:

(*i*) X es reflexivo.

(*ii*) Para toda sucesión acotada $(x_n)_{n \in \mathbb{N}}$ en X existe una subsucesión $(x_{n_k})_{k \in \mathbb{N}}$ de (x_n) y $x \in X$ tales que $x_{n_k} \rightharpoonup x$.

Demostración. Ver Eidelman, Vitalis, Tsolomitis [4] (*pag. 152, Teorema 9.8.1*).

□

1.4. Conjuntos convexos en espacios normados

Sean X un e.v.n y S un subconjunto de X.

Definición 1.24. Decimos que S es un conjunto convexo, si $\lambda x + (1-\lambda) y \in S$, para todo $x, y \in S$ y $\lambda \in [0,1]$.

Ejemplos.

- En $X = \mathbb{R}$ todo intervalo es convexo. Además por convención, el conjunto vacío $\emptyset$ es convexo.

- Sea $X = \mathbb{R}^n$, el conjunto $S = \{x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n / x_i \geq 0\}$ llamado primer ortante en $\mathbb{R}^n$ es un conjunto convexo.
- En $X = C[0,1]$, el conjunto $W = \{f \in C[0,1] / f(x) < 0\}$ es convexo.
- Si $X^\star$ es el espacio dual de X y $a \in X$, el conjunto $T = \{f \in X^\star / f(a) > 0\}$ es convexo en $X^\star$.
- Las bolas abiertas y cerradas en un espacio normado, son conjuntos convexos.

Los siguientes lemas establecen propiedades algebraicas y topológicas de los conjuntos convexos.

Lema 1.25. Sean X un e.v.n y A, B subconjuntos convexos de X. Se cumplen:

(i) $A \cap B$ es convexo. En general si $\{A_i\}_{i \in I}$ es una familia de conjuntos convexos en X entonces $\bigcap\limits_{i \in I} A_i$ es un conjunto convexo.

(ii) Si α, β son escalares entonces el conjunto $\alpha A + \beta B = \{\alpha x + \beta y / x \in A \wedge y \in B\}$ es convexo. En particular $-A = \{-x / x \in A\}$ es convexo.

(iii) Si $T : X \longrightarrow Y$ es una aplicación lineal, entonces $T(A)$ es convexo.

Demostración. Ver Berkovitz [2] (*pag. 37, Lema 2.1 y 2.2*).

□

Lema 1.26. Sean X un e.v.n y A un subconjunto convexo de X. Entonces:

(i) $\bar{A}$ es convexo.

(ii) $\overset{\circ}{A}$ es convexo.

(iii) Si $\overset{\circ}{A} \neq \emptyset$ entonces $\overline{\overset{\circ}{A}} = \bar{A}$ y $\overset{\circ}{\bar{A}} = \overset{\circ}{A}$.

Demostración. Ver Berkovitz [2] (*pag. 37, Lema 2.4 y 2.5*).

□

Un concepto más general respecto a la convexidad de un conjunto, viene dada por la siguiente definición.

Definición 1.27. Sean X un e.v.n, $S \subseteq X$ no vacío y $\bar{x} \in S$. Decimos que S es estrellado respecto a $\bar{x}$, si $\lambda x + (1 - \lambda)\bar{x} \in S$ para todo $x \in S$ y $0 \leq \lambda \leq 1$.

Observaciones.

(i) Si S es convexo entonces S es estrellado respecto a todos sus puntos.

(ii) El recíproco también se cumple. Si S es estrellado respecto a todos sus puntos entonces S es convexo.

Definición 1.28. Sean X un e.v.n y $C \subseteq X$ no vacío. Decimos que:

(i) C es un cono si para todo $\lambda \geq 0$ y $x \in C$ entonces $\lambda x \in C$.

(ii) C es un cono puntual si $x \in C$ y $-x \in C$ implica que $x = 0$.

(iii) C es un cono convexo si es cono y conjunto convexo.

Ejemplos.

- $C = \{x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n / x_i \geq 0\}$ es un cono convexo en $\mathbb{R}^n$.
- Sea $X = C\,[0,1]$ y $C = \{f \in X / f(t) \geq 0\}$ es un cono puntual.

El siguiente resultado caracteriza a los conos convexos.

Teorema 1.29. Sean X un e.v.n y $C \subseteq X$ un cono. Entonces C es un cono convexo si y sólo si $x + y \in C$ para cualesquiera $x, y \in C$.

Demostración.

Supongamos que C es un cono convexo entonces $\frac{1}{2}(x+y) = \frac{1}{2}x + \frac{1}{2}y \in C$. Luego $x + y = 2\left(\frac{1}{2}(x+y)\right)$ pertenece a C.

Reciprocamente, tomemos $\lambda \in [0,1]$ y $x, y \in C$. Como C es cono λx y $(1-\lambda)y$ pertenecen a C, y por hipótesis $\lambda x + (1-\lambda)y \in C$, así C es convexo. Por lo tanto C es cono convexo.

$\square$

Definición 1.30. Sean X un e.v.n y $S \subseteq X$ no vacío. Definimos el cono generado por S, denotado por cono(S), al conjunto

$$\text{cono}(S) = \{\lambda x / \lambda \geq 0 \wedge x \in S\}$$

Claramente, cono(S) es convexo. Además si S es un cono entonces cono$(S) = S$.

Ejemplos.

- Sea $S = B(0_X, 1)$ la bola unitaria en un espacio normado $(X, \|\ \|)$, entonces cono$(S) = X$.
- Sea $S = \{(x, f(x) / x \in \mathbb{R}\}$ un subconjunto de $\mathbb{R}^2$, donde f esta definido por

$$f(x) = \begin{cases} x \operatorname{sen}\left(\dfrac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

  Tenemos que cono$(S) = \big\{(x, y) \in \mathbb{R}^2 / \,|y| \leq |x|\big\}$.

1.5. Funcionales convexas en espacios normados

Sea X un espacio vectorial normado.

Definición 1.31. Decimos que una aplicación $f : X \longrightarrow Y$ es una funcional si $Y = \mathbb{R}$. La funcional f será llamada funcional lineal, si es una aplicación lineal.

Ejemplos.

- Sean $X = C[0,1]$ y las aplicaciones $f : X \longrightarrow \mathbb{R}$ y $g : X \longrightarrow \mathbb{R}$ dadas por:

$$f(x) = \int_0^1 x(t)dt$$
$$g(x) = x^2(0)$$

  para cada $x \in X$. Tenemos que f y g son funcionales lineal y no lineal respectivamente.

- Sea $X = l^1(\mathbb{R})$ y $f : X \longrightarrow \mathbb{R}$ definida por $f(x) = \max_{n\in\mathbb{N}} |a_n|$ donde $x = (a_n) \in X$, es una funcional (no lineal).

Definición 1.32. Sean X un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional. Diremos que f es una funcional sublineal si:

(i) $f(x + y) \leq f(x) + f(y)$ para todo $x, y \in X$.

(ii) $f(\alpha x) = \alpha f(x)$ para todo $x \in X$ y $\alpha \in \mathbb{R}$.

Si f satisface la condición $f(\alpha x) = \alpha f(x)$ para todo $x \in X$ y $\alpha \geq 0$, diremos que f es una funcional *positivamente homogénea*.

Definición 1.33. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Decimos que f es *Lipschitz continua* en $\bar{x} \in S$, si existen $k > 0$ y $\delta > 0$ tales que

$$|f(x) - f(y)| \leq k \|x - y\|$$

para todo $x, y \in S \cap B(\bar{x}, \delta)$.

La funcional f se denominará Lipschitz continua en S si existe $K > 0$ tal que $|f(x) - f(y)| \leq K \|x - y\|$ para todo $x, y \in S$.

Claramente, si f es Lipschitz continua en S entonces f es continua en S. Además la Definición 1.10 y el Teorema 1.11 son válidos para funcionales, basta considerar el caso en que $Y = \mathbb{R}$.

Definición 1.34. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Definimos:

(i) El epígrafe de f como el conjunto $\operatorname{epi}(f) = \{(x, \alpha) \in S \times \mathbb{R} / f(x) \leq \alpha\}$.

(ii) El conjunto de subnivel $S_\alpha = \{x \in S / f(x) \leq \alpha\}$, donde $\alpha \in \mathbb{R}$.

Pasaremos a definir el concepto de funcional convexa en espacios normados.

Definición 1.35. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Decimos que la funcional f es *convexa* en S si para cualesquiera $x, y \in S$ y $\lambda \in [0, 1]$, se cumple $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.

Decimos que f es cóncava en S si $-f$ es convexa en S.

Observemos que en la definición de convexidad, es importante el hecho que el conjunto S sea convexo.

Ejemplos.

- Si X es un espacio normado, entonces la funcional norma $\| \| : X \longrightarrow \mathbb{R}$ es convexa.
- Sea $X = C[0, 1]$ y consideremos la funcional $f : X \longrightarrow \mathbb{R}$ definida por $f(x) = \int_0^1 (x(t))^2 dt$, por la convexidad de la función real $g(u) = u^2$, tenemos que f es convexa en X.

El siguiente Teorema caracteriza las funcionales convexas en espacios normados.

Teorema 1.36. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Entonces f es convexa si y sólo si $\operatorname{epi}(f)$ es un conjunto convexo.

Demostración.

$\Longrightarrow)$

Supongamos que f es convexa y sean (x_1, α_1) y (x_2, α_2) dos puntos arbitrarios en $\operatorname{epi}(f)$, tenemos que $f(x_1) \leq \alpha_1$ y $f(x_2) \leq \alpha_2$. Así para cada $\lambda \in [0, 1]$, se tiene que

$$\lambda(x_1, \alpha_1) + (1 - \lambda)(x_2, \alpha_2) = (\lambda x_1 + (1 - \lambda)x_2, \lambda\alpha_1 + (1 - \lambda)\alpha_2) \tag{1.1}$$

Por la convexidad de S, el conjunto $\mathbb{R}$ y la funcional f, de la identidad (1.1) tenemos que $\lambda(x_1, \alpha_1) + (1 - \lambda)(x_2, \alpha_2) \in S \times \mathbb{R}$ y

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)x_2) &\leq \lambda f(x_1) + (1 - \lambda)f(x_2) \\ &\leq \lambda\alpha_1 + (1 - \lambda)\alpha_2 \end{aligned}$$

Por lo tanto $\lambda(x_1, \alpha_1) + (1 - \lambda)(x_2, \alpha_2) \in \operatorname{epi}(f)$. Así $\operatorname{epi}(f)$ es convexo.

$\Longleftarrow$)

Por hipótesis, epi(f) es convexo. Sean $x_1, x_2 \in S$, tenemos que los pares $(x_1, f(x_1))$ y $(x_2, f(x_2))$ son dos puntos en epi(f). Sea $\lambda \in [0,1]$, por la convexidad de epi(f), se sigue que

$$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) = (\lambda x_1 + (1-\lambda)x_2, \lambda f(x_1) + (1-\lambda)f(x_2))$$

pertenece a epi(f), por lo tanto $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$, demostrando de esta manera, la convexidad de f.

$\square$

Lema 1.37. Sean $(X, \|\ \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Si f es convexa, entonces S_α es un conjunto convexo.

Demostración.

Supongamos que f es convexa y sean $x, y \in S_\alpha$. Por la Definición 1.34, tenemos $f(x) \leq \alpha$ y $f(y) \leq \alpha$. Como S es convexo, tomando $\lambda \in [0,1]$ obtenemos

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\leq \lambda f(x) + (1-\lambda)f(y) \\ &\leq \lambda\alpha + (1-\lambda)\alpha \\ &= \alpha \end{aligned}$$

Así, $\lambda x + (1-\lambda)y \in S_\alpha$, quedando demostrado el lema.

$\square$

Pasaremos a definir un concepto mas débil respecto a la convexidad de funcionales.

Definición 1.38. Sean $(X, \|\ \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Decimos que f es cuasiconvexa, si para cada $\alpha \in \mathbb{R}$, el conjunto S_α es convexo.

En particular, gracias al Lema 1.37 toda funcional convexa es cuasiconvexa.

Ejemplo.

Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ definida $f(x) = x^3$. Como $S_\alpha = \{x \in \mathbb{R} / f(x) \leq \alpha\} = (-\infty, \sqrt[3]{\alpha})$ es un convexo en $\mathbb{R}$, se sigue de la Definición 1.37 que f es cuasiconvexa. La función f no es convexa puesto que su epígrafe no es un conjunto convexo. El ejemplo mostrado prueba que existen funciones cuasiconvexas pero que no son convexas.

El siguiente teorema caracteriza a una funcional cuasiconvexa, definida en un espacio vectorial normado.

Teorema 1.39. Sean $(X, \|\ \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Entonces f es cuasiconvexa si y sólo si $f(\lambda x_1 + (1-\lambda)x_2) \leq \text{máx}\,\{f(x_1), f(x_2)\}$ para cualesquiera $x_1, x_2 \in S$ y $0 \leq \lambda \leq 1$.

Demostración.

$\Longrightarrow)$

Supongamos que f es cuasiconvexa y $0 \leq \lambda \leq 1$. Sin pérdida de generalidad, supongamos que $\max\{f(x_1), f(x_2)\} = f(x_1)$. Si $\alpha = f(x_1)$ tenemos que $f(x_2) \leq f(x_1)$ es decir $x_2 \in S_\alpha$ y obviamente $x_1 \in S_\alpha$. Como $S_{f(x_1)}$ es convexo, el punto $\lambda x_1 + (1-\lambda)x_2$ pertenece a $S_{f(x_1)}$, luego $f(\lambda x_1 + (1-\lambda)x_2) \leq f(x_1) = \max\{f(x_1), f(x_2)\} = f(x_1)$.

$\Longleftarrow)$

Sea $\lambda \in [0,1]$ y $x_1, x_2 \in S_\alpha$, entonces $\max\{f(x_1), f(x_2)\} \leq \alpha$ puesto que $f(x_1) \leq \alpha$ y $f(x_2) \leq \alpha$. Luego por hipótesis $f(\lambda x_1 + (1-\lambda)x_2) \leq \max\{f(x_1), f(x_2)\} \leq \alpha$. En consecuencia $\lambda x_1 + (1-\lambda)x_2 \in S_\alpha$, demostrando así la convexidad de S_α. Por lo tanto f es cuasiconvexa.

$\square$

1.6. Separación de conjuntos convexos

Definición 1.40. Sean X un e.v.n, $f \in X'$ una funcional no nula y $\alpha \in \mathbb{R}$. El conjunto

$$H_{f,\alpha} = \{x \in X / f(x) = \alpha\}$$

se denomina hiperplano determinado por f y α. La expresión $f(x) = \alpha$ será llamada la ecuación del hiperplano $H_{f,\alpha}$. El siguiente resultado, nos muestra una condición necesaria y suficiente para que un hiperplano sea cerrado.

Teorema 1.41. Sea $(X, \|\ \|)$ un e.v.n. Un hiperplano $H_{f,\alpha}$ es cerrado si y sólo si $f \in X^\star$.

Demostración. Ver Luenberger [13] (*pag. 130, Proposición 3*).

$\square$

Definición 1.42. Sean $(X, \|\ \|)$ un e.v.n, $A, B \subseteq X$ subconjuntos no vacíos y $H = H_{f,\alpha}$ un hiperplano en X.

(*i*) Decimos que H separa a A y B en el sentido débil si $f(x) \leq \alpha$ para todo $x \in A$ y $f(x) \geq \alpha$ para todo $x \in B$.

(*ii*) Decimos que A y B son fuertemente separados por H si existe $\epsilon > 0$ tal que $f(x) \leq \alpha - \epsilon$ para todo $x \in A$ y $f(x) \geq \alpha + \epsilon$ para todo $x \in B$.

Evidentemente toda separación fuerte, es una separación en el sentido débil. Enunciaremos el Teorema de Hahn Banach y sus consecuencias, que darán un sentido algebraico y geométrico al concepto de separación de conjuntos convexos.

Teorema 1.43. (Hahn - Banach) Sean X un espacio vectorial real y $p : X \longrightarrow \mathbb{R}$ una funcional sublineal. Si $M \subseteq X$ es un subespacio vectorial y $f : M \longrightarrow \mathbb{R}$ es una funcional lineal tal que $f(x) \leq p(x)$ para todo $x \in M$. Entonces existe $F \in X'$ tal que:

(i) $F(x) \leq p(x)$ para todo $x \in X$.

(ii) $F(x) = f(x)$ para todo $x \in M$.

Demostración. Ver Kreyszig [11] (*pag. 214, Teorema 4.2.1*).

□

En particular, del Teorema 1.43 tomando $M = X$ tenemos el

Corolario 1.44. Sean X un espacio vectorial real y $f : X \longrightarrow \mathbb{R}$ un funcional sublineal entonces existe $l \in X'$ tal que $l(x) \leq f(x)$ para todo $x \in X$.

Los siguientes resultados son consecuencias del Teorema 1.43 y son conocidos como la primera y segunda forma geométrica del Teorema de Hahn Banach.

Teorema 1.45. (Primera forma geométrica del Teorema de Hahn - Banach)

Sean X un e.v.n y $A, B \subseteq X$ dos conjuntos convexos, disjuntos y no vacíos. Si A es abierto entonces existe un hiperplano cerrado $H_{f,\alpha}$ que separa A y B en forma débil.

Demostración. Ver Botelho [3] (*pag. 42, Teorema 2.2.9*).

□

Teorema 1.46. (Segunda forma geométrica del Teorema de Hahn - Banach)

Sean X un e.v.n y $A, B \subseteq X$ dos conjuntos convexos, disjuntos y no vacíos. Si A es cerrado y B es compacto (en el sentido fuerte) entonces existe un hiperplano cerrado $H_{f,\alpha}$ que separa fuertemente A y B.

Demostración. Ver Botelho [3] (*pag. 44, Teorema 2.2.12*).

□

El siguiente teorema caracteriza la separación débil de dos conjuntos convexos en un espacio normado real.

Teorema 1.47. (Teorema de Separación de Eidelheit)

Sean X un e.v.n, $S, T \subseteq X$ dos conjuntos convexos no vacíos tales que $\overset{\circ}{S} \neq \emptyset$. Son equivalentes:

(i) $\overset{\circ}{S} \cap T = \emptyset$.

(ii) Existen $l \in X^{\star} \setminus \{0_{X^{\star}}\}$ y $\alpha \in \mathbb{R}$ tales que:

   $a)$ $l(s) \leq \alpha \leq l(t)$ para todo $s \in S$ y $t \in T$.

   $b)$ $l(s) < \alpha$ para todo $s \in \overset{\circ}{S}$.

Demostración. Ver Werner [17] (*pag. 71 Teorema 3.2.4*).

$\square$

Los siguientes resultados son consecuencias del Teorema 1.47.

Teorema 1.48. Consideremos X un e.v.n y $S \subseteq X$ un subconjunto convexo y cerrado. Entonces $x \in X \setminus S$ si y sólo si existe $l \in X^\star$ no nulo tal que $l(x) < \inf_{s \in S} l(s)$.

Demostración.

$\Longrightarrow)$

Supongamos que $x \in X \setminus S$. Como S es cerrado, $X \setminus S$ es abierto, luego existe una bola abierta $B(x, r)$ tal que $B(x, r) \cap S = \emptyset$. Por el Teorema 1.47 existe $l \in X^\star$ no nulo y $\alpha \in \mathbb{R}$ tales que $l(x) < \alpha \leq l(s)$ para todo $s \in S$. De esta última desigualdad, tenemos

$$\begin{aligned} l(x) &< \alpha \\ &\leq \inf_{s \in S} l(s) \end{aligned}$$

De donde resulta la implicación.

$\Longleftarrow)$

Tomemos $x \in X$. Por hipótesis tenemos que existe $l \in X^\star$ no nulo tal que $l(x) < \inf_{s \in S} l(s)$. Si $x \in S$ entonces de la desigualdad anterior, tenemos $l(x) < l(x)$ lo cual es un absurdo. Por lo tanto $x \in X \setminus S$.

$\square$

Teorema 1.49. Sea X un e.v.n. Si $x \in X$ entonces existe $l \in X^\star$ tales que $\|l\|_{X^\star} = 1$ y $l(x) = \|x\|$.

Demostración.

Se presentan dos casos:

Caso 1

Si $x = 0$ es evidente, basta considerar $l \in X^\star$ tal que $\|l\|_{X^\star} = 1$.

Caso 2

Si $x \neq 0$, consideremos $S = \bar{B}(0, \|x\|)$ y $T = \{x\}$. Claramente S es convexo y $\overset{\circ}{S} \cap T = \emptyset$. Por el Teorema 1.47 existe $\bar{l} \in X^\star$ no nulo y $\alpha \in \mathbb{R}$ tales que $\bar{l}(s) \leq \alpha \leq \bar{l}(x)$ para todo $s \in S$. Definimos $l = \dfrac{\bar{l}}{\|\bar{l}\|_{X^\star}}$, así tenemos que $\|l\|_{X^\star} = 1$ y $l(s) \leq l(x)$ para todo $s \in S$. Como $\sup_{\|y\| \leq 1} |l(y)| = 1$ logramos

$$
\begin{aligned}
\|x\| &= \|x\| \sup_{\|y\|\leq 1} |l(y)| \\
&= \sup_{\|y\|\leq 1} |l(\|x\|\, y)| \\
&= \sup_{s\in S} |l(s)| \\
&= \sup_{s\in S} l(s) \leq l(x) \ (-s \in S)
\end{aligned}
$$

Obteniendo la desigualdad $\|x\| \leq l(x)$. Por la definición de la norma $\| \|_{X^\star}$ tenemos que $l(y) \leq \|y\|$ para todo $y \in X$. En particular $l(x) \leq \|x\|$, quedando demostrado el teorema.

□

Capítulo 2

Minimización de funcionales en espacios de Banach

Consideremos $(X, \|\ \|)$ un espacio real normado, $S \subseteq X$ un subconjunto no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Bajo estas hipótesis, analizaremos el problema de optimización

$$\min_{x \in S} f(x) \tag{2.1}$$

es decir, encontrar los puntos en S donde f alcanza su mínimo. En general no siempre es posible encontrar una solución al problema (2.1). Como ejemplo, tenemos el caso en que $X = S = \mathbb{R}$ y el funcional continuo $f : S \longrightarrow \mathbb{R}$ dado por $f(x) = e^x$.

Para el caso del problema de maximización,

$$\max_{x \in S} f(x)$$

basta considerar $-f$ en vez de f para llevar el problema al caso del problema (2.1).

2.1. Existencia de puntos minimizantes

Uno de los principales teoremas de existencia de mínimo (y máximo) en espacios abstractos es el Teorema de Weierstrass, el cual afirma que dado un conjunto compacto K en la topología de la norma y una funcional continua $f : K \longrightarrow \mathbb{R}$, entonces el problema (2.1) admite una solución factible en K. En esta sección estudiaremos condiciones más generales que permitan demostrar la existencia de solución para el problema (2.1) sobre conjuntos que no son necesariamente compactos y sobre funcionales que no siempre serán continuas. En principio daremos las siguientes definiciones.

Definición 2.1. Sean X un e.v.n, $S \subseteq X$ un subconjunto no vacío. Decimos que f posee un *mínimo global o absoluto* en el punto $\bar{x} \in S$ si $f(\bar{x}) \leq f(x)$ para todo $x \in S$. En tal sentido, denotaremos por $f(\bar{x}) = \min_{x \in S} f(x)$ al valor mínimo de f en S y $\bar{x} = \operatorname{argmin}_{x \in S} f(x)$ al punto minimizante $\bar{x}$.

Decimos que f posee un *mínimo local o relativo* en el punto $\bar{x} \in S$ si existe $\delta > 0$ tal que $f(\bar{x}) \leq f(x)$ para todo $x \in S \cap B(\bar{x}, \delta)$.

Definición 2.2. Sean X un espacio real normado, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Decimos que f es *débilmente semicontinua inferior* en el punto $\bar{x} \in X$ si para cada sucesión $(x_n) \subseteq S$ tal que $x_n \rightharpoonup \bar{x}$ se cumple

$$\liminf_{n\to\infty} f(x_n) \geq f(\bar{x})$$

Ejemplo.

Consideremos la función $f : \mathbb{R} \longrightarrow \mathbb{R}$ definida por

$$f(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Claramente, f es discontinua en $x = 0$. Afirmamos que f es débilmente semicontinua inferior en $x = 0$. En efecto sea $x_n \rightharpoonup 0$, de la definición de f tenemos $f(x_n) \in \{0, 1\}$. Por lo tanto

$$\liminf_{n\to\infty} f(x_n) \geq 0 = f(0)$$

Mostrando así la semicontinuidad débil inferior de f en $x = 0$.

El siguiente Teorema garantiza la existencia de solución para el problema (2.1).

Teorema 2.3. Sean X un espacio normado y $f : S \longrightarrow \mathbb{R}$ una funcional. Si $S \subseteq X$ es débilmente compacto por sucesiones y f es débilmente semicontinua inferior, entonces existe al menos un punto $\bar{x} \in S$ tal que $f(\bar{x}) \leq f(x)$ para todo $x \in S$. Es decir el problema de minimización (2.1) admite una solución.

Demostración.

En principio afirmamos que existe $\beta = \inf_{x\in S} f(x)$. En efecto, supongamos que f no es acotada inferiormente (hipótesis auxiliar). Así para cada $n \in \mathbb{N}$, podemos encontrar $x_n \in S$ tal que $f(x_n) < -n$. Como $\{x_n\} \subseteq S$ y siendo este un conjunto compacto débil en X, existe una subsucesión (x_{n_k}) de (x_n) y un punto $w \in S$ tales que $x_{n_k} \rightharpoonup w$ si $k \to \infty$. Por lo tanto

$$\begin{aligned} f(w) &\leq \liminf_{k\to\infty} f(x_{n_k}) \\ &< -n_k \\ &< -\infty \end{aligned}$$

Lo cual es una contradicción. Por lo tanto tenemos la afirmación.

Consideremos una sucesión $(z_n) \subseteq S$ tal que $f(z_n) \longrightarrow \beta$ si $n \longrightarrow \infty$. Por la compacidad débil de S existe una subsucesión (z_{n_k}) de (z_n) y un punto $\bar{x} \in S$ tales que $z_{n_k} \rightharpoonup \bar{x}$. Luego por la semicontinuidad débil inferior de f tenemos

$$\begin{aligned} f(\bar{x}) &\leq \liminf_{k\to\infty} f(z_{n_k}) \\ &= \beta \end{aligned}$$

Por lo tanto $f(\bar{x}) \leq \beta = \inf_{x\in S} f(x)$. Así tenemos $f(\bar{x}) \leq f(x)$ para todo $x \in S$. En conclusión existe $\min_{x\in S} f(x)$.

□

Lema 2.4. Sean X un e.v.n y S un subconjunto no vacío, cerrado y convexo. Entonces S es débilmente cerrado por sucesiones.

Demostración.

Consideremos una sucesión $(x_n) \subseteq S$ débilmente convergente a un punto $x \in X$. Veamos que $x \in S$. En efecto, supongamos que $x \notin S$ (hipótesis auxiliar). Por el Teorema 1.48 existe $f \in X^{\star}$ no nulo tal que $f(x) < \inf_{s\in S} f(s)$. Por la continuidad de f, la Definición 1.18 y la desigualdad anterior, tenemos

$$\begin{aligned} f(x) &= \lim_{n\to\infty} f(x_n) \\ &\geq \inf_{s\in S} f(s) \\ &> f(x) \end{aligned}$$

que es un absurdo. Por lo tanto $x \in S$, así S es débilmente cerrado por sucesiones.

□

El siguiente resultado caracteriza la semicontinuidad débil inferior de una funcional f.

Lema 2.5. Sean X un e.v.n, S un subconjunto débilmente cerrado no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Son equivalentes:

(i) f es débilmente semicontinua inferior.

(ii) epi(f) es débilmente cerrado por sucesiones en $S \times \mathbb{R}$.

(iii) Si $\alpha \in \mathbb{R}$ y S_α es no vacío entonces S_α es débilmente cerrado por sucesiones.

Demostración.

$i) \Longrightarrow ii)$

Supongamos que f sea débilmente semicontinua inferior y consideremos una sucesión (x_n, α_n) en epi(f) tal que $(x_n, \alpha_n) \rightharpoonup (x, \alpha)$ con $(x, \alpha) \in S \times \mathbb{R}$. Veamos que $(x, \alpha) \in$ epi(f). En efecto, por la convergencia débil de (x_n, α_n) a (x, α) en $X \times \mathbb{R}$, tenemos que $x_n \rightharpoonup x$ en X y $\alpha_n \rightharpoonup \alpha$ en $\mathbb{R}$ y siendo S débilmente cerrado, tenemos que $x \in S$. Como $\mathbb{R}$ es un espacio finito dimensional por el Teorema 1.20 se cumple que $\alpha_n \longrightarrow \alpha$ si $n \to \infty$, así dado $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $\alpha_n < \alpha + \epsilon$ para todo $n \geq n_0$. Por la Definición 1.34 tenemos $f(x_n) \leq \alpha_n$ para todo $n \in \mathbb{N}$. En particular $f(x_n) \leq \alpha_n < \alpha + \epsilon$ para todo $n \geq n_0$. Así por la semicontinuidad débil inferior de f y la desigualdad anterior conseguimos

$$\begin{aligned} f(x) &\leq \liminf_{n\to\infty} f(x_n) \\ &\leq \liminf_{n\to\infty} f(\alpha_n) \\ &= \alpha < \alpha + \epsilon \end{aligned}$$

resultando $f(x) < \alpha + \epsilon$ para todo $\epsilon > 0$. Por lo tanto $f(x) \leq \alpha$, con lo cual $(x, \alpha) \in \operatorname{epi}(f)$.

ii) $\Longrightarrow$ *iii)*

Supongamos que $\operatorname{epi}(f)$ es débilmente cerrado por sucesiones y tomemos $\alpha \in \mathbb{R}$ tal que $S_\alpha \neq \emptyset$. Como S es débilmente cerrado en X y $\{\alpha\}$ es débilmente cerrado (por el Lema 2.4) tenemos que $S \times \{\alpha\}$ es débilmente cerrado en $X \times \mathbb{R}$.

Afirmación: $S_\alpha \times \{\alpha\} = \mathbf{epi}(f) \cap (S \times \{\alpha\})$.

En efecto

$$\begin{aligned} (x,y) \in S_\alpha \times \{\alpha\} &\Longrightarrow x \in S_\alpha \ \wedge \ y = \alpha \\ &\Longrightarrow f(x) \leq \alpha \ \wedge \ y = \alpha \ \ (x \in S) \\ &\Longrightarrow (x,\alpha) \in \operatorname{epi}(f) \ \wedge \ (x,\alpha) \in (S \times \{\alpha\}) \\ &\Longrightarrow (x,y) \in \operatorname{epi}(f) \cap (S \times \{\alpha\}) \end{aligned}$$

Por otro lado

$$\begin{aligned} (x,y) \in \operatorname{epi}(f) \cap (S \times \{\alpha\}) &\Longrightarrow (x,y) \in \operatorname{epi}(f) \ \wedge \ (x,y) \in S \times \{\alpha\} \\ &\Longrightarrow f(x) \leq y \ \wedge \ x \in S \ \wedge y = \alpha \\ &\Longrightarrow f(x) \leq \alpha \ \wedge \ y = \alpha \\ &\Longrightarrow x \in S_\alpha \ \wedge \ y = \alpha \\ &\Longrightarrow (x,y) \in S_\alpha \times \{\alpha\} \end{aligned}$$

Con lo cual se tiene la afirmación. Como la intersección de dos conjuntos débilmente cerrados por sucesiones es débilmente cerrado por sucesiones, de la afirmación anterior concluimos que $S_\alpha \times \{\alpha\}$ es débilmente cerrado por sucesiones. Por lo tanto S_α es débilmente cerrado por sucesiones.

$iii) \Longrightarrow i)$

Supongamos que f no sea semicontinua débil inferior (hipótesis auxiliar), entonces existe una sucesión $(x_n) \subseteq S$ tal que $x_n \rightharpoonup \bar{x}$ y $\liminf_{n\to\infty} f(x_n) < f(\bar{x})$. Tomemos $\alpha \in \mathbb{R}$ tal que $\liminf_{n\to\infty} f(x_n) < \alpha < f(\bar{x})$. Por la definición de límite inferior, existe una subsucesión (x_{n_j}) de (x_n) tal que $x_{n_j} \in S_\alpha$, para todo $i \in \mathbb{N}$ y $x_{n_j} \rightharpoonup \bar{x}$ si $j \to \infty$. Pero de la desigualdad anterior, tenemos que $f(\bar{x}) > \alpha$, con lo cual $\bar{x} \notin S_\alpha$ que contradice la hipótesis.

$\square$

Lema 2.6. Sean $S \subseteq X$ un conjunto cerrado y convexo de un espacio normado X y $f : X \longrightarrow \mathbb{R}$ una funcional. Si f es continua y cuasiconvexa. Entonces f es débilmente semicontinua inferior.

Demostración.

Por el Lema 2.5 bastara ver que S_α (no vacío) es cerrado débil por sucesiones. En efecto, sea $\alpha \in \mathbb{R}$ tal que $S_\alpha \neq \emptyset$. Siendo f continua y S un conjunto cerrado, entonces S_α es cerrado y además convexo por la Definición 1.38. Por el Lema 2.4 se sigue que S_α es cerrado débil por sucesiones, de donde se sigue el resultado.

$\square$

Lema 2.7. Sean $(X, \| \|)$ un espacio de Banach reflexivo y $S \subseteq X$ un conjunto no vacío, cerrado, convexo y acotado. Entonces S es débilmente compacto por sucesiones.

Demostración.

Tomemos una sucesión $(x_n) \subseteq S$ arbitraria. Como S es acotado tenemos que (x_n) es acotada. Así por el Teorema 1.23 y la reflexividad del espacio X, existen una subsucesión (x_{n_j}) de (x_n) y un punto $x \in X$ tales que $x_{n_j} \rightharpoonup x$ si $j \to \infty$. Siendo S un conjunto cerrado y convexo, por el Lema 2.4 tenemos que S es débilmente cerrado por sucesiones. Por lo tanto $x \in S$, mostrando así la compacidad débil por sucesiones de S.

$\square$

Gracias a los resultados anteriores, enunciaremos una condición necesaria para el punto minimizante de una funcional definida sobre un espacio reflexivo.

Teorema 2.8. Sea $(X, \| \|)$ un espacio de Banach reflexivo y $S \subseteq X$ un subconjunto no vacío, convexo, cerrado y acotado. Si $f : S \longrightarrow \mathbb{R}$ es una función continua y cuasiconvexa. Entonces f posee un punto mínimo en S.

Demostración.

Por las hipótesis sobre S y el Lema 2.7 tenemos que S es secuencialmente débil compacto. Así por el Lema 2.6 f es semicontinua débil inferior. Luego se cumplen las hipótesis del Teorema 2.3, por lo tanto existe $\bar{x} \in S$ tal que $f(\bar{x}) = \min_{x \in S} f(x)$.

$\square$

2.2. Conjunto de puntos minimizantes

En esta sección revisaremos algunas propiedades del conjunto de puntos minimizantes de una funcional f definida en un espacio normado.

Definición 2.9. Sean X un espacio normado, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. El conjunto de puntos minimizantes de f en S esta definido y será denotado como $M(f) = \{\bar{x} \in S / f(\bar{x}) \leq f(x)$, para todo $x \in S\}$.

Teorema 2.10. Consideremos X un e.v.n y $S \neq \emptyset$ un conjunto convexo. Si $f : S \longrightarrow \mathbb{R}$ es cuasiconvexa, entonces $M(f)$ es un conjunto convexo.

Demostración.

Consideremos dos casos. Si $M(f)$ es vacío, no hay nada que probar. Supongamos que $\bar{x} \in M(f)$ y definamos el conjunto

$$M = \{x \in S / f(x) \leq f(\bar{x})\}$$

Por la cuasiconvexidad de f y la Definición 1.38 se sigue que M es convexo. Veamos que $M = M(f)$. En efecto si $w \in M$ entonces $f(w) \leq f(\bar{x})$ en consecuencia $f(w) \leq f(x)$ para todo $x \in S$, por tanto $w \in M(f)$. Por otra parte, si $w \in M(f)$ entonces $f(w) \leq f(x)$ para todo $x \in S$, en particular $f(w) \leq f(\bar{x})$, es decir $w \in M$ logrando el resultado.

□

El siguiente resultado muestra las condiciones para que un mínimo local de una funcional f sea un mínimo global.

Teorema 2.11. Sean X un e.v.n, S un subconjunto no vacío y convexo y $f : S \longrightarrow \mathbb{R}$ una funcional convexa. Si $\bar{x}$ es un punto mínimo local de f entonces $\bar{x}$ es un punto mínimo global de f en S.

Demostración.

Sea $\bar{x} \in S$ un punto mínimo local de f, entonces existe $\delta > 0$ tal que $\bar{x}$ es punto mínimo de f en $S \cap B(\bar{x}, \delta)$, es decir $f(\bar{x}) \leq f(x)$ para todo $x \in S \cap B(\bar{x}, \delta)$. Afirmamos que $f(\bar{x}) \leq f(x)$ para todo $x \in S \setminus B(\bar{x}, \delta)$. En efecto, sea $x \in S$ tal que $x \notin B(\bar{x}, \delta)$, entonces el número real $\lambda = \delta / \|x - \bar{x}\|$ pertenece al intervalo $(0, 1)$, así el punto $x_\lambda = \lambda x + (1 - \lambda)\bar{x}$ por hipótesis, pertenece a S. Luego

$$\begin{aligned} \|x_\lambda - \bar{x}\| &= \|\lambda x + (1 - \lambda)\bar{x} - \bar{x}\| \\ &= \lambda \|x - \bar{x}\| \\ &= \delta \end{aligned}$$

demostrando que x_λ pertenece a $S \cap B(\bar{x}, \delta)$. De la convexidad de f y la hipótesis de minimalidad local, tenemos

$$\begin{aligned} f(\bar{x}) &\leq f(x_\lambda) \\ &= f(\lambda x + (1-\lambda)\bar{x}) \\ &\leq \lambda f(x) + (1-\lambda) f(\bar{x}) \end{aligned} \tag{2.2}$$

Así, de la desigualdad (2.2) tenemos $\lambda f(\bar{x}) \leq \lambda f(x)$ por lo tanto $f(\bar{x}) \leq f(\bar{x})$ puesto que $\lambda \in (0,1)$, con lo cual $f(\bar{x}) \leq f(x)$ para todo $x \in S$, mostrando de esta forma la minimalidad global de $\bar{x}$ en S.

□

2.3. Aplicación a un problema de proximalidad

Sean X un espacio normado, S un subconjunto no vacío de X y $\hat{x} \in X$ un punto arbitrario. Se presenta el problema de proximalidad o punto proximizante, como el problema de encontrar un punto $\bar{x} \in S$ que minimice la funcional

$$\begin{aligned} f_{\hat{x}} : S &\longrightarrow \mathbb{R} \\ x &\longmapsto f_{\hat{x}}(x) = \|x - \hat{x}\| \end{aligned} \tag{2.3}$$

es decir $\|\bar{x} - \hat{x}\| \leq \|x - \hat{x}\|$ para todo $x \in S$. Comenzaremos con la siguiente definición.

Definición 2.12. Sean $(X, \| \, \|)$ un espacio normado y S un subconjunto no vacío de X. Decimos que S es un *conjunto proximinal* si para cada $\hat{x} \in X$ existe $\bar{x} \in S$ tal que $\bar{x} = \underset{x \in S}{\operatorname{argmin}} f_{\hat{x}}(x)$. En tal sentido decimos que $\bar{x}$ es la mejor aproximación o punto proximal a $\hat{x}$ desde S.

El siguiente resultado considera las condiciones necesarias para que un conjunto sea proximinal dentro de un espacio normado.

Teorema 2.13. Sean $(X, \| \, \|)$ un espacio de Banach reflexivo y $S \subseteq X$ no vacío. Si S es cerrado y convexo entonces S es proximinal.

Demostración.

Sea S un subconjunto cerrado, convexo y no vacío de X y $\hat{x} \in X$ un punto arbitrario. Veamos que existe una solución para el problema de minimización, respecto a la funcional $f_{\hat{x}}(x) = \|x - \hat{x}\|$ en S.

Afirmamos que $f_{\hat{x}}$ es continua y convexa. En efecto:

- $f_{\hat{x}}$ es continua: Si $x, y \in S$ tenemos

$$\begin{aligned}|f_{\hat{x}}(x) - f_{\hat{x}}(y)| &= |\|x - \hat{x}\| - \|y - \hat{x}\|| \\ &\leq \|(x - \hat{x}) - (y - \hat{x}\| = \|x - y\|\end{aligned}$$

- $f_{\hat{x}}$ es convexa: Si $x, y \in S$ y $\lambda \in [0, 1]$ entonces

$$\begin{aligned}f_{\hat{x}}(\lambda x + (1-\lambda)y) &= \|\lambda x + (1-\lambda)y - \hat{x}\| \\ &= \|\lambda(x - \hat{x}) + (1-\lambda)(y - \hat{x})\| \\ &\leq \lambda \|x - \hat{x}\| + (1-\lambda)\|y - \hat{x}\| \\ &= \lambda f_{\hat{x}}(x) + (1-\lambda) f_{\hat{x}}(y)\end{aligned}$$

Por lo tanto f es continua y convexa (cuasiconvexa). Tomemos $\tilde{x} \in S$ fijo pero arbitrario. Definimos el conjunto $\tilde{S} = \{x \in S / f_{\hat{x}}(x) \leq f_{\hat{x}}(\tilde{x})\}$. Por la convexidad de f tenemos que $\tilde{S}$ es convexo en X y además para cada $x \in \tilde{S}$ se tiene

$$\begin{aligned}\|x\| &= \|x - \hat{x} + \hat{x}\| \\ &\leq \|x - \hat{x}\| + \|\hat{x}\| \\ &= f_{\hat{x}}(x) + \|\hat{x}\| \leq f(\tilde{x}) + \|\hat{x}\| \qquad (2.4)\end{aligned}$$

Por la desigualdad (2.4) tenemos que $\tilde{S}$ es acotado. Como $f_{\hat{x}}$ es continua y S es cerrado entonces $\tilde{S}$ es cerrado. Así estamos en las hipótesis del Teorema 2.8, por lo tanto, existe $\bar{x} \in \tilde{S}$ tal que $f_{\hat{x}}(\bar{x}) \leq f_{\hat{x}}(x)$ para todo $x \in \tilde{S}$. Por otro lado como $\tilde{S} \subseteq S$ entonces $\bar{x} \in S$ y además para cada $x \notin \tilde{S}$ tenemos que $f_{\hat{x}}(x) > f_{\hat{x}}(\tilde{x}) \geq f(\bar{x})$, por lo tanto $f_{\hat{x}}(\bar{x}) \leq f_{\hat{x}}(x)$ para todo $x \in S$. Como $\hat{x} \in X$ fue tomado arbitrario, tenemos que S es un conjunto proximinal.

□

2.4. Problema de aproximación de Chebyshev

Consideremos M un espacio métrico completo y el espacio de Banach

$$C(M) = \{f : M \longrightarrow \mathbb{R} / f \text{ es continua en } M\}$$

bajo la norma $\|f\|_{C(M)} = \max\limits_{t \in M} |f(t)|$.

Consideremos $S \subseteq C(M)$ un subconjunto no vacío, cerrado y convexo y $\hat{f} \in C(M)$. El problema de minimización en este caso, respecto al problema (2.3) consiste en encontrar

$\bar{f} \in S$ tal que $\left\|\bar{f} - \hat{f}\right\| \leq \left\|f - \hat{f}\right\|$ para todo $f \in S$. Como $C(M)$ es un espacio de Banach no reflexivo, el Teorema 2.13 no puede aplicarse directamente, sin embargo se presenta el siguiente resultado.

Teorema 2.14. Sea $S \subseteq C(M)$ un subconjunto no vacío, cerrado y convexo, tal que para cualquier $\tilde{f} \in S$ el subespacio generado por $S - \{\tilde{f}\} = \{f - \tilde{f}/f \in S\}$ es reflexivo. Entonces S es un conjunto proximinal.

Demostración.

Sean $\tilde{f} \in S$ y $\hat{f} \in C(M)$ fijo, tenemos

$$\begin{aligned} \inf_{f \in S} \left\|f - \hat{f}\right\| &= \inf_{f \in S} \left\|f - \tilde{f} - (\hat{f} - \tilde{f})\right\| \\ &= \inf_{f \in S - \{\tilde{f}\}} \left\|f - (\hat{f} - \tilde{f})\right\| \end{aligned} \tag{2.5}$$

Sea V el subespacio generado por $\tilde{f} - \hat{f}$ y $S - \{\tilde{f}\}$. Por hipótesis $S - \{\tilde{f}\}$ es reflexivo, así tenemos que V es reflexivo (puesto que el espacio generado por el vector $\tilde{f} - \hat{f}$ es finito dimensional, por tanto reflexivo). Como S es cerrrado y convexo, $S - \{\tilde{f}\}$ es cerrado y convexo. Así por el Teorema 2.13 $S - \{\hat{f}\}$ es proximinal y por la identidad (2.5), S es proximinal y dicho punto mínimo pertenece a S.

□

Observación.

En general el subespacio generado por $S - \{\tilde{f}\}$ es finito dimensional y por tanto reflexivo, existiendo el caso en que S sea un subespacio generado por combinaciones lineales finitas de elementos de $C(M)$, tales como por ejemplo si tomamos $M = [a, b]$, que serán combinaciones lineales de funciones de la forma $f_n(t) = t^n$. Para este último caso, el problema de aproximación de Chebyshev, posee una solución.

2.5. Aplicación a problemas de control óptimo

En esta sección aplicaremos el Teorema 2.8 a un problema del Control Óptimo. En principio mostraremos un problema de Control Óptimo que no posee solución.

Consideremos el sistema dinámico

$$\begin{cases} x'(t) = -u^2(t) \quad \text{c.t.p en } [0, 1] \\ x(0) = 1 \\ x(1) = 0 \end{cases} \tag{2.6}$$

con una función de control $u \in L^2[0, 1]$. Sabemos que la solución del sistema (2.6) viene dada por

$$x(t) = c - \int_0^t u^2(s)ds \quad t \in [0,1]$$

Por las condiciones iniciales dadas en (2.6), obtenemos

$$\begin{aligned} x(0) = 1 &\Longrightarrow 1 = c - 0 \\ &\Longrightarrow c = 1 \end{aligned}$$

Con lo cual $x(t) = 1 - \int_0^t u^2(s)ds$. Por la segunda condición inicial tenemos

$$\begin{aligned} x(1) = 0 &\Longrightarrow 1 - \int_0^1 u^2(s)ds = 0 \\ &\Longrightarrow \int_0^1 u^2(s)ds = 1 \end{aligned} \tag{2.7}$$

Consideremos la funcional $f(u) = \int_0^1 t^2u^2(t)dt$ para cada $u \in L^2(0,1)$ y el conjunto S definido por

$$S = \left\{ u \in L^2(0,1) / \int_0^1 u^2(s)ds = 1 \right\}$$

es decir, S es la esfera unitaria en $L^2(0,1)$. Como $t^2u^2(t) \geq 0$ para todo $t \in [0,1]$, tenemos que $\int_0^1 t^2u^2(t)dt \geq 0$, así $\inf\limits_{u \in S} f(u) \geq 0$.

Sea la sucesión de funciones admisibles $(u_n)_{n\in\mathbb{N}}$ dada por

$$u_n(t) = \begin{cases} n, & \text{c.t.p en } [0, \dfrac{1}{n^2}), \\ 0, & \text{c.t.p en } [\dfrac{1}{n^2}, 1] \end{cases}$$

Si $n \in \mathbb{N}$ entonces

$$\begin{aligned} \|u_n\|^2_{L^2(0,1)} &= \int_0^1 u_n^2(t)dt \\ &= \int_0^{\frac{1}{n^2}} n^2 dt = 1 \end{aligned}$$

Así, $u_n \in S$ para todo $n \in \mathbb{N}$. Por otra parte, tenemos

$$\begin{aligned} f(u_n) &= \int_0^1 t^2u_n^2(t)dt \\ &= \int_0^{\frac{1}{n^2}} t^2n^2dt = \frac{1}{3n^4} \end{aligned}$$

y además

$$\begin{aligned}\lim_{n\to\infty} f(u_n) &= \lim_{n\to\infty} \frac{1}{3n^4}\\ &= 0 = \inf_{u\in S} f(u)\end{aligned}$$

Supongamos que f posea un minimizante $\bar{u} \in S$, entonces $f(\bar{u}) = 0$, es decir $\int_0^1 t^2\bar{u}^2(t)dt = 0$. Como $t^2u^2(t) \geq 0$ entonces $\bar{u}(t) = 0$ c.t.p en $[0,1]$, así $\bar{u} = 0$ en $L^2(0,1)$ en contradicción con el hecho que $\bar{u} \in S$. Por lo tanto f no posee punto minimizante en S.

Veamos a continuación un problema especial de Control Óptimo bajo un sistema de ecuaciones diferenciales.

Sean $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ y $x_0 \in \mathbb{R}^n$. Consideremos el sistema de ecuaciones diferenciales

$$\begin{cases} \dot{x} = Ax(t) + Bu(t) \quad \text{c.t.p en } [t_0, t_1] \\ x(t_0) = x_0 \end{cases} \tag{2.8}$$

donde $-\infty < t_0 < t_1 < +\infty$ y $x(t) \in \mathbb{R}^n$ para cada $t \in [t_0, t_1]$, el vector solución del sistema (2.8). Sea $u \in L^2_m[t_0, t_1]$ una función de control, donde

$$L^2_m[t_0,t_1] = \left\{ w = (w_1(t), w_2(t), \cdots, w_m(t))/w_i \in L^2(t_0,t_1) \right\}$$

La solución del sistema de ecuaciones diferenciales (2.8) viene dada por

$$x(t) = x_0 + \int_{t_0}^{t} e^{(t-s)A}Bu(s)ds$$

donde e^A es la matriz exponencial de A y la integral en el sentido componente a componente. Definimos el conjunto admisible

$$S = \left\{ u \in L^2_m[t_0,t_1]/ \, \|u(t)\| \leq 1 \text{ c.t.p en } [t_0,t_1] \right\} \tag{2.9}$$

donde $\|\cdot\|$ denota la l_2-norma en $\mathbb{R}^n$. Definimos el funcional

$$\begin{aligned} f : S &\longrightarrow \mathbb{R} \\ u &\longrightarrow f(u) = \int_{t_0}^{t_1} g(x(t)) + h(u(t))dt \end{aligned} \tag{2.10}$$

donde $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $h : \mathbb{R}^m \longrightarrow \mathbb{R}$ y el vector $x(t)$ es solución del sistema (2.8). El problema planteado es establecer la existencia de puntos minimizantes de f sobre S. El siguiente Teorema muestra la existencia de tales puntos minimizantes.

Teorema 2.15. Consideremos el problema de minimización

$$\min_{u \in S} f(u) \tag{2.11}$$

donde S, f vienen dadas por (2.9) y (2.10) respectivamente y $x(t)$ dado por el sistema (2.8). Supongamos que:

(*i*) g es convexa y continua.

(*ii*) h es convexa, continua y Lipschitziana en la bola cerrada unitaria de $\mathbb{R}^m$.

Entonces el problema (2.11) admite una solución en S.

Demostración.

Sea $X = L^2_m[t_0, t_1]$. Como X es de Hilbert, entonces es Banach reflexivo. Siendo S la bola unitaria en $L^2_m[t_0, t_1]$ entonces S es cerrado, convexo y acotado. Veamos que la funcional f es cuasiconvexa y continua. En efecto:

- f es cuasiconvexa.

  Definimos

$$\begin{aligned} L : S &\longrightarrow AC^n[t_0, t_1] \\ u &\longrightarrow L(u(t)) = \int_{t_0}^{t} e^{(t-s)A} Bu(s)ds \end{aligned}$$

  donde $AC^n[t_0, t_1]$ es el espacio de las funciones vectoriales absolutamente continuas de $[t_0, t_1]$ en $\mathbb{R}^n$, dotado con la norma del máximo.

  Afirmamos que $L(\lambda u_1 + (1-\lambda)u_2)(t) = \lambda L(u_1(t)) + (1-\lambda)L(u_2(t))$. Sean $u_1, u_2 \in S$ y $\lambda \in [0, 1]$, tenemos

$$\begin{aligned} L(\lambda u_1 + (1-\lambda)u_2)(t) &= \int_{t_0}^{t} e^{(t-s)A} B \left(\lambda u_1(s) + (1-\lambda)u_2(s)\right) ds \\ &= \int_{t_0}^{t} \lambda e^{(t-s)A} Bu_1(s) + (1-\lambda) e^{(t-s)A} Bu_2(s) ds \\ &= \lambda \int_{t_0}^{t} e^{(t-s)A} Bu_1(s) + (1-\lambda) \int_{t_0}^{t} e^{(t-s)A} Bu_2(s) ds \\ &= \lambda L(u_1(t)) + (1-\lambda) L(u_2(t)) \end{aligned} \tag{2.12}$$

  resultando nuestra afirmación. De la igualdad (2.12), obtenemos por la convexidad de g

$$
\begin{aligned}
g(x_0 + L(\lambda u_1 + (1-\lambda)u_2)(t)) &= g(x_0 + \lambda L(u_1(t)) + (1-\lambda)L(u_2(t))) \\
&= g(\lambda\{x_0 + L(u_1(t))\} + (1-\lambda)\{x_0 + L(u_2(t)\}) \quad (2.13)\\
&\leq \lambda g(x_0 + \lambda L(u_1(t))) + (1-\lambda)g(x_0 + L(u_2(t)))
\end{aligned}
$$

Mostrando la convexidad de la funcional $g(x_0 + L(\cdot))$.

Afirmamos que $S_\alpha = \{u \in S / f(u) \leq \alpha\}$ es convexo, para cada $\alpha \in \mathbb{R}$.

En efecto, tomemos $u_1, u_2 \in S_\alpha$ y $\lambda \in [0,1]$. Por (2.13) y la convexidad de las funciones g y h tenemos

$$
\begin{aligned}
f(\lambda u_1 + (1-\lambda)u_2) &= \int_{t_0}^{t_1} g(x_0 + \lambda L(u_1(t)) + (1-\lambda)L(u_2(t))) \\
&\quad + h(\lambda u_1(t) + (1-\lambda)u_2(t))dt \\
&\leq \int_{t_0}^{t_1} \lambda g(x_0 + L(u_1(t))) + (1-\lambda)g(x_0 + L(u_2(t))) \\
&\quad + \lambda h(u_1(t)) + (1-\lambda)h(u_2(t))dt \\
&= \lambda \int_{t_0}^{t_1} g(x_0 + L(u_1(t))) + h(u_1(t))dt + \\
&\quad (1-\lambda)\int_{t_0}^{t_1} g(x_0 + L(u_2(t))) + h(u_2(t))dt \\
&= \lambda f(u_1) + (1-\lambda)f(u_2)
\end{aligned}
$$

Mostrando la convexidad de la funcional f. Luego por el Lema 1.37, S_α es convexo y por la Definición 1.38 tenemos que f es cuasiconvexa.

- f es continua

Para cada $u \in S$ tenemos por definición de L y la norma $L^2_m[t_0, t_1]$

$$
\begin{aligned}
\|L(u)\|_{AC^n[t_0,t_1]} &= \left\| \int_{t_0}^{t} e^{(t-s)A} Bu(s)ds \right\|_{AC^n[t_0,t_1]} \\
&\leq C \|u\|_{L^2_m[t_0,t_1]} \qquad (2.14)
\end{aligned}
$$

donde $C > 0$ esta en términos de las normas matriciales $\|A\|$ y $\|B\|$.

Sean $\bar{u} \in S$ y una sucesión $(u_n)_{n\in\mathbb{N}} \subseteq S$ tales que $u_n \longrightarrow \bar{u}$ si $n \to \infty$, por la definición de f tenemos

$$
\begin{aligned}
f(u_n) - f(\bar{u}) &= \int_{t_0}^{t_1} g(x_0 + L(u_n(t)) + h(u_n(t))dt - \int_{t_0}^{t_1} g(x_0 + L(\bar{u}(t)) + h(\bar{u}(t))dt \\
&= \int_{t_0}^{t_1} g(x_0 + L(u_n(t)) - g(x_0 + L(\bar{u}(t))dt + \int_{t_0}^{t_1} h(u_n(t)) - h(\bar{u}(t))dt
\end{aligned} \tag{2.15}
$$

Por (2.14) y la continuidad de la función g tenemos

$$
\|L(u_n - \bar{u})\|_{AC^n[t_0,t_1]} \leq C \|u_n - \bar{u}\|_{L_2^m[t_0,t_1]} \tag{2.16}
$$

$$
|g(x_0 + L(u_n)(t)) - g(x_0 + L(\bar{u})(t))| \leq L_0 C \|u_n - \bar{u}\|_{L_2^m[t_0,t_1]} \tag{2.17}
$$

donde L_0 es una constante local de Lipschitz.

Si $n \to \infty$ en (2.17) tenemos que $g(x_0+L(u_n)(t) \longrightarrow g(x_0+L(\bar{u})(t)$. Por otra parte, como $\|u_n\|_{L_2^m[t_0,t_1]} \leq 1$ y $\|\bar{u}\|_{L_2^m[t_0,t_1]} \leq 1$, tenemos que el segundo miembro de la desigualdad (2.17) es acotado. Luego por el teorema de la convergencia dominada (Ver Jost [9], pag. 209, Teorema 16.5) obtenemos la convergencia

$$
\int_{t_0}^{t_1} g(x_0 + L(u_n)(t))dt \longrightarrow \int_{t_0}^{t_1} g(x_0 + L(\bar{u})(t))dt, \text{ si } n \to \infty
$$

es decir

$$
\int_{t_0}^{t_1} g(x_0 + L(u_n(t)) - g(x_0 + L(\bar{u}(t))dt \longrightarrow 0, \text{ si } n \to \infty \tag{2.18}
$$

Como h es Lipschitziana, tenemos para la segunda integral de (2.15)

$$
\begin{aligned}
\int_{t_0}^{t_1} |h(u_n(t)) - h(\bar{u}(t))|\, dt &\leq C_1 \int_{t_0}^{t_1} \|u_n(t) - \bar{u}(t)\|\, dt \\
&\leq C_1 \|u_n(t) - \bar{u}(t)\|_{L_2^m[t_0,t_1]}
\end{aligned} \tag{2.19}
$$

Tomando $n \to \infty$ en (2.19) tenemos

$$
\int_{t_0}^{t_1} |h(u_n(t)) - h(\bar{u}(t))|\, dt \longrightarrow 0 \tag{2.20}
$$

De (2.15), (2.18) y (2.20) tenemos que $f(u_n) \longrightarrow f(\bar{u})$ si $n \to \infty$. Por el Teorema 1.11 f es continua en $\bar{u}$, por lo tanto f es continua en S.

Así, f es continua y cuasiconvexa y siendo S un conjunto no vacío, acotado, convexo y cerrado en el espacio reflexivo $L_2^m[t_0, t_1]$, por el Teorema 2.8 existe $\tilde{u} \in S$ tal que $f(\tilde{u}) = \min_{u \in S} F(u)$.

$\square$

Capítulo 3

Condiciones de optimalidad diferenciable y no diferenciable

En este capítulo revisaremos los conceptos de la derivada direccional, las derivadas de Fréchet y Gateaux en espacios normados. Asimismo, el importante concepto del subgradiente de una funcional, estableciendo condiciones necesarias y suficientes para puntos minimizantes.

3.1. Derivadas direccionales en espacios normados

Definición 3.1. Consideremos X un espacio vectorial. Sean $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una aplicación. Decimos que f es derivable en $\bar{x} \in S$ en la dirección $h \in X$, si existe el límite

$$f'(\bar{x})h := \lim_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda h) - f(\bar{x})\}$$

Decimos que f es *direccionalmente diferenciable* en $\bar{x}$ si el límite anterior existe para todo $h \in X$.

Observaciones.

(*i*) Si $h = 0$ entonces $f'(\bar{x})h = 0$, para todo $\bar{x} \in S$.

(*ii*) En términos secuenciales si consideramos una sucesión $(\lambda_n) \subseteq \mathbb{R}^+$ entonces $\bar{x} + \lambda_n h \in S$ para todo $n \in \mathbb{N}$, problema que se evita si $S = X$.

Ejemplo.

Sea la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$f(x_1, x_2) = \begin{cases} x_1^2 + \dfrac{x_1^2}{x_2}, & \text{si } x_2 \neq 0 \\ 0, & \text{si } x_2 = 0 \end{cases}$$

Evidentemente f no es continua en $(0,0)$. Sea $h = (h_1, h_2) \in \mathbb{R}^2$ una dirección arbitraria.

Si $h_2 \neq 0$ entonces

$$\begin{aligned} f'(0_{\mathbb{R}^2})(h_1, h_2) &= \lim_{\lambda \to 0^+} \frac{1}{\lambda} f(\lambda(h_1, h_2)) \\ &= \lim_{\lambda \to 0^+} \frac{1}{\lambda} \left\{ (\lambda h_1)^2 + \frac{(\lambda h_1)^2}{\lambda h_2} \right\} \\ &= \lim_{\lambda \to 0^+} \lambda h_1^2 + \frac{h_1^2}{h_2} = \frac{h_1^2}{h_2} \end{aligned}$$

Si $h_2 = 0$ tenemos

$$\begin{aligned} f'(0_{\mathbb{R}^2})(h_1, h_2) &= \lim_{\lambda \to 0^+} \frac{1}{\lambda} f(\lambda h_1, 0) \\ &= \lim_{\lambda \to 0^+} \frac{1}{\lambda} \{0\} = 0 \end{aligned}$$

Por lo tanto

$$f'(0_{\mathbb{R}^2})(h_1, h_2) = \begin{cases} \dfrac{h_1^2}{h_2}, & \text{si } h_2 \neq 0 \\ 0, & \text{si } h_2 = 0 \end{cases}$$

Observemos en este ejemplo que $f'(0_{\mathbb{R}^2})$ es una aplicación continua y no lineal.

Lema 3.2. Consideremos X un espacio vectorial y $f : X \longrightarrow \mathbb{R}$ una funcional convexa. Si $\bar{x}$ y h pertenecen a X, entonces la función

$$\begin{aligned} \varphi : \mathbb{R}^+ \setminus \{0\} &\longrightarrow \mathbb{R} \\ \lambda &\longrightarrow \varphi(\lambda) = \frac{1}{\lambda} \{f(\bar{x} + \lambda h) - f(\bar{x})\} \end{aligned}$$

es monótona creciente, es decir si $0 < s \leq t$ entonces $\varphi(s) \leq \varphi(t)$.

Demostración.

Sean $0 < s \leq t$. Por la convexidad de f tenemos

$$\begin{aligned} f(\bar{x} + sh) - f(\bar{x}) &= f\left(\frac{s}{t}(\bar{x} + th) - \frac{t-s}{t}\bar{x}\right) - f(\bar{x}) \\ &\leq \frac{s}{t} f(\bar{x} + th) + \frac{t-s}{t} f(\bar{x}) - f(\bar{x}) \\ &= \frac{s}{t} \{f(\bar{x} + th) - f(\bar{x})\} \end{aligned}$$

Así, obtenemos

$$\frac{f(\bar{x}+sh)-f(\bar{x})}{s} \leq \frac{f(\bar{x}+th)-f(\bar{x})}{t}$$

Por lo tanto $\varphi(s) \leq \varphi(t)$.

$\square$

Lema 3.3. Sean $X \subseteq \mathbb{R}$ y $f : X \longrightarrow \mathbb{R}$ una función monótona y acotada. Si $a \in X'_+$ (el conjunto derivado lateral derecho de X definido por $X'_+ = \{a \in \mathbb{R}/(a, a+\epsilon) \cap X \neq \emptyset \text{ para todo } \varepsilon > 0\}$) entonces existe $\lim\limits_{x \to a^+} f(x)$.

Demostración. Ver Lima [12] (*pag. 167, Teorema 12*).

$\square$

Lema 3.4. Consideremos X un espacio vectorial y $f : X \longrightarrow \mathbb{R}$ una funcional convexa. Para cada $\bar{x} \in X$, dada cualquier dirección $h \in X$, existe la derivada direccional $f'(\bar{x})h$, es decir f es direccionalmente diferenciable en X.

Demostración.

Sean $\bar{x}, h \in X$. Definimos la aplicación $\varphi : \mathbb{R}^+ \longrightarrow \mathbb{R}$ por

$$\varphi(\lambda) = \frac{1}{\lambda}\left\{f(\bar{x}+\lambda h) - f(\bar{x})\right\}$$

Para $\lambda > 0$ obtenemos por la convexidad de f

$$\begin{aligned} f(\bar{x}) &= f\left(\frac{1}{1+\lambda}(\bar{x}+\lambda h) + \frac{\lambda}{1+\lambda}(\bar{x}-h)\right) \\ &\leq \frac{1}{1+\lambda} f(\bar{x}+\lambda h) + \frac{\lambda}{1+\lambda} f(\bar{x}-h) \end{aligned} \tag{3.1}$$

Luego, de (3.1) tenemos

$$(1+\lambda) f(\bar{x}) \leq f(\bar{x}+\lambda h) + \lambda f(\bar{x}-h)$$

De la desigualdad anterior, transponiendo términos conseguimos

$$\begin{aligned} f(\bar{x}) - f(\bar{x}-h) &\leq \frac{1}{\lambda}\left\{f(\bar{x}+\lambda h) - f(\bar{x})\right\} \\ &= \varphi(\lambda) \end{aligned} \tag{3.2}$$

Como el primer miembro de la desigualdad (3.2) no depende de λ, tenemos que φ es acotada inferiormente, además por el Lema 3.2 es monótona creciente. Así por el Lema 3.3 existe $\lim\limits_{\lambda \to 0+} \varphi(\lambda)$, en consecuencia por la Definición 3.1, existe $f'(\bar{x})h$ para todo $\bar{x}$ y h en X.

$\square$

Lema 3.5. Sean X un espacio vectorial real y $f : X \longrightarrow \mathbb{R}$ una funcional convexa. Entonces para cada $\bar{x} \in X$, la aplicación

$$\begin{aligned} f'(\bar{x}) : X &\longrightarrow \mathbb{R} \\ h &\longrightarrow f'(\bar{x})h \end{aligned}$$

es una funcional sublineal.

Demostración.

Por el Lema 3.4 existe $f'(\bar{x})h$ para todo $\bar{x}, h \in X$. Observemos que $f'(\bar{x})0_X = 0$ y en consecuencia $f'(\bar{x})(\alpha 0_X) = 0$ para todo $\alpha > 0$. Si $h \in X$ y $\alpha > 0$ entonces

$$\begin{aligned} f'(\bar{x})(\alpha h) &= \lim_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda\alpha h) - f(\bar{x})\} \\ &= \alpha \lim_{\lambda \to 0^+} \frac{1}{\alpha\lambda} \{f(\bar{x} + \lambda\alpha h) - f(\bar{x})\} \\ &= \alpha f'(\bar{x})h \end{aligned}$$

Mostrando así que $f'(\bar{x})(\alpha h) = \alpha f'(\bar{x})h$ para cada $h \in X$ y $\alpha > 0$.

Sean $h_1, h_2 \in X$ y $\alpha > 0$, por la convexidad de f tenemos

$$\begin{aligned} f(\bar{x} + \lambda(h_1 + h_2)) &= f\left(\frac{1}{2}(\bar{x} + 2\lambda h_1) + \frac{1}{2}(\bar{x} + 2\lambda h_2)\right) \\ &\leq \frac{1}{2} f(\bar{x} + 2\lambda h_1) + \frac{1}{2} f(\bar{x} + 2\lambda h_2) \end{aligned} \tag{3.3}$$

Luego de la desigualdad (3.3) y agregando el término $-f(\bar{x})$ a ambos miembros, obtenemos

$$f(\bar{x} + \lambda(h_1 + h_2)) - f(\bar{x}) \leq \frac{1}{2} f(\bar{x} + 2\lambda h_1) - \frac{1}{2} f(\bar{x}) + \frac{1}{2} f(\bar{x} + 2\lambda h_2) - \frac{1}{2} f(\bar{x}) \tag{3.4}$$

Multiplicando (3.4) por $\dfrac{1}{\lambda}$ a ambos miembros, resulta

$$\begin{aligned} \frac{1}{\lambda} \{f(\bar{x} + \lambda(h_1 + h_2)) - f(\bar{x})\} &\leq \frac{1}{2\lambda} \{f(\bar{x} + 2\lambda h_1) - f(\bar{x})\} \\ &\quad + \frac{1}{2\lambda} \{f(\bar{x} + 2\lambda h_2) - f(\bar{x})\} \end{aligned} \tag{3.5}$$

Tomando límite en (3.5) cuando $\lambda \to 0^+$, por la Definición 3.1 logramos

$$f'(\bar{x})(h_1 + h_2) \leq f'(\bar{x})(h_1) + f'(\bar{x})(h_2)$$

Demostrando así la sublinealidad de la aplicación $f'(\bar{x})$.

$\square$

Observación.

Si una funcional f no esta definida en todo el espacio X, tan sólo esta definida en un subconjunto no vacío S, requerimos que $\bar{x} + \lambda(x - \bar{x}) \in S$ para un $\lambda > 0$ suficientemente pequeño a fin de encontrar la derivada direccional de f en la dirección $x - \bar{x}$. Gracias a la Definición 1.27 es suficiente considerar a S como un conjunto estrellado.

El siguiente resultado nos muestra dos condiciones de optimalidad para funcionales direccionalmente diferenciables.

Teorema 3.6. Sean X un espacio normado, $S \subseteq X$ un subconjunto no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Se cumplen:

(i) Si $\bar{x} \in S$ es punto mínimo de f en S y la funcional f posee derivadas direccionales en $\bar{x}$ respecto a toda dirección $x - \bar{x}$ ($x \in S$), entonces $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$.

(ii) Supongamos que S es convexo y f una funcional convexa. Si f posee derivadas direccionales en $\bar{x}$ en toda dirección $x - \bar{x}$ ($x \in S$) y se satisface que $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$ entonces $\bar{x}$ es punto mínimo de f en S.

Demostración.

(i) Sea $x \in S$. Como f es direccionalmente diferenciable en $\bar{x}$ en la dirección $x - \bar{x}$ tenemos

$$f'(\bar{x})(x - \bar{x}) = \lim_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})\}$$

Además, por hipótesis $\bar{x}$ es punto minimizante de f en S, es decir existe $\lambda_0 > 0$ tal que $f(\bar{x}+\lambda(x-\bar{x})) \geq f(\bar{x})$ para todo $\lambda \in (0, \lambda_0)$. Gracias a esta desigualdad, si $\lambda \to 0^+$ obtenemos

$$\begin{aligned} f'(\bar{x})(x - \bar{x}) &= \lim_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})\} \\ &\geq 0 \end{aligned}$$

Con lo cual $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$.

(ii) Sean $x \in S$ y $0 \leq \lambda \leq 1$. Por la convexidad de f y S tenemos

$$\begin{aligned} f(\bar{x} + \lambda(x - \bar{x})) &= f(\lambda x + (1 - \lambda)\bar{x}) \\ &\leq \lambda f(x) + (1 - \lambda)f(\bar{x}) \end{aligned}$$

Obteniendo de esta última desigualdad

$$f(\bar{x}+\lambda(x-\bar{x}))-f(\bar{x}) \leq \lambda f(x)-\lambda f(\bar{x}) \tag{3.6}$$

Transponiendo los términos convenientemente en (3.6), logramos

$$f(x)-f(\bar{x}) \geq \frac{1}{\lambda}\left\{f(\bar{x}+\lambda(x-\bar{x}))-f(\bar{x})\right\}$$

Tomando el límite cuando $\lambda \to 0^+$ en la desigualdad anterior y por hipótesis obtenemos

$$\begin{aligned} f(x)-f(\bar{x}) &\geq \lim_{\lambda\to 0^+}\frac{1}{\lambda}\left\{f(\bar{x}+\lambda(x-\bar{x}))-f(\bar{x})\right\} \\ &= f'(\bar{x})(x-\bar{x}) \geq 0 \end{aligned}$$

Por lo tanto, $f(x) \geq f(\bar{x})$ para todo $x \in S$ obteniendo el resultado.

□

3.2. Funcionales Gateaux y Fréchet diferenciables

Definición 3.7. Sean $(X, \| \|_X)$ e $(Y, \| \|_Y)$ dos espacios normados, $S \subseteq X$ un subconjunto abierto no vacío y $f : S \longrightarrow Y$ una aplicación. Decimos que f es *Gateaux diferenciable* en $\bar{x} \in S$ si la aplicación

$$f'(\bar{x})h := \lim_{\lambda\to 0}\frac{1}{\lambda}\left\{f(\bar{x}+\lambda h)-f(\bar{x})\right\}$$

existe para todo $h \in X$ y $f'(\bar{x}) \in L(X,Y)$. La aplicación $f'(\bar{x})$ se denomina la *derivada de Gateaux* de f en $\bar{x}$.

Ejemplos.

- Sean $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase $C^1(\mathbb{R}^n)$ y $\bar{x}, h \in \mathbb{R}^n$. Tenemos

$$\begin{aligned} f'(\bar{x})h &= \lim_{\lambda\to 0}\frac{1}{\lambda}\left\{f(\bar{x}+\lambda h)-f(\bar{x})\right\} \\ &= \frac{d}{d\lambda}f(\bar{x}+\lambda h)\mid_{\lambda=0} \\ &= \nabla f(\bar{x}+\lambda h)^T.h\mid_{\lambda=0} \\ &= \nabla f(\bar{x})^T.h \end{aligned}$$

para todo $h \in \mathbb{R}^n$. Por lo tanto f es Gateaux diferenciable en $\bar{x} \in \mathbb{R}^n$.

- Sean $(X, \| \|_X)$ e $(Y, \| \|_Y)$ dos espacios normados y $l \in L(X,Y)$. Si $\bar{x}, h \in X$ entonces

$$\begin{aligned} l'(\bar{x})h &= \lim_{\lambda \to 0} \frac{1}{\lambda} \{l(\bar{x} + \lambda h)) - l(\bar{x})\} \\ &= \lim_{\lambda \to 0} \frac{1}{\lambda} \{l(\lambda h)\} = l(h) \end{aligned}$$

es decir, $l'(\bar{x})h = l(h)$ para todo $h \in X$, por lo tanto l es Gateaux diferenciable en $\bar{x}$ y $l'(\bar{x}) = l$.

Definición 3.8. Sean $(X, \| \|_X)$ e $(Y, \| \|_Y)$ dos espacios normados, $S \subseteq X$ un subconjunto abierto no vacío y $f : S \longrightarrow Y$ una aplicación. Decimos que f es *Fréchet diferenciable* en $\bar{x} \in S$ si existe una aplicación lineal continua $f'(\bar{x}) : X \longrightarrow Y$ tal que

$$\lim_{\|h\|_X \to 0} \frac{\|f(\bar{x} + h) - f(\bar{x}) - f'(\bar{x})h\|_Y}{\|h\|_X} = 0$$

$f'(\bar{x})$ será llamada la *derivada de Fréchet* de f en $\bar{x}$.

De acuerdo a esta definición, denotaremos

$$f(\bar{x} + h) = f(\bar{x}) + f'(\bar{x})h + \circ(\|h\|_X)$$

donde $\circ(\|h\|_X)$ posee la propiedad

$$\lim_{\|h\|_X \to 0} \frac{\circ(\|h\|_X)}{\|h\|_X} = 0_Y$$

Ejemplo.

Consideremos $l : \mathbb{R}^3 \longrightarrow \mathbb{R}$ una función continua con respecto a sus tres variables y con derivadas parciales continuas con respecto a sus dos primeras variables. Sea la funcional $(-\infty < a < b < +\infty)$

$$\begin{aligned} f : C^1[a,b] &\longrightarrow \mathbb{R} \\ x &\longrightarrow f(x) = \int_a^b l(x(t), x'(t), t)dt \end{aligned}$$

Si $\bar{x}, h \in C^1[a,b]$, tenemos

$$\begin{aligned} f(\bar{x} + h) - f(\bar{x}) &= \int_a^b l((\bar{x} + h)(t), (\bar{x} + h)'(t), t)dt - \int_a^b l(x(t), x'(t), t)dt \\ &= \int_a^b l(\bar{x}(t) + h(t), \bar{x}'(t) + h'(t), t)dt - \int_a^b l(x(t), x'(t), t)dt \end{aligned}$$

Como l posee derivadas parciales continuas con respecto a las dos primeras variables, tenemos que

$$f(\bar{x}+h) - f(\bar{x}) = \int_a^b \left\{ l_x(\bar{x}(t), \bar{x}'(t), t)h(t) + l_{x'}(\bar{x}(t), \bar{x}'(t), t)h'(t) \right\} dt + \circ(\|h\|_{C^1[a,b]})$$

Por lo tanto, f es Fréchet diferenciable y además

$$f'(\bar{x})h = \int_a^b \left\{ l_x(\bar{x}(t), \bar{x}'(t), t)h(t) + l_{x'}(\bar{x}(t), \bar{x}'(t), t)h'(t) \right\} dt$$

Teorema 3.9. Sean $(X, \| \|_X)$ e $(Y, \| \|_Y)$ dos espacios normados, $S \subseteq X$ un subconjunto abierto no vacío y $f : S \longrightarrow Y$ una aplicación. Si f es Fréchet diferenciable en $\bar{x} \in S$, entonces f es Gateaux diferenciable en $\bar{x}$ y las derivadas de Gateaux y Fréchet coinciden.

Demostración.

Supongamos que f es Fréchet diferenciable en $\bar{x} \in S$. Por la Definición 3.8 tenemos

$$\lim_{\lambda \to 0} \frac{\|f(\bar{x}+\lambda h) - f(\bar{x}) - f'(\bar{x})(\lambda h)\|_Y}{\|\lambda h\|_X} = 0 \quad \text{para todo } h \in X \setminus \{0\}$$

Así,

$$\begin{aligned} &\lim_{\lambda \to 0} \frac{1}{|\lambda|} \left\| f(\bar{x}+\lambda h) - f(\bar{x}) - f'(\bar{x})(\lambda h) \right\|_Y = 0 \quad \text{para todo } h \in X \setminus \{0\} \\ \Longrightarrow &\lim_{\lambda \to 0} \frac{1}{\lambda} \left\{ f(\bar{x}+\lambda h) - f(\bar{x}) - f'(\bar{x})(\lambda h) \right\} = 0 \quad \text{para todo } h \in X \setminus \{0\} \end{aligned} \tag{3.7}$$

Como $f'(\bar{x})$ es lineal, de la implicación dada en (3.7) tenemos

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \{f(\bar{x}+\lambda h) - f(\bar{x})\} = f'(\bar{x})h \quad \text{para todo } h \in X \setminus \{0\} \tag{3.8}$$

Por otra parte, como $f'(\bar{x})0_X = 0_Y$ concluimos que

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \{f(\bar{x}+\lambda h) - f(\bar{x})\} = f'(\bar{x})h \quad \text{para todo } h \in X$$

Por lo tanto f es Gateaux diferenciable en $\bar{x}$ y $f'_G(\bar{x})h = f'(\bar{x})h$ para todo $h \in X$.

□

Corolario 3.10. Sean $(X, \| \|_X)$ e $(Y, \| \|_Y)$ dos espacios normados, $S \subseteq X$ un subconjunto abierto no vacío y $f : S \longrightarrow Y$ una aplicación. Si f es Fréchet diferenciable en $\bar{x} \in S$ entonces:

(i) La derivada de Fréchet es única.

(ii) f es continua en $\bar{x}$.

Demostración.

(i) Por el Teorema 3.9 las derivadas de Fréchet y Gateaux coinciden en $\bar{x}$. Como la derivada de Gateaux es única (por unicidad del límite), la derivada de Fréchet es única.

(ii) Como f es Fréchet diferenciable en $\bar{x} \in S$, por la Definición 3.8 existe el límite

$$\lim_{\|h\|_X \to 0} \frac{\|f(\bar{x}+h) - f(\bar{x}) - f'(\bar{x})h\|_Y}{\|h\|_X}$$

y es igual a 0. Dado $\epsilon > 0$ existe $\delta > 0$ tal que $\bar{x} + h \in B(\bar{x}, \delta)$ y

$$\left\|f(\bar{x}+h) - f(\bar{x}) - f'(\bar{x})h\right\|_Y \leq \epsilon \left\|h\right\|_X \tag{3.9}$$

Desde que $f'(\bar{x}) \in L(X, Y)$, existe $\alpha > 0$ tal que $\|f'(\bar{x})h\|_Y \leq \alpha \|h\|_X$. De (3.9) y la desigualdad anterior conseguimos

$$\begin{aligned} \|f(\bar{x}+h) - f(\bar{x})\|_Y &= \left\|f(\bar{x}+h) - f(\bar{x}) - f'(\bar{x})h + f'(\bar{x})h\right\|_Y \\ &\leq \left\|f(\bar{x}+h) - f(\bar{x}) - f'(\bar{x})h\right\|_Y + \left\|f'(\bar{x})h\right\|_Y \\ &\leq \epsilon \left\|h\right\|_X + \alpha \left\|h\right\|_X \end{aligned} \tag{3.10}$$

Si $\|h\|_X \to 0$ logramos en (3.10)

$$\lim_{\|h\|_X \to 0} \|f(\bar{x}+h) - f(\bar{x})\|_Y = 0$$

Mostrando la continuidad de f en $\bar{x}$.

$\square$

El siguiente teorema caracteriza a las funciones convexas Gateaux diferenciables.

Teorema 3.11. Sean X un e.v.n, $S \subseteq X$ un abierto convexo no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Si f es Gateaux diferenciable en todo $x \in S$ entonces f es convexa si y sólo si $f(y) \geq f(x) + f'(x)(y - x)$ para todo $x, y \in S$.

Demostración.

$\Longrightarrow)$

Supongamos que f es convexa. Tomemos $x, y \in S$ y $0 \leq \lambda \leq 1$, se tiene que

$$\begin{aligned} f(x + \lambda(y - x)) &= f(\lambda y + (1 - \lambda)x) \\ &\leq \lambda f(y) + (1 - \lambda)f(x) \end{aligned}$$

Transponiendo términos conseguimos $\lambda(f(y) - f(x)) \geq f(x + \lambda(y - x)) - f(x)$. Por lo tanto,

$$f(y) \geq f(x) + \frac{1}{\lambda}\{f(x + \lambda(y - x)) - f(x)\} \tag{3.11}$$

Como f es Gateaux diferenciable, tomando el límite cuando $\lambda \to 0$ en (3.11) por la Definición 3.7 conseguimos $f(y) \geq f(x) + f'(x)(y - x)$.

$\Longleftarrow$)

Supongamos que $f(y) \geq f(x) + f'(x)(y - x)$ para todo $x, y \in S$. Por la convexidad de S si $0 \leq \lambda \leq 1$ entonces

$$\begin{aligned} f(x) &\geq f(\lambda x + (1 - \lambda)y) + f'(\lambda y + (1 - \lambda)x)(x - (\lambda y + (1 - \lambda)x)) \\ &= f(\lambda x + (1 - \lambda)y) + f'(\lambda y + (1 - \lambda)x)((1 - \lambda)(x - y)) \end{aligned} \tag{3.12}$$

Análogamente,

$$f(y) \geq f(\lambda x + (1 - \lambda)y) + f'(\lambda y + (1 - \lambda)x)(-\lambda(x - y)) \tag{3.13}$$

Denotemos por $u = \lambda x + (1 - \lambda)y$. Multiplicando por λ y $(1 - \lambda)$ a (3.12) y (3.13) respectivamente, sumando las desigualdades resultantes y por la linealidad de $f'(\cdot)$ tenemos

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) &\geq \lambda f(u) + \lambda f'(u)((1 - \lambda)(x - y)) + (1 - \lambda)f(u) \\ &\quad + (1 - \lambda)f'(u)(-\lambda(x - y)) \\ &= f(u) \end{aligned} \tag{3.14}$$

Por lo tanto, de la desigualdad (3.14) y la notación de u tenemos $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$, mostrando la convexidad de f.

$\square$

Ejemplo.

Sea $S \subseteq \mathbb{R}^n$ abierto y convexo y $f : S \longrightarrow \mathbb{R}$ una función de clase $C^1(S)$. Por el teorema anterior tenemos que $f(y) \geq f(x) + \nabla f(x)^T(y - x)$ para todo $x, y \in S$.

Teorema 3.12. Sean X un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional Gateaux diferenciable en $\bar{x}$. Si $\bar{x}$ es punto mínimo de f en X entonces $f'(\bar{x})h = 0$ para todo $h \in X$.

Demostración.

Sea $h \in X$. Como $\bar{x}$ es punto mínimo de f en S y f es Gateaux diferenciable en $\bar{x}$, por el Teorema 3.6 -(i) se tiene que $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$. En particular para $x = h + \bar{x}$ tenemos que $f'(\bar{x})h \geq 0$. Asimismo, para $x = -h + \bar{x}$ se sigue que $f'(\bar{x})(-h) \geq 0$ para todo $h \in X$, de esta última desigualdad por la linealidad de $f'(\bar{x})$ tenemos que $f'(\bar{x})h \leq 0$. En consecuencia $f'(\bar{x})h = 0$ para todo $h \in X$.

$\square$

3.3. Aplicación a un problema del cálculo variacional

En esta sección aplicaremos el resultado expuesto en el Teorema 3.12 a un problema del Cálculo de Variaciones.

Consideremos $l : \mathbb{R}^3 \longrightarrow \mathbb{R}$ una función continua respecto a todas sus variables, con derivadas parciales continuas respecto a sus dos primeras variables. Sea la funcional

$$\begin{array}{rcl} f : C^1[a,b] & \longrightarrow & \mathbb{R} \\ x & \longrightarrow & f(x) = \displaystyle\int_a^b l(x(t), x'(t), t)dt \end{array}$$

donde $-\infty < a < b < +\infty$. Definimos las condiciones de transversalidad $x(a) = x_1$ y $x(b) = x_2$ con $x_1, x_2 \in \mathbb{R}$ y el conjunto factible

$$S = \left\{ x \in C^1[a,b] / x(a) = x_1 \wedge x(b) = x_2 \right\}$$

Claramente, S es un conjunto convexo en $C^1[a,b]$. Supongamos que $\bar{x} \in S$ es el punto minimizante de f. Por el ejemplo dado en la Definición 3.8, f es Fréchet diferenciable en $\bar{x}$; por los Teoremas 3.9 y 3.6-(i) tenemos que f es Gateaux diferenciable en $\bar{x}$ y $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$. Esta última desigualdad es equivalente a $f'(\bar{x})h \geq 0$ para todo $h \in \tilde{S} = S - \{\bar{x}\}$. Afirmamos que $\tilde{S} = \left\{ x \in C^1[a,b] / x(a) = 0 \wedge x(b) = 0 \right\}$. En efecto, si $x \in \tilde{S}$ entonces $x = \hat{x} - \bar{x}$ con $\hat{x} \in S$. Luego,

$$\begin{aligned} x(a) &= \tilde{x}(a) - \bar{x}(a) = x_1 - x_1 = 0 \\ x(b) &= \tilde{x}(b) - \bar{x}(b) = x_2 - x_2 = 0 \end{aligned}$$

En consecuencia, $x \in \left\{ x \in C^1[a,b] / x(a) = 0 \wedge x(b) = 0 \right\}$. Como la otra inclusión es evidente, se tiene la afirmación. Además, si $h \in \tilde{S}$ entonces $-h \in \tilde{S}$, en consecuencia $f'(\bar{x})(-h) \geq 0$ y por la linealidad de $f'(\bar{x})$ tenemos que $f'(\bar{x})h \leq 0$ para todo $h \in \tilde{S}$, obteniendo así la identidad $f'(\bar{x})h = 0$ para todo $h \in \tilde{S}$.

Por el Teorema 3.9, la derivada de Gateaux de f en $\bar{x}$ esta dada por

$$f'(\bar{x})h = \int_a^b \left\{ \frac{\partial l}{\partial x}(\bar{x}(t), \bar{x}'(t), t)h(t) + \frac{\partial l}{\partial x'}(\bar{x}(t), \bar{x}'(t), t)h'(t) \right\} dt$$

Siendo $f'(\bar{x})h = 0$ tenemos finalmente que

$$\int_a^b \left\{ \frac{\partial l}{\partial x}(\bar{x}(t), \bar{x}'(t), t)h(t) + \frac{\partial l}{\partial x'}(\bar{x}(t), \bar{x}'(t), t)h'(t) \right\} dt = 0 \tag{3.15}$$

para todo $h \in \tilde{S}$. Para obtener el resultado precisamos de los siguientes lemas.

Lema 3.13. Sea el conjunto $\tilde{S} = \{x \in C^1[a,b]/x(a) = 0 \wedge x(b) = 0\}$ con $-\infty < a < b < +\infty$. Si existe $x \in C[a,b]$ tal que $\int_a^b x(t)h'(t)dt = 0$ para todo $h \in \tilde{S}$, entonces x es constante en $[a,b]$.

Demostración.

Sean

$$c = \frac{1}{b-a}\int_a^b x(t)dt \quad y \quad h(t) = \int_a^t (x(s)-c)ds \tag{3.16}$$

Observemos que c es constante y $h'(t) = x(t) - c$, así $x \in C[a,b]$. Además se cumplen que $h(a) = 0$ y

$$\begin{aligned} h(b) &= \int_a^b (x(s)-c)ds \\ &= \int_a^b x(s) - c(b-a) \\ &= 0 \end{aligned} \tag{3.17}$$

Resultando que $h \in \tilde{S}$. Así por la hipótesis sobre x en $\tilde{S}$, el Teorema Fundamental del Cálculo y las identidades (3.16) y (3.17) obtenemos

$$\begin{aligned} \int_a^b (x(t)-c)^2 dt &= \int_a^b (x(t)-c)h'(t)dt \\ &= \int_a^b x(t)h'(t)dt - c(h(b)-h(a)) \\ &= -ch(b) \\ &= 0 \end{aligned}$$

Por lo tanto como $(x(t)-c)^2$ es continua tenemos que $x(t) = c$ para todo $t \in [a,b]$. En consecuencia, x es constante.

□

Lema 3.14. Bajo las hipótesis del Lema 3.13, si existen $x, y \in \tilde{S}$ tales que

$$\int_a^b \{x(t)h(t) + y(t)h'(t)\}dt = 0$$

para todo $h \in \tilde{S}$. Entonces $y \in C^1[a,b]$ y $y'(t) = x(t)$ para todo $t \in [a,b]$.

Demostración.

Definimos

$$\begin{aligned} \varphi : [a,b] &\longrightarrow \mathbb{R} \\ t &\longrightarrow \varphi(t) = \int_a^t x(s)ds \end{aligned}$$

Sea $h \in \tilde{S}$, por integración por partes, tenemos

$$\begin{aligned} \int_a^b x(t)h(t)dt &= \varphi(t)h(t) \mid_a^b - \int_a^b \varphi(t)h'(t)dt \\ &= \varphi(b)h(b) - \varphi(a)h(a) - \int_a^b \varphi(t)h'(t)dt \\ &= -\int_a^b \varphi(t)h'(t)dt \end{aligned} \tag{3.18}$$

De la hipótesis del teorema y la identidad (3.18) conseguimos

$$\int_a^b (-\varphi(t) + y(t))h'(t)dt = 0$$

para todo $h \in \tilde{S}$. Por el Lema 3.13 existe $c \in \mathbb{R}$ tal que $y(t) - \varphi(t) = c$ para todo $t \in [a,b]$. Así $y(t) = \varphi(t) + c$, con lo cual $y \in C^1[a,b]$ y además $y'(t) = x(t)$ para todo $t \in [a,b]$ quedando demostrado el lema.

□

Usando el Lema 3.14 obtenemos el siguiente resultado, conocido como el Teorema Fundamental del Cálculo Variacional.

Teorema 3.15. Sea $l : \mathbb{R}^3 \longrightarrow \mathbb{R}$ una función continua con respecto a todas sus variables y con derivadas parciales continuas respecto a sus dos primeras variables. Consideremos la funcional

$$\begin{aligned} f : C^1[a,b] &\longrightarrow \mathbb{R} \\ x &\longrightarrow f(x) = \int_a^b l(x(t), x'(t), t)dt \end{aligned}$$

Si $S = \{x \in C^1[a,b] / x(a) = x_1 \wedge x(b) = x_2\}$ y $\bar{x}$ es el punto mínimo de f en S entonces

$$\frac{d}{dt}\{l_{x'}(\bar{x}(t), \bar{x}'(t), t)\} = l_x(\bar{x}(t), \bar{x}'(t), t)$$

para todo $t \in [a,b]$.

Demostración.

Sea $\bar{x}$ el punto mínimo de f en S. Por (3.15) tenemos

$$\int_a^b \left\{ l_x(\bar{x}(t), \bar{x}'(t), t)h(t) + l_{x'}(\bar{x}(t), \bar{x}'(t), t)h'(t) \right\} dt = 0$$

para todo $h \in S$. Por lo tanto por el Lema 3.14 para $x(t) = l_x(\bar{x}(t), \bar{x}'(t), t)$ e $y(t) = l_{x'}(\bar{x}(t), \bar{x}'(t), t)$ conseguimos

$$\frac{d}{dt} \left\{ l_{x'}(\bar{x}(t), \bar{x}'(t), t) \right\} = l_x(\bar{x}(t), \bar{x}'(t), t)$$

para cada $t \in [a,b]$, logrando la prueba del Teorema.

□

Ejemplo.

Determinaremos la curva $x \in C^1[a,b]$ que une los puntos (a, x_1) y (b, x_2) con la menor longitud. En otras palabras debemos encontrar el punto minimizante de la funcional

$$f(x) = \int_a^b \sqrt{1 + (x'(t))^2} dt$$

sobre $S = \left\{ x \in C^1[a,b] / x(a) = x_1 \wedge x(b) = x_2 \right\}$. En este caso la funcional l viene dada por $l(x, x', t) = \sqrt{1 + x'^2}$ la cual es continua y con la derivada parcial respecto a la variable x' continua, puesto que $l_{x'} = \dfrac{x'}{\sqrt{1 + x'^2}}$ y ademas $l_x = 0$. Si $\bar{x}$ es el punto minimizante de f por el Teorema 3.15 tenemos

$$\begin{aligned} \frac{d}{dt} \left\{ l_{x'}(\bar{x}(t), \bar{x}'(t), t) \right\} &= l_x(\bar{x}(t), \bar{x}'(t), t) \\ &= 0 \end{aligned}$$

Es decir $\dfrac{d}{dt} \left\{ \dfrac{\bar{x}'}{\sqrt{1 + \bar{x}'^2}} \right\} = 0$ para todo $t \in [a,b]$, en consecuencia $\dfrac{\bar{x}'}{\sqrt{1 + \bar{x}'^2}} = c$ donde c es una constante. Despejando $\bar{x}'$ conseguimos

$$\bar{x}' = \sqrt{\frac{c^2}{1 - c^2}} = k$$

De esta última identidad, resulta que $\bar{x}'(t) = k$ e integrando tenemos $\bar{x}(t) = kt + k_1$. Por las condiciones de transversalidad $\bar{x}(a) = x_1$ y $\bar{x}(b) = x_2$ concluimos que $\bar{x}$ es la función lineal (linea recta) que pasa por los puntos (a, x_1) y (b, x_2).

3.4. Subgradientes de funcionales en espacios normados

Definición 3.16. Sean $(X, \| \|)$ un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional convexa. Decimos que $l \in X^\star$ es un *subgradiente* de f en $\bar{x} \in X$ si

$$f(x) \geq f(\bar{x}) + l(x - \bar{x})$$

para todo $x \in X$. Al conjunto de todos los subgradientes en un punto $\bar{x}$ será llamado el *subdiferencial* de f en $\bar{x}$, denotado por

$$\partial f(\bar{x}) = \{l \in X^\star / f(x) \geq f(\bar{x}) + l(x - \bar{x}) \text{ para todo } x \in X\}$$

Ejemplos.

- Sean X un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional convexa Gateaux diferenciable. Si $x \in X$, por el Teorema 3.11 tenemos

$$f(y) \geq f(x) + f'(x)(y - x)$$

Como $f'(x) \in X^\star$ se sigue que $f'(x) \in \partial f(x)$.

- Sean $(X, \| \|)$ un e.v.n y $X^\star$ su espacio dual topológico con la norma usual. Consideremos la funcional norma $f(x) = \|x\|$. Si $x \in X$, se cumple

$$\partial f(x) = \begin{cases} \{l \in X^\star / l(x) = \|x\| \wedge \|l\|_{X^\star} = 1\}, & si\ x \neq 0 \\ \{l \in X^\star / \|l\|_{X^\star} \leq 1\}, & si\ x = 0 \end{cases}$$

En efecto. En principio veamos el caso en que $x = 0$. Sea $l \in \partial f(0)$ entonces $f(x) \geq f(0) + l(x - 0)$, es decir $f(x) \geq l(x)$ para todo $x \in X$. Así $\|x\| \geq l(x)$ para todo $x \in X$. De esta última desigualdad, por la definición de la norma en el espacio dual $X^\star$ conseguimos

$$\begin{aligned} \frac{l(x)}{\|x\|} \leq 1 &\Longrightarrow \sup_{x \neq 0} \frac{l(x)}{\|x\|} \leq 1 \\ &\Longrightarrow \|l\|_{X^\star} \leq 1 \end{aligned}$$

Resultando la primera inclusión. Para la otra inclusión, si $\|l\|_{X^\star} \leq 1$ tenemos que $|l(x)| \leq \|x\|$ para todo $x \neq 0$, de donde $l(x) \leq \|x\|$, siendo esta desigualdad evidente para $x = 0$. Usando esta desigualdad resulta que $f(x) \geq f(0) + l(x - 0)$ por lo tanto $l \in \partial f(0)$.

Ahora consideremos el caso en que $x \neq 0$. Por el Teorema 1.49 existe $l \in X^\star$ tal que $l(x) = \|x\|$ y $\|l\|_{X^\star} = 1$. Así conseguimos $l(x) \leq \|x\|$ para todo $x \in X$. Usando esta desigualdad y las condiciones sobre l, obtenemos

$$\begin{aligned}\|x\| + l(y-x) &= \|x\| - l(x) + l(y)\\ &= l(y) \leq \|y\|\end{aligned}$$

de donde $\|y\| \geq \|x\| + l(y-x)$, por lo tanto $l \in \partial f(x)$. Para la otra inclusión, si $l \in \partial f(x)$ entonces

$$\begin{aligned}\|x\| - l(x) &= 2\,\|x\| - \|x\| - l(x)\\ &= \|2x\| - \|x\| - l(2x-x) \geq 0\end{aligned} \tag{3.19}$$

Análogamente

$$\begin{aligned}-\|x\| + l(x) &= \|0\| - \|x\| - l(0-x)\\ &= \geq 0\end{aligned} \tag{3.20}$$

De las desigualdades (3.19) y (3.20) logramos $l(x) = \|x\|$. Por otra parte, usando esta última identidad, para $y \in X$ obtenemos

$$\begin{aligned}\|y\| &\geq \|x\| + l(y-x)\\ &= \|x\| + l(y) - l(x)\\ &= l(y)\end{aligned} \tag{3.21}$$

De la desigualdad (3.21) resulta que $\|l\|_{X^\star} = \sup\limits_{x\neq 0} \dfrac{|l(x)|}{\|x\|} \leq 1$ y siendo $l(x) = \|x\|$ concluimos que $\|l\|_{X^\star} = 1$.

Lema 3.17. Consideremos $(X, \|\ \|)$ un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional convexa. Entonces para cada $\bar{x} \in X$ se cumple

$$\partial f(\bar{x}) = \left\{ l \in X^\star / f'(\bar{x})h \geq l(h) \text{ para todo } h \in X \right\}$$

donde $f'(\bar{x})h$ es la derivada direccional de f en $\bar{x}$ en la dirección h.

Demostración.

Por el Lema 3.4 existe la derivada direccional $f'(\bar{x})h$ para todo $\bar{x}, h \in X$. Sea $l \in \partial f(\bar{x})$, para $\lambda > 0$ tenemos $f(\bar{x} + \lambda h) - f(\bar{x}) \geq l(\lambda h)$. Como l es una aplicación lineal, tenemos la desigualdad $\lambda^{-1}\{f(\bar{x}+\lambda h) - f(\bar{x})\} \geq l(h)$. Tomando límite cuando $\lambda \to 0^+$ en la desigualdad anterior, logramos

$$\lim_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda h) - f(\bar{x})\} \geq l(h)$$

Por la Definición 3.1 tenemos que $f'(\bar{x})h \geq l(h)$ para todo $h \in X$.

Veamos la otra inclusión. Supongamos que $l \in X^\star$ y $f'(\bar{x})h \geq l(h)$ para todo $h \in X$. Por el Lema 3.2 se cumple que $\varphi(1) \geq \lim_{\lambda \to 0^+} \varphi(\lambda)$ donde $\varphi(\lambda) = \frac{1}{\lambda} \{f(\bar{x} + \lambda h) - f(\bar{x})\}$. Por lo tanto $f(\bar{x} + h) - f(\bar{x}) \geq f'(\bar{x})h \geq l(h)$ para todo $h \in X$. De esta manera, conseguimos

$$f(\bar{x} + h) \geq f(\bar{x}) + l(h) \tag{3.22}$$

para todo $h \in X$. Tomando $u = \bar{x} + h$ en (3.22) tenemos

$$f(u) \geq f(\bar{x}) + l(u - \bar{x})$$

para todo $u \in X$ mostrando así que $l \in \partial f(\bar{x})$.

□

El siguiente resultado, muestra que bajo ciertas condiciones el subdiferencial de una funcional, es un conjunto no vacío.

Teorema 3.18. Sean $(X, \| \|)$ un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional convexa y continua. Entonces $\partial f(\bar{x}) \neq \emptyset$ para todo $\bar{x} \in X$.

Demostración.

Sea $\bar{x} \in X$, por hipótesis f es continua en $\bar{x}$, por tanto existen una bola abierta $B(\bar{x}, \delta)$ y $\bar{\alpha} > 0$ tales que $|f(x)| \leq \bar{\alpha}$ para todo $x \in B(\bar{x}, \delta)$. Consideremos el epígrafe de f, $\text{epi}(f) = \{(x, \alpha) \in X \times \mathbb{R} / f(x) \leq \alpha\}$. Por la convexidad de f, epi(f) es un conjunto convexo y además se cumplen $(\bar{x}, \bar{\alpha}+1) \in int(\text{epi}(f))$ y $(\bar{x}, f(\bar{x})) \notin int(\text{epi}(f))$. Luego por el Teorema de Separación de Eidelheit (Teorema 1.47) existen una funcional lineal continua $T : X \times \mathbb{R} \longrightarrow \mathbb{R}$ con $T \neq 0_{(X \times \mathbb{R})^\star}$ y $\gamma \in \mathbb{R}$ tales que

$$T(x, \alpha) \leq \gamma \leq T(\bar{x}, \bar{\alpha}) \tag{3.23}$$

para todo $(x, \alpha) \in \text{epi}(f)$. Observemos que para cada $(x, \alpha) \in X \times \mathbb{R}$ podemos escribir

$$T(x, \alpha) = T(x, 0) + \alpha T(0, 1)$$

Denotando por $T(x, 0) = l(x)$ y $\beta = T(0, 1)$ tenemos que $l \in X^\star$, $\beta \in \mathbb{R}$ y el par $(l, \beta) \neq (0_{X^\star}, 0)$, con lo cual $T(x, \alpha) = l(x) + \beta\alpha$, quedando la desigualdad (3.23) expresada como

$$l(x) + \beta\alpha \leq \gamma \leq l(\bar{x}) + \beta f(\bar{x}) \tag{3.24}$$

para todo $(x, \alpha) \in \text{epi}(f)$. Tomando $x = \bar{x}$ en (3.24) tenemos que $\beta\alpha \leq \beta f(\bar{x})$ para todo

$\alpha \geq f(\bar{x})$, en consecuencia $\beta \leq 0$. Supongamos que $\beta = 0$, de (3.24) tenemos

$$l(x) \leq l(\bar{x})$$

para todo $x \in X$, de donde resulta

$$\|l\|_{X^\star} = \sup_{x \neq \bar{x}} \frac{|l(x - \bar{x})|}{\|x - \bar{x}\|} \leq 0$$

Obteniendo que $\|l\|_{X^\star} = 0$, en contradicción con el hecho que $(l, \beta) \neq (0_{X^\star}, 0)$. Por lo tanto $\beta < 0$. Así, de (3.24) dividiendo por β obtenemos

$$\frac{1}{\beta} l(x) + \alpha \geq \frac{1}{\beta} l(\bar{x}) + f(\bar{x}) \tag{3.25}$$

para todo $(x, \alpha) \in \text{epi}(f)$. Tomando $\alpha = f(x)$ en (3.25) conseguimos

$$\begin{aligned} \frac{1}{\beta} l(x) + f(x) \geq \frac{1}{\beta} l(\bar{x}) + f(\bar{x}) \\ \Longrightarrow f(x) \geq f(\bar{x}) - \frac{1}{\beta} l(x - \bar{x}) \end{aligned}$$

para todo $x \in X$. Por la Definición 3.16 y la desigualdad anterior concluimos que $-\frac{1}{\beta} l \in \partial f(\bar{x})$. En consecuencia $\partial f(\bar{x}) \neq \emptyset$.

$\square$

Teorema 3.19. Sean $(X, \| \, \|)$ un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional convexa. Entonces $\bar{x} \in X$ es punto mínimo de f si y sólo si $0_{X^\star} \in \partial f(\bar{x})$.

Demostración.

$$\begin{aligned} \bar{x} = \underset{x \in X}{argmin}\, f(x) &\Longleftrightarrow f(x) \geq f(\bar{x}) \text{ para todo } x \in X \\ &\Longleftrightarrow f(x) \geq f(\bar{x}) + 0_{X^\star}(x - \bar{x}) \text{ para todo } x \in X \\ &\Longleftrightarrow 0_{X^\star} \in \partial f(\bar{x}) \end{aligned}$$

$\square$

Teorema 3.20. Sean $(X, \| \, \|)$ un e.v.n y $f : X \longrightarrow \mathbb{R}$ una funcional convexa y continua. Entonces para cualesquiera $\bar{x}$ y h en X, la derivada direccional de f en $\bar{x}$ en la dirección h es

$$f'(\bar{x})h = \max_{l \in \partial f(\bar{x})} l(h)$$

Demostración.

Tomemos $\bar{x} \in X$ un punto arbitrario y $h \in X$ una dirección arbitraria. Evidentemente, el teorema es válido si $h = 0$. Supongamos que $h \neq 0$, por hipótesis, los Lemas 3.3 y 3.17 y el Teorema 3.18 tenemos que $f'(\bar{x})h$ existe, $\partial f(\bar{x})$ es no vacío y $f'(\bar{x})h \geq l(h)$ para todo $l \in \partial f(\bar{x})$. Veamos la existencia de un elemento $w \in \partial f(\bar{x})$ tal que $f'(\bar{x})h \geq w(h)$. En efecto, consideremos el conjunto

$$T = \big\{(\bar{x} + \lambda h, f(\bar{x}) + \lambda f'(\bar{x})h) \in X \times \mathbb{R} / \lambda \geq 0\big\}$$

- $T \neq \emptyset$ pues $(\bar{x}, f(\bar{x})) \in T$.

- $\operatorname{epi}(f) \cap T = \emptyset$.

  En efecto, por el Lema 3.2 la función $\varphi(\lambda) = \dfrac{1}{\lambda}\{f(\bar{x} + \lambda h) - f(\bar{x})\}$ es monótona creciente y $\lim\limits_{\lambda \to 0^+} \varphi(\lambda) = f'(\bar{x})h$. Así $\varphi(\lambda) \geq f'(\bar{x})h$ para todo $\lambda \geq 0$. De esta última desigualdad obtenemos $f(\bar{x}+\lambda h) \geq f(\bar{x})+\lambda f'(\bar{x})h$ para todo $\lambda \geq 0$. Así el par $(\bar{x}+\lambda h, f(\bar{x})+\lambda f'(\bar{x})h)$ no pertenece a $int(\operatorname{epi}(f))$ para ningún $\lambda \geq 0$, mostrando la afirmación.

- T es convexo.

  Efectivamente, si $u_1, u_2 \in T$ y $0 \leq \lambda \leq 1$ tenemos

$$\begin{aligned}
\lambda u_1 + (1-\lambda)u_2 &= \lambda\big\{(\bar{x} + \lambda_1 h, f(\bar{x}) + \lambda_1 f'(\bar{x})h)\big\} \\
&\quad + (1-\lambda)\big\{(\bar{x} + \lambda_2 h, f(\bar{x}) + \lambda_2 f'(\bar{x})h)\big\} \\
&= (\bar{x} + \lambda\lambda_1 + (1-\lambda)\lambda_2, f(\bar{x}) + \{\lambda\lambda_1 + (1-\lambda)\lambda_2\} f'(\bar{x})h)
\end{aligned}$$

  donde λ_1 y λ_2 pertenecen a $[0, +\infty)$. Como $\lambda\lambda_1 + (1-\lambda)\lambda_2 \geq 0$ se sigue la afirmación.

De estas afirmaciones, tenemos por el Teorema 1.47 y cálculos similares efectuados en (3.23) y (3.24), que existen $l \in X^\star$, $\beta, \gamma \in \mathbb{R}$ con $(l, \beta) \neq (0_{X^\star}, 0)$ tales que

$$l(x) + \beta\alpha \leq \gamma \leq l(\bar{x} + \lambda h) + \beta(f(\bar{x}) + \lambda f'(\bar{x})h) \tag{3.26}$$

para todo $(x, \alpha) \in \operatorname{epi}(f)$ y $\lambda \geq 0$. Usando los mismos cálculos efectuados en la demostración del Teorema 3.19, tomando $x = \bar{x}$ y $\lambda = 0$ obtenemos que $\beta < 0$. De esta forma, multiplicando por $\dfrac{1}{\beta}$ a los dos extremos de la desigualdad (3.26) y transponiendo términos convenientemente conseguimos

$$\frac{1}{\beta} l(x - \bar{x} - \lambda h) + \alpha \geq f(\bar{x}) + \lambda f'(\bar{x})h \tag{3.27}$$

Tomando $\alpha = f(x)$ y $\lambda = 0$ en (3.27) y despejando términos obtenemos

$$f(x) \geq f(\bar{x}) - \frac{1}{\beta} l(x - \bar{x})$$

mostrando que $-\frac{1}{\beta} l \in \partial f(\bar{x})$.

Haciendo $x = \bar{x}$, $\alpha = f(x)$ y $\lambda = 1$ en (3.27) tenemos

$$\frac{1}{\beta} l(\bar{x} - \bar{x} - h) + f(\bar{x}) \geq f(\bar{x}) + f'(\bar{x})h$$

En consecuencia por la linealidad de $-\frac{1}{\beta} l$ logramos $-\frac{1}{\beta} l(h) \geq f'(\bar{x})h$. Denotando por $w = -\frac{1}{\beta} l$ tenemos que $f'(\bar{x})h \leq w(h)$. Con lo cual queda demostrado el teorema.

□

Corolario 3.21. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ no vacío y $f : X \longrightarrow \mathbb{R}$ una funcional convexa y continua.

(*i*) Si $\bar{x}$ es el punto mínimo de f en S y S es estrellado respecto a $\bar{x}$, entonces

$$\max_{l \in \partial f(\bar{x})} l(x - \bar{x}) \geq 0 \tag{3.28}$$

para todo $x \in S$.

(*ii*) Si para algún $\bar{x} \in S$ la desigualdad (3.28) es válida, entonces $\bar{x}$ es punto mínimo de f en S.

Demostración.

(*i*) Como $\bar{x}$ es punto mínimo de f, por el Teorema 3.6-(*i*) tenemos $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$. Luego por el Teorema 3.20 tenemos

$$f'(\bar{x})(x - \bar{x}) = \max_{l \in \partial f(\bar{x})} l(x - \bar{x}) \geq 0$$

para todo $x \in S$.

(*ii*) Por hipótesis, el Lema 3.2 y el Teorema 3.20 conseguimos

$$\begin{aligned} \frac{1}{\lambda} \{f(\bar{x} + \lambda(x - \bar{x}) - f(\bar{x})\} &\geq f'(\bar{x})(x - \bar{x}) \\ &= \max_{l \in \partial f(\bar{x})} l(x - \bar{x}) \geq 0 \end{aligned} \tag{3.29}$$

para todo $x \in S$ y $\lambda > 0$. Haciendo $\lambda = 1$ en la desigualdad (3.29) obtenemos $f(x) \geq f(\bar{x})$ para cada $x \in S$, por lo tanto $\bar{x}$ es punto mínimo de f en S.

□

3.5. Condiciones de optimalidad para funcionales Clarke diferenciables

En esta sección estudiaremos a las funcionales Clarke diferenciables, siendo considerada una versión débil de derivada direccional en espacios abstractos. Asimismo desarrollaremos resultados que caracterizan a los puntos minimizantes para esta clase de funcionales.

Definición 3.22. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Decimos que f es *Clarke diferenciable* en $\bar{x} \in S$ si el límite superior

$$f_c'(\bar{x})h := \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda h) - f(x)\}$$

existe para todo $h \in X$. El valor $f_c'(\bar{x})h$ será llamado la derivada de Clarke de f en la dirección h.

Observaciones.

- La derivada de Clarke se diferencia de la derivada direccional en el sentido que el valor de x varia en el cociente incremental, es decir $x \to \bar{x}$.

- La definición de la derivada de Clarke es una versión débil del límite definido en la derivada de Gateaux. En consecuencia si f es una funcional Gateaux diferenciable, entonces es Clarke diferenciable.

- Para que la expresión del límite superior en la derivada de Clarke, este bien definida debemos exigir que $x \in S$ y $x + \lambda h \in S$ para valores de λ suficientemente pequeños. Por este motivo $x \in \overset{\circ}{S}$.

Ejemplo.

Sea la función real de variable real $f(x) = |x|$. Sabemos que esta función no es derivable en el punto $x = 0$. Sin embargo, esta aplicación es Clarke diferenciable en $x = 0$ y además dado $h \in \mathbb{R}$ se cumple $f_c'(0)h = |h|$. En efecto, por la desigualdad triangular tenemos

$$\begin{aligned} f_c'(0)h &= \limsup_{\substack{x \to 0 \\ \lambda \to 0^+}} \frac{1}{\lambda} \{|x + \lambda h| - |x|\} \\ &\leq \limsup_{\substack{x \to 0 \\ \lambda \to 0^+}} \frac{1}{\lambda} \{|x| + \lambda |h| - |x|\} \\ &\leq |h| \end{aligned}$$

de donde $f_c'(0)h \leq |h|$. Por otra parte si $x = \lambda h$, siendo el límite superior el mayor valor adherente tenemos

$$
\begin{aligned}
f_c'(0)h &= \limsup_{\substack{x\to\bar{x}\\ \lambda\to 0^+}} \frac{1}{\lambda}\{|x+\lambda h|-|x|\}\\
&\geq \limsup_{\lambda\to 0^+} \frac{1}{\lambda}\{2\lambda\,|h|-\lambda\,|h|\}\\
&\geq \limsup_{\lambda\to 0^+}\{|h|\} = |h|
\end{aligned}
$$

Con lo cual $f_c'(0)h \geq |h|$, que demuestra nuestra afirmación.

Lema 3.23. Sean $(X, \|\ \|)$ un e.v.n, $S \subseteq X$ con interior no vacío y $\bar{x} \in \overset{\circ}{S}$. Si f es una funcional Lipschitz continua en $\bar{x}$ con constante de Lipschitz k, entonces f es Clarke diferenciable en $\bar{x}$ y $|f_c'(\bar{x})h| \leq k\,\|h\|$ para todo $h \in X$.

Demostración.

Como f es Lipschitz continua en $\bar{x}$ existe $\delta_1 > 0$ tal que $|f(x)-f(y)| \leq k\,\|x-y\|$ para todo $x, y \in B(\bar{x}, \delta_1) \cap S$. Además por hipótesis $\bar{x} \in \overset{\circ}{S}$, si $x \to \bar{x}$ y $\lambda \to 0^+$ podemos encontrar una bola abierta $B(\bar{x}, \delta_2)$ tal que $x, x+\lambda h \in B(\bar{x}, \delta_2) \subseteq S$. Tomando $\delta = \text{mín}\{\delta_1, \delta_2\}$, tenemos

$$
\begin{aligned}
\left|\frac{1}{\lambda}\{f(x+\lambda h)-f(x)\}\right| &\leq \frac{k}{\lambda}\,\|x+\lambda h - x\|\\
&\leq \frac{k}{\lambda}\,\|\lambda h\| = k\,\|h\| \qquad (3.30)
\end{aligned}
$$

Así la expresión $\left|\frac{1}{\lambda}\{f(x+\lambda h)-f(x)\}\right|$ esta acotada superiormente, en consecuencia existe

$$
f_c'(\bar{x})h = \limsup_{\substack{x\to\bar{x}\\ \lambda\to 0^+}} \frac{1}{\lambda}\{f(x+\lambda h)-f(x)\}
$$

Luego, gracias a (3.30) tenemos

$$
\begin{aligned}
\left|f_c'(\bar{x})h\right| &= \left|\limsup_{\substack{x\to\bar{x}\\ \lambda\to 0^+}} \frac{1}{\lambda}\{f(x+\lambda h)-f(x)\}\right|\\
&\leq \limsup_{\substack{x\to\bar{x}\\ \lambda\to 0^+}} \left|\frac{1}{\lambda}\{f(x+\lambda h)-f(x)\}\right| \leq k\,\|h\|
\end{aligned}
$$

Así, $|f_c'(\bar{x})h| \leq k\,\|h\|$ para todo $h \in X$.

$\square$

Lema 3.24. Sean $(X, \|\ \|)$ un e.v.n, $S \subseteq X$ con interior no vacío y $\bar{x} \in \overset{\circ}{S}$. Si $f : S \longrightarrow \mathbb{R}$ es una funcional Clarke diferenciable en $\bar{x}$ entonces la aplicación $f_c'(\bar{x})$ es una funcional sublineal.

Demostración.

Claramente $f'_c(\bar{x})(0_X) = 0$. Consideremos $h \in X$ no nulo y $\alpha > 0$ entonces por las propiedades del límite superior tenemos

$$\begin{aligned} f'_c(\bar{x})(\alpha h) &= \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda \alpha h) - f(x)\} \\ &= \alpha \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\alpha \lambda} \{f(x + \lambda \alpha h) - f(x)\} \\ &= \alpha \limsup_{\substack{x \to \bar{x} \\ \mu \to 0^+}} \frac{1}{\mu} \{f(x + \mu h) - f(x)\} = \alpha f'_c(\bar{x})(h) \end{aligned} \tag{3.31}$$

Tomando $h_1, h_2 \in X$, resulta que

$$\begin{aligned} f'_c(\bar{x})(h_1 + h_2) &= \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda(h_1 + h_2)) - f(x)\} \\ &= \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda(h_1 + h_2)) - f(x + \lambda h_2) + f(x + \lambda h_2) - f(x)\} \\ &\leq \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda(h_1 + h_2)) - f(x + \lambda h_2)\} \\ &\quad + \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda h_2) - f(x)\} \end{aligned} \tag{3.32}$$

Como $x + \lambda h_2 \to \bar{x}$ si $x \to \bar{x}$ y $\lambda \to 0^+$ de (3.32) logramos

$$\begin{aligned} f'_c(\bar{x})(h_1 + h_2) &\leq \limsup_{\substack{u \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(u + \lambda h_1) - f(u)\} + \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(x + \lambda h_2) - f(x)\} \\ &= f'_c(\bar{x})(h_1) + f'_c(\bar{x})(h_2) \end{aligned} \tag{3.33}$$

Por lo tanto, de la identidad (3.31) y la desigualdad (3.33) se sigue el resultado.

$\square$

Teorema 3.25. Sean $(X, \| \ \|)$ un e.v.n y $f : X \longrightarrow \mathbb{R}$ es una funcional convexa y Lipschitz continua en $\bar{x} \in X$. Entonces las derivadas de Clarke y direccional de f en el punto $\bar{x}$ coinciden.

Demostración.

Consideremos $h \in X$ una dirección arbitraria. Por el Lema 3.23 existe la derivada de Clarke $f'_c(\bar{x})h$; por la convexidad de f y el Lema 3.4 existe la derivada direccional $f'(\bar{x})$ en la dirección h. Así tenemos

$$
\begin{aligned}
f'(\bar{x})h &= \lim_{\lambda \to 0^+} \frac{1}{\lambda}\{f(\bar{x}+\lambda h) - f(\bar{x})\} \\
&\leq \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda}\{f(x+\lambda h) - f(x)\} = f'_c(\bar{x})h
\end{aligned}
$$

Resultando la desigualdad $f'(\bar{x})h \leq f'_c(\bar{x})h$ para todo $h \in X$.

Por otra parte, por el Lema 3.2 ($\varphi(\lambda) \leq \varphi(\epsilon)$) y la definición de la derivada de Clarke, se tiene que

$$
\begin{aligned}
f'_c(\bar{x})h &= \limsup_{\substack{x \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda}\{f(x+\lambda h) - f(x)\} \\
&= \lim_{\substack{\delta \to 0^+ \\ \epsilon \to 0^+}} \sup_{\|x-\bar{x}\|<\delta} \sup_{0<\lambda<\epsilon} \frac{1}{\lambda}\{f(x+\lambda h) - f(x)\} \\
&= \lim_{\substack{\delta \to 0^+ \\ \epsilon \to 0^+}} \sup_{\|x-\bar{x}\|<\delta} \frac{1}{\epsilon}\{f(x+\epsilon h) - f(x)\} \qquad (3.34)
\end{aligned}
$$

Sea $\alpha > 0$ arbitrario. De la igualdad (3.34) tenemos ($\delta = \epsilon\alpha$)

$$
f'_c(\bar{x})h = \lim_{\epsilon \to 0^+} \sup_{\|x-\bar{x}\|<\epsilon\alpha} \frac{1}{\epsilon}\{f(x+\epsilon h) - f(x)\} \qquad (3.35)
$$

Por hipótesis, f es Lipchitiz continua en $\bar{x}$, luego para $\epsilon > 0$ suficientemente pequeño obtenemos la desigualdad

$$
\begin{aligned}
\left|\frac{1}{\epsilon}\{f(x+\epsilon h) - f(x)\} - \frac{1}{\epsilon}\{f(\bar{x}+\epsilon h) - f(\bar{x})\}\right| &\leq \frac{1}{\epsilon}|f(x+\epsilon h) - f(\bar{x}+\epsilon h)| \\
&\quad + \frac{1}{\epsilon}|f(x) - f(\bar{x})| \\
&\leq \frac{k}{\epsilon}\|x-\bar{x}\| + \frac{k}{\epsilon}\|x-\bar{x}\| \qquad (3.36) \\
&= \frac{k}{\epsilon}(2\epsilon\alpha) = 2k\alpha
\end{aligned}
$$

donde k es la constante de Lipschitz de f.

Por lo tanto

$$
\frac{1}{\epsilon}\{f(x+\epsilon h) - f(x)\} \leq \frac{1}{\epsilon}\{f(\bar{x}+\epsilon h) - f(\bar{x})\} + 2k\alpha \qquad (3.37)
$$

De la desigualdad (3.37), la identidad (3.35) y la Definición 3.1 tenemos

$$\begin{aligned} f_c'(\bar{x})h &= \lim_{\epsilon\to 0^+} \sup_{\|x-\bar{x}\|<\epsilon\alpha} \frac{1}{\epsilon}\{f(x+\epsilon h)-f(x)\} \\ &\leq \lim_{\epsilon\to 0^+} \frac{1}{\epsilon}\{f(\bar{x}+\epsilon h)-f(\bar{x})\}+2k\alpha \\ &= f'(\bar{x})h+2k\alpha \end{aligned}$$

Por lo tanto $f_c'(\bar{x})h \leq f'(\bar{x})h + 2k\alpha$ para todo $\alpha \geq 0$. En consecuencia $f_c'(\bar{x})h \leq f'(\bar{x})h$. De esta forma, obtenemos la igualdad $f_c'(\bar{x})h{=}f'(\bar{x})h$, mostrando el resultado.

$\square$

Definición 3.26. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ un subconjunto con interior no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional Lipschitz continua en $\bar{x} \in \overset{\circ}{S}$. Definimos el gradiente generalizado o *subdiferencial de Clarke* de f en $\bar{x}$ como el conjunto

$$\partial_{cl} f(\bar{x}) = \left\{ l \in X^{\star} / f_c'(\bar{x}) \geq l(h) \text{ para todo } h \in X \right\}$$

Teorema 3.27. Consideremos $(X, \| \|)$ un e.v.n, $f : X \longrightarrow \mathbb{R}$ una funcional convexa y Lipschitz continua en $\bar{x} \in X$. Entonces el subdiferencial $\partial f(\bar{x})$ de f en $\bar{x}$ coincide con el gradiente generalizado $\partial_{cl} f(\bar{x})$ de f en $\bar{x}$.

Demostración.

Por los Teoremas 3.20, 3.25 y la Definición 3.26 tenemos

$$\begin{aligned} l \in \partial f(\bar{x}) &\Longleftrightarrow f'(\bar{x})h \geq l(h) \text{ para todo } h \in X \\ &\Longleftrightarrow f_c'(\bar{x})h \geq l(h) \text{ para todo } h \in X \\ &\Longleftrightarrow l \in \partial_{cl} f(\bar{x}) \end{aligned}$$

De donde $\partial_{cl} f(\bar{x}) = \partial f(\bar{x})$.

$\square$

Teorema 3.28. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ un subconjunto con interior no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional Lipschitz continua en $\bar{x} \in \overset{\circ}{S}$. Entonces $\partial_{cl} f(\bar{x})$ es un subconjunto no vacío.

Demostración.

Por los Lemas 3.23 y 3.24, f es Clarke diferenciable y la derivada de Clarke es una funcional sublineal. Luego, por el Corolario 1.44 existe una funcional lineal $l : X \longrightarrow \mathbb{R}$ tal que $f_c'(\bar{x})h \geq l(h)$ para todo $h \in X$. Veamos la continuidad de l en X. Por la Definición 3.26 y el Lema 3.23 tenemos

$$
\begin{aligned}
l(h) &\leq f'_c(\bar{x})h \\
&\leq \left| f'_c(\bar{x})h \right| \\
&\leq k \left\| h \right\|
\end{aligned} \tag{3.38}
$$

Donde k es la constante de Lipschitz de f. Por la linealidad de l tenemos

$$
\begin{aligned}
l(-h) &= -l(h) \\
&\leq f'_c(\bar{x})(-h) \\
&\leq k \left\| -h \right\| = k \left\| h \right\|
\end{aligned} \tag{3.39}
$$

Por lo tanto de (3.38) y (3.39) conseguimos la desigualdad $-k \left\| h \right\| \leq l(h) \leq k \left\| h \right\|$, en consecuencia $|l(h)| \leq k \left\| h \right\|$ para todo $h \in X$. Por el Teorema 1.12, l es continua en X con lo cual $l \in X^{\star}$. Por la Definición 3.26 tenemos que $l \in \partial_{cl} f(\bar{x})$.

□

El siguiente teorema establece una condición necesaria para los puntos minimizantes de una funcional Clarke diferenciable.

Teorema 3.29. Sean $(X, \| \|)$ un e.v.n, $T \subseteq X$ un subconjunto con interior no vacío, $f : T \longrightarrow \mathbb{R}$ una funcional y $S \subseteq T$. Supongamos que $\bar{x} \in S \cap \overset{\circ}{T}$ satisface:

(i) $\bar{x}$ es punto mínimo de f en T.

(ii) S es estrellado respecto a $\bar{x}$.

(iii) f es Lipschitz continua en $\bar{x}$.

Entonces la derivada de Clarke $f'_c(\bar{x})$ satisface $f'_c(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$.

Demostración.

Como S es estrellado respecto a $\bar{x}$ entonces $\bar{x} + \lambda(x - \bar{x}) \in S$ para $0 < \lambda < 1$. Por hipótesis f es Lipschitz continua con constante k en $\bar{x} \in \overset{\circ}{T}$, por el Lema 3.23 f es Clarke diferenciable en $\bar{x}$. Si $x \in S$ tenemos

$$
\begin{aligned}
\left| \frac{1}{\lambda} \left\{ f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x}) \right\} \right| &\leq \frac{k}{\lambda} \left\| \lambda(x - \bar{x}) \right\| \\
&= k \left\| x - \bar{x} \right\|
\end{aligned} \tag{3.40}
$$

Tomando límite cuando $\lambda \to 0^+$ por la monotonia de la aplicación $\varphi(\cdot)$, en (3.40) logramos la existencia de $f'(\bar{x})(x - \bar{x})$. Como $\bar{x} = \underset{x \in S}{argmin} f(x)$ pues $S \subseteq T$, tenemos la desigualdad $f(\bar{x} + \lambda(x - \bar{x})) \geq f(\bar{x})$ por lo tanto

$$\limsup_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})\} \geq 0$$

Luego,

$$\begin{aligned}\limsup_{\substack{y \to \bar{x} \\ \lambda \to 0^+}} \frac{1}{\lambda} \{f(y + \lambda(x - \bar{x})) - f(y)\} &\geq \limsup_{\lambda \to 0^+} \frac{1}{\lambda} \{f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})\} \\ &\geq 0\end{aligned}$$

Por la Definición 3.22 logramos $f'_c(\bar{x})(x - \bar{x}) \geq 0$.

$\square$

Observacion.

En el teorema anterior si consideramos $S = X$ y $f : X \longrightarrow \mathbb{R}$ una funcional Clarke diferenciable entonces $\bar{x} = \underset{x \in X}{argmin} f(x)$ y $0_{X^\star} \in \partial_{cl} f(\bar{x})$.

Capítulo 4

Optimización no convexa en espacios de Banach

4.1. Conos tangentes

Definición 4.1. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ no vacío y $\bar{x} \in \bar{S}$.

(i) Diremos que $h \in X$ es un *vector tangente* a S en $\bar{x}$ si existen dos sucesiones $(x_n)_{n\in\mathbb{N}} \subseteq S$ y $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^+$ tales que $\bar{x} = \lim\limits_{n\to\infty} x_n$ y $h = \lim\limits_{n\to\infty} \lambda_n(x_n - \bar{x})$.

(ii) El conjunto $T(S, \bar{x}) = \{h \in X/\, h$ es vector tangente a S en $\bar{x}\}$ será llamado *cono tangente secuencial de Bouligand* a S en $\bar{x}$ o *cono contenedor de* S en $\bar{x}$.

Observaciones.

- Por definición de $T(S, \bar{x})$ se exige que $\bar{x} \in \bar{S}$, sin embargo en lo sucesivo se considerará que $\bar{x} \in S$.

- Claramente $T(S, \bar{x})$ es un cono. En efecto, sean $h \in T(S, \bar{x})$ y $\lambda \geq 0$. Si $\lambda = 0$ evidentemente $\lambda h \in T(S, \bar{x})$ (basta considerar la sucesión constante $x_n = \bar{x}$). Por otro lado si $\lambda > 0$, tomemos las sucesiones (x_n) y (λ_n). Definamos $\mu_n = \lambda\lambda_n$, entonces tenemos que $\bar{x} = \lim\limits_{n\to\infty} x_n$ y

$$\begin{aligned}\lambda h &= \lambda \lim_{n\to\infty} \lambda_n(x_n - \bar{x}) \\ &= \lim_{n\to\infty} \underbrace{\lambda\lambda_n}_{\mu_n}(x_n - \bar{x})\end{aligned}$$

  con lo cual $\lambda h \in T(S, \bar{x})$, mostrando así nuestra afirmación.

Lema 4.2. Sean $(X, \| \|)$ un e.v.n y $S \subseteq X$ no vacío. Si S es estrellado respecto a $\bar{x} \in S$, entonces $\text{cono}(S - \{\bar{x}\}) \subseteq T(S, \bar{x})$.

Demostración.

En principio, si S es un cono, entonces $\text{cono}(S) = S$. Tomemos $x \in S$. Como S es estrellado respecto a $\bar{x} \in S$ tenemos que el punto

$$\begin{aligned} x_n &= \bar{x} + \frac{1}{n}(x - \bar{x}) \\ &= \frac{1}{n}x + \left(1 - \frac{1}{n}\right)\bar{x} \end{aligned} \tag{4.1}$$

pertenece a S para cada $n \in \mathbb{N}$. Tomando el límite en (4.1) cuando $n \to \infty$ tenemos que $x_n \longrightarrow \bar{x}$ y además haciendo $\lambda_n = n$, nuevamente de (4.1) obtenemos

$$\begin{aligned} \lim_{n\to\infty} \lambda_n(x_n - \bar{x}) &= \lim_{n\to\infty} n(x_n - \bar{x}) \\ &= x - \bar{x} \end{aligned}$$

Por lo tanto $x - \bar{x} \in T(S, \bar{x})$, mostrando así que $S - \{\bar{x}\} \subseteq T(S, \bar{x})$. Por la afirmación dada al inicio de la prueba y la observación posterior a la Definición 4.1 se sigue que

$$\text{cono}(S - \{\bar{x}\}) \subseteq \text{cono}(T(S, \bar{x})) = T(S, \bar{x})$$

Con lo cual finaliza la demostración.

□

Lema 4.3. Sean $(X, \| \, \|)$ un e.v.n y $S \subseteq X$ no vacío. Si $\bar{x} \in S$ entonces $T(S, \bar{x}) \subseteq \overline{\text{cono}(S - \{\bar{x}\})}$.

Demostración.

Sean $\bar{x} \in S$ (fijo) y $h \in T(S, \bar{x})$. Por la Definición 4.1, existen $(x_n)_{n\in\mathbb{N}} \subseteq S$ y $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^+$ tales que $\bar{x} = \lim\limits_{n\to\infty} x_n$ y $h = \lim\limits_{n\to\infty} \lambda_n(x_n - \bar{x})$. Por la Definición 1.30 tenemos

$$\text{cono}(S - \{\bar{x}\}) = \{\lambda(x - \bar{x}) / \lambda \geq 0 \, \wedge \, x \in S\}$$

Observemos que la sucesión $z_n = \lambda_n(x_n - \bar{x})$ pertenece a $\text{cono}(S - \{\bar{x}\})$ y como $h = \lim\limits_{n\to\infty} z_n$ logramos que $h \in \overline{\text{cono}(S - \{\bar{x}\})}$, quedando demostrado el lema.

□

Observación.

Si S es un conjunto estrellado respecto a $\bar{x}$, de los Lemas 4.2 y 4.3 se sigue que

$$\text{cono}(S - \{\bar{x}\}) \subseteq T(S, \bar{x}) \subseteq \overline{\text{cono}(S - \{\bar{x}\})}$$

Teorema 4.4. Sean $(X, \| \, \|)$ un e.v.n y $S \subseteq X$ no vacío. Si $\bar{x} \in S$, entonces $T(S, \bar{x})$ es convexo.

Demostración.

Sea $\bar{x} \in S$ (fijo) y consideremos una sucesión de vectores tangentes $(h_n) \subseteq T(S, \bar{x})$ tal que $h = \lim_{n\to\infty} h_n$ donde $h \in X$. Veamos que $h \in T(S, \bar{x})$. Como $h_n \in T(S, \bar{x})$ para cada $n \in \mathbb{N}$ existen $(x_{n_i})_{i\in\mathbb{N}} \subseteq S$ y $(\lambda_{n_i})_{i\in\mathbb{N}} \subseteq \mathbb{R}^+$ tales que $\bar{x} = \lim_{i\to\infty} x_{n_i}$ y $h_n = \lim_{i\to\infty} \lambda_{n_i}(x_{n_i} - \bar{x})$. Gracias a estas convergencias, para cada $n \in \mathbb{N}$ existe $i(n) \in \mathbb{N}$ tal que

$$\|x_{n_i} - \bar{x}\| \leq \frac{1}{n} \tag{4.2}$$

$$\|\lambda_{n_i}(x_{n_i} - \bar{x}) - h_n\| \leq \frac{1}{n} \tag{4.3}$$

para todo $i \geq i(n)$.

Definamos las sucesiones $y_n = x_{n_{i(n)}}$ y $t_n = \lambda_{n_{i(n)}}$ para cada $n \in \mathbb{N}$. Por (4.2) tenemos que $\bar{x} = \lim_{n\to\infty} y_n$, además, de las desigualdades (4.3), (4.2) y la definición de t_n e y_n obtenemos

$$\begin{aligned} \|t_n(y_n - \bar{x}) - h\| &= \left\|\lambda_{n_{i(n)}}(x_{n_{i(n)}} - \bar{x}) - h_n + h_n - h\right\| \\ &\leq \frac{1}{n} + \|h_n - h\| \end{aligned} \tag{4.4}$$

Tomando límite cuando $n \to \infty$ en (4.4), logramos la convergencia $h = \lim_{n\to\infty} t_n(y_n - \bar{x})$. Por lo tanto de la Definición 4.1 tenemos que $h \in T(S, \bar{x})$ consiguiendo el resultado.

$\square$

Como consecuencia del Teorema 4.4 y la observación posterior al Lema 4.3 tenemos el

Corolario 4.5. Si S es estrellado respecto a $\bar{x} \in S$ entonces $T(S, \bar{x}) = \overline{\text{cono}(S - \{\bar{x}\})}$.

Teorema 4.6. Sean $(X, \| \, \|)$ un e.v.n y $S \subseteq X$ no vacío y convexo. Entonces $T(S, \bar{x})$ es un cono convexo para todo $\bar{x} \in S$.

Demostración.

Consideremos $\bar{x} \in S$ (fijo arbitrario) y sean $h_1, h_2 \in T(S, \bar{x})$. Si $h_1 = 0$ o $h_2 = 0$ entonces $\lambda h_1 + (1 - \lambda)h_2$ pertenece a $T(S, \bar{x})$, puesto que $T(S, \bar{x})$ es un cono. Consideremos el caso en que h_1, h_2 son no nulos. Por la Definición 4.1 existen (x_n) y (y_n) en S y (λ_n) y (μ_n) en $\mathbb{R}^+$ tales que

$$\bar{x} = \lim_{n\to\infty} x_n, \; h = \lim_{n\to\infty} \lambda_n(x_n - \bar{x}) \tag{4.5}$$

$$\bar{x} = \lim_{n\to\infty} y_n, \; h = \lim_{n\to\infty} \mu_n(x_n - \bar{x}) \tag{4.6}$$

Definimos las sucesiones $\theta_n = \lambda_n + \mu_n$ y $z_n = \dfrac{1}{\theta_n}(\lambda_n x_n + \mu_n y_n)$. Claramente $\theta_n \in \mathbb{R}^+$ y por la convexidad de S y la definición de θ_n tenemos que $z_n \in S$ para todo $n \in \mathbb{N}$. Por lo tanto de las convergencias (4.5), (4.6), la definición de θ_n y siendo las sucesiones (λ_n/θ_n) y (μ_n/θ_n) sucesiones reales acotadas, obtenemos

$$\begin{aligned}
\lim_{n\to\infty} z_n &= \lim_{n\to\infty} \frac{1}{\theta_n}(\lambda_n x_n + \mu_n y_n) \\
&= \lim_{n\to\infty} \frac{1}{\theta_n}(\lambda_n x_n - \lambda_n \bar{x} + \mu_n y_n - \mu_n \bar{x} + \lambda_n \bar{x} + \mu_n \bar{x}) \\
&= \lim_{n\to\infty} \frac{\lambda_n}{\theta_n}(x_n - \bar{x}) + \frac{\mu_n}{\theta_n}(y_n - \bar{x}) + \bar{x} = \bar{x} \qquad (4.7)
\end{aligned}$$

y

$$\begin{aligned}
\lim_{n\to\infty} \theta_n(z_n - \bar{x}) &= \lim_{n\to\infty} \lambda_n x_n + \mu_n y_n - \theta_n \bar{x} \\
&= \lim_{n\to\infty} \lambda_n x_n + \mu_n y_n - (\lambda_n + \mu_n)\bar{x} \\
&= \lim_{n\to\infty} \lambda_n (x_n - \bar{x}) + \mu_n (y_n - \bar{x}) = h_1 + h_2
\end{aligned}$$

Mostrando así que $h_1 + h_2 \in T(S, \bar{x})$. Por lo tanto $T(S, \bar{x})$ es un cono convexo.

$\square$

4.2. Condiciones de optimalidad sobre conos tangentes

En esta sección enunciaremos algunos resultados respecto a los puntos minimizantes de funcionales, respecto a los conos tangentes generados por los puntos mínimizantes. Comenzaremos con el siguiente resultado.

Teorema 4.7. Sean $(X, \| \|)$ un e.v.n, $S \subseteq X$ no vacío y $f : X \longrightarrow \mathbb{R}$ una funcional.

(i) Si f es continua y convexa entonces para cada punto de minimización $\bar{x} \in S$ de f en S se cumple

$$f(\bar{x}) \leq f(\bar{x} + h) \qquad (4.8)$$

para todo $h \in T(S, \bar{x})$.

(ii) Si S es estrellado respecto a $\bar{x} \in S$ y se satisface la desigualdad (4.8) entonces $\bar{x}$ es el punto mínimo de f en S.

Demostración.

(i) Sea $\bar{x} = \underset{x\in S}{argmin}\, f(x)$. Supongamos que exista $h_1 \in T(S, \bar{x})$ tal que $f(\bar{x}) > f(\bar{x} + h_1)$ (hipótesis auxiliar). Evidentemente $h_1 \neq 0$. Escogemos $\alpha > 0$ tal que $f(\bar{x}) - f(\bar{x} + h_1) > \alpha > 0$. Como $h_1 \in T(S, \bar{x})$ existen dos sucesiones $(x_n)_{n\in\mathbb{N}} \subseteq S$ y $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^+$ tales que $\bar{x} = \lim_{n\to\infty} x_n$

y $h_1 = \lim_{n\to\infty} \lambda_n(x_n - \bar{x})$. Sea la sucesión $h_n = \lambda_n(x_n - \bar{x})$, como $\lim_{n\to\infty} h_n = h_1$ y $h_1 \neq 0$, se tiene que $\lim_{n\to\infty} 1/\lambda_n = 0$, luego para n suficientemente grande $1/\lambda_n$ pertenece al intervalo $(0,1)$. Para dichos índices, por ser S un conjunto estrellado resulta que

$$\begin{aligned} \bar{x} + h_n &= \bar{x} + \lambda_n(x_n - \bar{x}) \\ &= \lambda_n x_n + (1 - \lambda_n)\overline{x} \end{aligned} \tag{4.9}$$

pertenece a S para n suficientemente grande. Luego por la convexidad de f, la definición de h_n y la identidad (4.9) para n suficientemente grande tenemos

$$\begin{aligned} f(x_n) &= f(\frac{1}{\lambda_n}\bar{x} + x_n - \bar{x} + \bar{x} - \frac{1}{\lambda_n}\bar{x}) \\ &= f(\frac{1}{\lambda_n}(\bar{x} + h_n) + (1 - \frac{1}{\lambda_n})\bar{x}) \\ &\leq \frac{1}{\lambda_n} f(\bar{x} + h_n) + (1 - \frac{1}{\lambda_n}) f(\bar{x}) \end{aligned} \tag{4.10}$$

Por la continuidad de f, la convergencia de (h_n) y la elección de α logramos de (4.10)

$$\begin{aligned} f(x_n) &\leq \frac{1}{\lambda_n} f(\bar{x} + h_n) + (1 - \frac{1}{\lambda_n}) f(\bar{x}) \\ &\leq \frac{1}{\lambda_n} f(\bar{x} + h_1) + (1 - \frac{1}{\lambda_n}) f(\bar{x}) \\ &\leq \frac{1}{\lambda_n} \{f(\bar{x} + h_1) + \alpha\} + (1 - \frac{1}{\lambda_n}) f(\bar{x}) \\ &\leq \frac{1}{\lambda_n} f(\bar{x}) + (1 - \frac{1}{\lambda_n}) f(\bar{x}) = f(\bar{x}) \end{aligned}$$

Lo que contradice la minimalidad de $\bar{x}$. Por lo tanto se tiene (4.8).

(ii) Como S es estrellado respecto de $\bar{x}$, por el Lema 4.2 tenemos que $S - \{\bar{x}\} \subseteq T(S, \bar{x})$. Así por la hipótesis tenemos $f(\bar{x}) \leq f(\bar{x} + h)$ para todo $h \in S - \{\bar{x}\}$. Si $h \in S - \{\bar{x}\}$ entonces existe $x_1 \in S$ tal que $h = x_1 - \bar{x}$, de donde $x_1 = h + \bar{x}$. Luego obtenemos $f(\bar{x}) \leq f(x_1)$ Así, dado $x_1 \in S$ concluimos que $f(\bar{x}) \leq f(x_1)$. En consecuencia $\bar{x} = \underset{x\in S}{argmin}\, f(x)$.

□

Teorema 4.8. Consideremos $(X, \| \ \|)$ un e.v.n, $S \subseteq X$ no vacío y $f : U \longrightarrow \mathbb{R}$ una funcional, donde $U \subseteq X$ es abierto y $S \subseteq U$. Si $\bar{x} \in S$ es punto minimizante de f en S y f es Fréchet diferenciable en $\bar{x}$, entonces $f'(\bar{x})h \geq 0$ para todo $h \in T(S, \bar{x})$.

Demostración.

Consideremos $\bar{x} = \underset{x\in S}{argmin}\, f(x)$ y $h \in T(S, \bar{x})$. Si $h = 0$ no hay nada que demostrar. Supongamos que $h \neq 0$. Por hipótesis existen dos sucesiones $(x_n)_{n\in\mathbb{N}} \subseteq S$ y $(\lambda_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^+$ tales que

$\bar{x} = \lim\limits_{n\to\infty} x_n$ y $h = \lim\limits_{n\to\infty} \lambda_n(x_n - \bar{x})$. Como f es Fréchet diferenciable en $\bar{x}$, por la linealidad y continuidad de $f'(\bar{x})$, tenemos

$$\begin{aligned} f'(\bar{x})h &= f'(\bar{x}) \lim_{n\to\infty} \lambda_n(x_n - \bar{x}) \\ &= \lim_{n\to\infty} \lambda_n f'(\bar{x})(x_n - \bar{x}) \\ &= \lim_{n\to\infty} \lambda_n(f(x_n) - f(\bar{x}) - \big\{f(x_n) - f(\bar{x}) - f'(\bar{x})(x_n - \bar{x})\big\} \end{aligned} \tag{4.11}$$

Como $x_n \in S$, por la minimalidad de $\bar{x}$ en S resulta que $f(x_n) - f(\bar{x}) \geq 0$. Por tanto de (4.11) logramos

$$\begin{aligned} f'(\bar{x})h &\geq -\lim_{n\to\infty} \lambda_n \left\{f(x_n) - f(\bar{x}) - f'(\bar{x})(x_n - \bar{x})\right\} \\ &= -\lim_{n\to\infty} \|h_n\| \left\{\frac{f(x_n) - f(\bar{x}) - f'(\bar{x})(x_n - \bar{x})}{\|x_n - \bar{x}\|}\right\} \end{aligned} \tag{4.12}$$

donde $\|h_n\| = \lambda_n \|x_{n-}\bar{x}\|$. De la desigualdad (4.12), la Definición 3.8 ($h = x_n - \bar{x}$) y siendo f Fréchet diferenciable en $\bar{x}$, obtenemos que $f'(\bar{x})h \geq 0$ para todo $h \in T(S, \bar{x})$.

□

4.3. Condiciones de optimalidad para funcionales pseudoconvexas

Definición 4.9. Sean X un espacio vectorial, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional con derivada direccional en $\bar{x} \in S$ para toda dirección $x - \bar{x}$ con $x \in S$. Diremos que f es *pseudoconvexa* en el punto $\bar{x}$ si para todo $x \in S$ se cumple

$$f'(\bar{x})(x - \bar{x}) \geq 0 \implies f(x) \geq f(\bar{x})$$

Diremos que f es pseudoconvexa en S si f es pseudoconvexa en x, para todo $x \in S$.

Observación.

En la Definicion 4.9 podemos considerar alternativamente que f es pseudoconvexa en $\bar{x}$ si $f(x) < f(\bar{x}) \implies f'(\bar{x})(x - \bar{x}) < 0$ para todo $x \in S$.

Ejemplo.

Sea la funcion real de variable real $f(x) = x^3 + x$, $x \in \mathbb{R}$. Afirmamos que f es pseudoconvexa en $\mathbb{R}$. En efecto, sea $\bar{x} \in \mathbb{R}$. Claramente $f'(\bar{x}) = 3\bar{x}^2 + 1 > 0$. Si $f(x) < f(\bar{x})$ tenemos

$$
\begin{aligned}
x^3 + x < \bar{x}^3 + \bar{x} &\Longrightarrow x^3 - \bar{x}^3 + x - \bar{x} < 0 \\
&\Longrightarrow (x - \bar{x})\underbrace{(x^2 + x\bar{x} + \bar{x}^2 + 1)}_{>0} < 0 \\
&\Longrightarrow (x - \bar{x}) < 0.
\end{aligned}
$$

Por lo tanto, $f\,'(\bar{x})(x - \bar{x}) < 0$ para todo $x \in \mathbb{R}$, mostrando así nuestra afirmación.

Lema 4.10. Consideremos X un espacio vectorial, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional convexa. Si existe la derivada direccional en $\bar{x} \in S$ en toda dirección $x - \bar{x}$, para todo $x \in S$. Entonces f es pseudoconvexa en $\bar{x}$.

Demostración.

Consideremos $x \in S$. Por la convexidad de f en S para $\lambda \in [0, 1]$ tenemos

$$
f(\lambda x + (1 - \lambda)\bar{x}) \leq \lambda f(x) + (1 - \lambda) f(\bar{x}) \tag{4.13}
$$

Transponiendo términos en (4.13) logramos la desigualdad

$$
\begin{aligned}
f(x) &\geq f(\bar{x}) + \frac{1}{\lambda}\{f(\lambda x + (1 - \lambda)\bar{x}) - f(\bar{x})\} \\
&= f(\bar{x}) + \frac{1}{\lambda}\{f(\bar{x} + \lambda(x - \bar{x})) - f(\bar{x})\} \qquad (4.14)
\end{aligned}
$$

Tomando el límite cuando $\lambda \to 0^+$ en (4.14) y por hipótesis obtenemos la desigualdad $f(x) \geq f(\bar{x}) + f'(\bar{x})(x - \bar{x})$. Por lo tanto,

$$
f(x) - f(\bar{x}) \geq f'(\bar{x})(x - \bar{x}) \tag{4.15}
$$

Si $f'(\bar{x})(x - \bar{x}) \geq 0$ entonces de (4.15) conseguimos $f(x) \geq f(\bar{x})$, demostrando la pseudoconvexidad de f en $\bar{x}$.

□

Teorema 4.11. Sean $(X, \|\ \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío y $f : U \longrightarrow \mathbb{R}$ una funcional, donde U es un conjunto abierto que contiene a S. Si f es Fréchet diferenciable y pseudoconvexa en S, entonces f es cuasiconvexa en S.

Demostración.

Sea $\alpha \in \mathbb{R}$ y supongamos que $S_\alpha = \{x \in S / f(x) \leq \alpha\}$ es no vacío. Afirmamos que S_α es convexo. En efecto, sean $x, y \in S_\alpha$. Supongamos lo contrario, entonces existe $\hat{\lambda} \in [0, 1]$ tal que $f(\hat{\lambda}x + (1 - \hat{\lambda})\bar{x}) > \alpha$. Para $\alpha \geq \max\{f(x), f(y)\}$, si tomamos $\hat{\lambda} = 0$ o $\hat{\lambda} = 1$, resulta que tanto $f(x) > \alpha = \text{máx}\{f(x), f(y)\}$ o $f(y) > \alpha = \text{máx}\{f(x), f(y)\}$ lo cual es absurdo. Por lo tanto $\hat{\lambda} \in (0,1)$. Como f es Fréchet diferenciable en S, por el Corolario 3.10 tenemos que f es continua en S. Así, la aplicación continua

$$\begin{aligned} \psi : (0,1) &\longrightarrow \mathbb{R} \\ \lambda &\longrightarrow f(\lambda x + (1-\lambda)y) \end{aligned}$$

alcanza su valor máximo en $(0,1)$, es decir existe $\bar{\lambda} \in (0,1)$ tal que $f(\bar{\lambda}x + (1-\bar{\lambda})y) \geq f(\lambda x + (1-\lambda)y)$. Denotando por $\bar{x} = \bar{\lambda}x + (1-\bar{\lambda})y$ tenemos que $\bar{x} = \underset{x \in \hat{S}}{argminf(x)}$ donde $\hat{S} = [x, y]$, el segmento cerrado que une los puntos x e y. Por los teoremas 3.6 -(i) y 3.9 (para el caso del punto máximo $\bar{x}$ en $\hat{S}$) obtenemos

$$f'(\bar{x})(x - \bar{x}) \leq 0 \text{ y } f'(\bar{x})(y - \bar{x}) \leq 0 \tag{4.16}$$

De la definición de $\bar{x}$, logramos las identidades:

$$x - \bar{x} = x - \bar{\lambda}x + (1-\bar{\lambda})y = (1-\bar{\lambda})(x - y) \tag{4.17}$$

$$y - \bar{x} = y - \bar{\lambda}x + (1-\bar{\lambda})y = -\bar{\lambda}(x - y) \tag{4.18}$$

En consecuencia, de (4.16), (4.17), (4.18) y la linealidad de $f'(\bar{x})$ resulta que:

$$0 \geq f'(\bar{x})(x - \bar{x}) = (1-\bar{\lambda})f'(\bar{x})(x - y) \Longrightarrow f'(\bar{x})(x - y) \leq 0 \tag{4.19}$$

$$0 \geq f'(\bar{x})(y - \bar{x}) = -\bar{\lambda}f'(\bar{x})(x - y) \Longrightarrow f'(\bar{x})(x - y) \geq 0 \tag{4.20}$$

De (4.19) y (4.20) conseguimos $f'(\bar{x})(x-y) = 0$. Luego de (4.18) obtenemos $f'(\bar{x})(y-\bar{x}) = 0$. Así por la pseudoconvexidad de f resulta que $f(y) - f(\bar{x}) \geq 0$. Por otra parte

$$\begin{aligned} f(y) - f(\bar{x}) &= f(y) - f(\bar{\lambda}x + (1-\bar{\lambda})y) \\ &\leq f(y) - f(\hat{\lambda}x + (1-\hat{\lambda})y) \\ &< f(y) - \alpha \leq 0 \end{aligned}$$

Con lo cual obtenemos $f(y) - f(\bar{x}) < 0$, lo que contradice la desigualdad anterior. Por lo tanto S_α es convexo y por la Definición 1.38 se sigue que f es cuasiconvexa.

□

Observación.

En el caso que $f : X \longrightarrow \mathbb{R}$ es una funcional Fréchet diferenciable en $\bar{x} \in X$, se tienen las implicaciones: f convexa $\Longrightarrow f$ es pseudoconvexa en todo $\bar{x} \in X \Longrightarrow f$ es cuasiconvexa.

Teorema 4.12. Sean $(X, \| \, \|)$ un e.v.n, $S \subseteq X$ un conjunto convexo no vacío, $U \subseteq X$ un conjunto abierto que contiene a S y $f : U \longrightarrow \mathbb{R}$ una funcional. Supongamos que:

(i) S es estrellado respecto a $\bar{x} \in S$.

(ii) f es direccionalmente diferenciable y pseudoconvexa en $\bar{x}$.

(iii) $f'(\bar{x})h \geq 0$ para todo $h \in T(S, \bar{x})$.

Entonces $\bar{x} = \underset{x \in S}{argmin}\, f(x)$.

Demostración.

Como S es estrellado respecto a $\bar{x}$ por el Lema 4.2 tenemos que $S - \{\bar{x}\} \subseteq T(S, \bar{x})$. Así por la hipótesis (ii) se cumple $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$ y por la pseudoconvexidad de f tenemos la desigualdad $f(x) - f(\bar{x}) \geq 0$ para todo $x \in S$. En consecuencia $\bar{x} = \underset{x \in S}{argmin}\, f(x)$. □

De los resultados de los Teoremas 4.12 y 3.6-(i) tenemos la siguiente caracterización.

Corolario 4.13. Sean X un e.v.n, $S \subseteq X$ no vacío y $f : S \longrightarrow \mathbb{R}$ una funcional. Si f posee derivada direccional en $\bar{x} \in S$ en toda dirección $x - \bar{x}$ con $x \in S$ y es pseudoconvexa en $\bar{x}$ entonces $\bar{x} = \underset{x \in S}{argmin}\, f(x)$ si y sólo si $f'(\bar{x})(x - \bar{x}) \geq 0$ para todo $x \in S$.

4.4. El teorema de Lyusternik en espacios de Banach

En esta sección estudiaremos una extensión del Teorema de Lyusternik, muy empleado en problemas de optimización en dimensión finita, a espacios de Banach de cualquier dimensión, sobre funcionales Fréchet diferenciables, siendo este resultado utilizado en el estudio de problemas de optimización restringida con condiciones suficientemente regulares.

Teorema 4.14. (Teorema de Lyusternik en espacios de Banach)

Sean $(X, \| \, \|_X)$ y $(Z, \| \, \|_Z)$ dos espacios de Banach reales y $h : X \longrightarrow Z$ una aplicación. Consideremos el conjunto factible $S = \{x \in X / h(x) = 0_Z\}$ y el punto $\bar{x} \in S$. Supongamos que:

(i) h es Fréchet diferenciable en una vecindad de $\bar{x}$.

(ii) $h'(\cdot)$ es continua en $\bar{x}$.

(iii) $h'(\bar{x})$ es una aplicación sobreyectiva.

Entonces $W = \{x \in X / h'(\bar{x})x = 0_Z\}$ es un subconjunto de $T(S, \bar{x})$.

Demostración.

Por hipótesis (iii) la aplicación $h'(\bar{x})$ es lineal, continua y sobreyectiva. Así por el teorema de la aplicación abierta (Teorema 1.17) $h'(\bar{x})$ es una funcional lineal abierta. Como $B(0_X, 1)$

es abierto en X entonces $h'(\bar{x})(B(0_X, 1))$ es abierto en Z por tanto existe $\rho_1 > 0$ tal que $B(0_Z, \rho_1) \subseteq h'(\bar{x})(B(0_X, 1))$ puesto que 0_Z es punto interior de $h'(\bar{x})(B(0_X, 1))$.

Afirmación 1 $h'(\bar{x})(B(0_X, 1))$ es acotado.

En efecto, sea $w \in h'(\bar{x})(B(0_X, 1))$, luego existe $u \in B(0_X, 1)$ tal que $w = h'(\bar{x})u$. Como $h'(\bar{x}) \in L(X, Z)$, tenemos que

$$\begin{aligned} \|w\|_Z &= \left\|h'(\bar{x})u\right\|_Z \\ &\leq \left\|h'(\bar{x})\right\|_{Z^\star} \|u\|_X \leq \left\|h'(\bar{x})\right\| \end{aligned}$$

resultando que $w \in B(0_Z, \|h'(\bar{x})\|_{Z^\star})$, demostrando nuestra afirmación.

Consideremos el conjunto

$$\Gamma = \{\rho > 0 / B(0_z, \rho) \subseteq h'(\bar{x})(B(0_X, 1))\}$$

Claramente $\rho_1 \in \Gamma$, así Γ es no vacío y por la Afirmación 1, acotado superiormente. Sea $\rho_0 = \sup \Gamma$. Consideremos $\epsilon \in (0, \frac{\rho_0}{2})$ arbitrario. Por la hipótesis (ii) existe $\delta > 0$ tal que

$$\left\|h'(\tilde{x}) - h'(\bar{x})\right\|_{L(X,Z)} \leq \epsilon \tag{4.21}$$

para todo $\tilde{x} \in B(\bar{x}, 2\delta)$. Tomemos $\tilde{x}_1, \tilde{x}_2 \in B(\bar{x}, 2\delta)$, como $h(\tilde{x}_2) - h(\tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1) \in Z$ por el Teorema 1.49, existe $l \in Z^\star$ tal que $\|l\|_{Z^\star} = 1$ y

$$l(h(\tilde{x}_2) - h(\tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)) = \left\|h(\tilde{x}_2) - h(\tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)\right\|_Z \tag{4.22}$$

Definimos la función

$$\begin{aligned} \varphi : [0,1] &\longrightarrow \mathbb{R} \\ t &\longmapsto \varphi(t) = l(h(\tilde{x}_1 + t(\tilde{x}_2 - \tilde{x}_1) - th'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)) \end{aligned}$$

Dado que $l \in Z^\star$ podemos escribir $\varphi(t)$ como

$$\varphi(t) = (l \circ h)(\tilde{x}_1 + t(\tilde{x}_2 - \tilde{x}_1)) - t(l \circ h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1))$$

para cada $t \in [0,1]$. Observemos que $\varphi(0) = l \circ h(\tilde{x}_1)$ y $\varphi(1) = l(h(\tilde{x}_2) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1))$. Por la convexidad de $B(\bar{x}, 2\delta)$ tenemos que $\tilde{x}_1 + t(\tilde{x}_2 - \tilde{x}_1) \in B(\bar{x}, 2\delta)$ para todo $t \in [0,1]$ y por hipótesis h es Fréchet diferenciable en una vecindad de $\bar{x}$. En consecuencia, la aplicación φ es diferenciable en $(0,1)$ y por la regla de la cadena logramos

$$\begin{aligned}\varphi'(t) &= l(h'(\tilde{x}_1 + t(\tilde{x}_2 - \tilde{x}_1))(\tilde{x}_2 - \tilde{x}_1) - l(h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)) \\ &= l(h'(\tilde{x}_1 + t(\tilde{x}_2 - \tilde{x}_1))(\tilde{x}_2 - \tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1))\end{aligned} \tag{4.23}$$

Luego, por el teorema del valor medio para derivadas, existe $\bar{t} \in (0,1)$ tal que $\varphi(1) - \varphi(0) = \varphi'(\bar{t})$. Notemos que

$$\varphi(1) - \varphi(0) = l(h(\tilde{x}_2) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)) - l \circ h(\tilde{x}_1) \tag{4.24}$$

Por lo tanto de (4.21), (4.22), (4.23) y (4.24) conseguimos la desigualdad

$$\begin{aligned}\left\|h(\tilde{x}_2) - h(\tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)\right\|_Z &= l(h(\tilde{x}_2) - h(\tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)) \\ &= \varphi(1) - \varphi(0) \\ &= \varphi'(\bar{t}) \\ &= l(h'(\tilde{x}_1 + \bar{t}(\tilde{x}_2 - \tilde{x}_1))(\tilde{x}_2 - \tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)) \\ &\leq \|l\|_{Z^\star} \left\|h'(\tilde{x}_1 + \bar{t}(\tilde{x}_2 - \tilde{x}_1)) - h'(\bar{x})\right\|_{L(X,Z)} \|\tilde{x}_2 - \tilde{x}_1\|_X\end{aligned}$$

De esta desigualdad, considerando que $\|l\|_{Z^\star} = 1$ y $\tilde{x}_1 + \bar{t}(\tilde{x}_2 - \tilde{x}_1) \in B(\bar{x}, 2\delta)$ logramos

$$\left\|h(\tilde{x}_2) - h(\tilde{x}_1) - h'(\bar{x})(\tilde{x}_2 - \tilde{x}_1)\right\|_Z \leq \epsilon \|\tilde{x}_2 - \tilde{x}_1\|_X \tag{4.25}$$

para cualesquiera $\tilde{x}_1, \tilde{x}_2 \in B(\bar{x}, 2\delta)$. Por la elección de ϵ tenemos que $\dfrac{\epsilon}{\rho_0} < \dfrac{1}{2}$. Tomemos $\alpha > 1$ tal que $\alpha\left(\dfrac{\epsilon}{\rho_0} + \dfrac{1}{2}\right) < 1$. Sea $x \in W$. Si $x = 0_X$ el teorema se satisface puesto que $0_X \in T(S, \bar{x})$. Consideremos el caso en que $x \neq 0_X$ y tomemos $\lambda \in (0, \hat{\lambda}]$ fijo pero arbitrario, donde $\hat{\lambda} = \dfrac{\delta}{\|x\|_X}$. Como $h'(\bar{x})$ es sobreyectiva, podemos definir las sucesiones (r_n) y (u_n) de la siguiente manera:

Para $r_1 = 0_X$, existe $u_1 \in X$ tal que $h'(\bar{x})u_1 = h(\bar{x} + \lambda x + r_1)$. Definimos $r_2 = r_1 - u_1$, para este elemento, existe $u_2 \in X$ tal que $h'(\bar{x})u_2 = h(\bar{x} + \lambda x + r_2)$, del mismo modo, para $r_3 = r_2 - u_2$ existe $u_3 \in X$ tal que $h'(\bar{x})u_3 = h(\bar{x} + \lambda x + r_3)$. En consecuencia, para cada $n \in \mathbb{N}$ tenemos que (r_n) y (u_n) satisfacen:

$$\begin{aligned}h'(\bar{x})u_n &= h(\bar{x} + \lambda x + r_n) \\ r_{n+1} &= r_n - u_n\end{aligned} \tag{4.26}$$

Observemos que por la definición de Γ y $\alpha > 1$ tenemos $B(0_Z, \dfrac{\rho_0}{\alpha}) \subseteq h'(\bar{x})(B(0_X, 1))$.

Afirmación 2 $\|u_n\|_X \leq \dfrac{\alpha}{\rho_0} \|h(\bar{x} + \lambda x + r_n)\|_Z$ para todo $n \in \mathbb{N}$.

En efecto, si $u_n = 0_X$ la desigualdad es evidente. Supongamos que $u_n \neq 0_X$ y definamos $w_n = \dfrac{u_n}{\|u_n\|_X}$. De esta forma $w_n \in S(0_X, 1)$, siendo este conjunto la frontera de la bola unitaria en X. Así tenemos

$$\begin{aligned} B(0_Z, \frac{\rho_0}{\alpha}) &\subseteq B(0_Z, \rho_0) \\ &\subseteq h'(\bar{x})(B(0_X, 1)) \end{aligned}$$

De esta última inclusión, como $w_n \in S(0_X, 1)$ tenemos

$$\frac{\rho_0}{\alpha} \leq \left\| h'(\bar{x}) \frac{u_n}{\|u_n\|_X} \right\|_Z \tag{4.27}$$

Por (4.26) y (4.27) tenemos

$$\begin{aligned} \|u_n\| &\leq \frac{\alpha}{\rho_0} \left\| h'(\bar{x}) u_n \right\|_Z \\ &= \frac{\alpha}{\rho_0} \left\| h(\bar{x} + \lambda x + r_n) \right\|_Z \end{aligned}$$

Logrando de esta manera la Afirmación 2.

Definimos $d(\lambda) = \|h(\bar{x} + \lambda x)\|_Z$ y $q = \dfrac{\epsilon\alpha}{\rho_0}$. Por la elección de λ obtenemos

$$\begin{aligned} \|\lambda x\|_X &= \lambda \|x\|_X \\ &\leq \hat{\lambda} \|x\|_X \\ &= \delta \end{aligned}$$

Por lo tanto $\|\lambda x\|_X \leq \delta$, en consecuencia $\|\bar{x} + \lambda x - \bar{x}\| \leq \delta < 2\delta$. Luego por las hipótesis sobre x y $\bar{x}$, la desigualdad anterior y (4.25), logramos

$$\begin{aligned} d(\lambda) &= \left\| h(\bar{x} + \lambda x) - h(\bar{x}) - h'(\bar{x})(\lambda x) \right\|_Z \\ &\leq \epsilon \|\lambda x\|_X \leq \epsilon\delta \end{aligned} \tag{4.28}$$

Como $\alpha > 1$ tenemos $q = \dfrac{\epsilon\alpha}{\rho_0} \leq 1 - \dfrac{\alpha}{2} < \dfrac{1}{2}$.

Afirmación 3 Se cumplen las siguientes desigualdades:

(a) $\|r_n\|_X \leq \dfrac{\alpha}{\rho_0} d(\lambda) \left(\dfrac{1 - q^{n-1}}{1 - q} \right)$.

(b) $\|h(\bar{x} + \lambda x + r_n)\|_X \leq d(\lambda) q^{n-1}$.

(c) $\|u_n\|_X \leq \dfrac{\alpha}{\rho_0} d(\lambda) q^{n-1}$.

para cualquier $n \in \mathbb{N}$. En efecto, por inducción matemática sobre n. Para $n = 1$ tenemos:

(a) Como $r_1 = 0_X$ tenemos $0 = \|r_1\|_X \leq \dfrac{\alpha}{\rho_0} d(\lambda)(0)$.

(b) $\|h(\bar{x} + \lambda x + r_1)\|_Z = \|h(\bar{x} + \lambda x)\|_Z = d(\lambda)$ (por definición de $d(\lambda)$).

(c) De la Afirmación 2 y (b) tenemos

$$
\begin{aligned}
\|u_1\|_X &\leq \frac{\alpha}{\rho_0} \|h(\bar{x} + \lambda x + r_1)\|_Z \\
&= \frac{\alpha}{\rho_0} \|h(\bar{x} + \lambda x)\|_Z \\
&= \frac{\alpha}{\rho_0} d(\lambda)
\end{aligned}
$$

Supongamos válidas las tres afirmaciones para n (hipótesis inductiva). Para $n+1$ tenemos,

(a) Por la definición de r_n, la desigualdad triangular y las hipótesis inductivas (a) y (c) conseguimos

$$
\begin{aligned}
\|r_{n+1}\|_X &= \|r_n - u_n\|_X \\
&\leq \|r_n\|_X + \|u_n\|_X \\
&\leq \frac{\alpha}{\rho_0} d(\lambda) \left(\frac{1 - q^{n-1}}{1 - q} \right) + \frac{\alpha}{\rho_0} d(\lambda) q^{n-1} \\
&= \frac{\alpha}{\rho_0} d(\lambda) \left(\frac{1 - q^{n-1}}{1 - q} + q^{n-1} \right) = \frac{\alpha}{\rho_0} d(\lambda) \left(\frac{1 - q^n}{1 - q} \right)
\end{aligned}
$$

(b) Por hipótesis inductiva, la definición de q y (4.28) tenemos

$$
\begin{aligned}
\|\lambda x + r_n\|_X &\leq \|\lambda x\|_X + \|r_n\|_X \\
&\leq \delta + \frac{\alpha}{\rho_0} d(\lambda) \left(\frac{1 - q^{n-1}}{1 - q} \right) \\
&\leq \delta + \frac{\alpha}{\rho_0} (\epsilon\delta) \left(\frac{1 - q^{n-1}}{1 - q} \right) \\
&\leq \delta + \frac{\delta q}{1 - q} \left(1 - q^{n-1} \right) \\
&\leq \delta \left(1 + \frac{q}{1 - q} \left\{ 1 - q^{n-1} \right\} \right) \leq 2\delta
\end{aligned}
\tag{4.29}
$$

y

$$
\begin{aligned}
\|\lambda x + r_n - u_n\|_X &= \|\lambda x + r_{n+1}\|_X \\
&\leq \|\lambda x\|_X + \|r_{n+1}\|_X \\
&\leq \delta + \frac{\alpha}{\rho_0} d(\lambda) \left(\frac{1 - q^n}{1 - q} \right) \\
&= \delta \left(1 + \frac{q}{1 - q} \left\{ 1 - q^{n-1} \right\} \right) \leq 2\delta
\end{aligned}
\tag{4.30}
$$

puesto que el producto $\frac{q}{1-q}\left\{1-q^{n-1}\right\}$ es menor que 1 y $q=\frac{\epsilon\alpha}{\rho_0}$.

Luego, por la definición de u_n y la linealidad de $h'(\bar{x})$ conseguimos

$$\begin{aligned}\|h(\bar{x}+\lambda x+r_{n+1})\|_Z &= \|h(\bar{x}+\lambda x+r_n-u_n)\|_Z\\ &= \|h(\bar{x}+\lambda x+r_n)-h(\bar{x}+\lambda x+r_n)+h(\bar{x}+\lambda x+r_n-u_n)\|_Z\\ &= \|-h(\bar{x})(-u_n)-h(\bar{x}+\lambda x+r_n)+h(\bar{x}+\lambda x+r_n-u_n)\|_Z\\ &= \|h(\bar{x}+\lambda x+r_n-u_n)-h(\bar{x}+\lambda x+r_n)-h(\bar{x})(-u_n)\|_Z\end{aligned} \tag{4.31}$$

Tomando $\tilde{x}_2=\bar{x}+\lambda x+r_n-u_n$ y $\tilde{x}_1=\bar{x}+\lambda x+r_n$, de (4.29) y (4.30) logramos

$$\begin{aligned}\tilde{x}_2-\tilde{x}_1 &= (\bar{x}+\lambda x+r_n-u_n)-(\bar{x}+\lambda x+r_n)=-u_n\\ \|\tilde{x}_2-\bar{x}\|_X &= \|(\bar{x}+\lambda x+r_n-u_n)-\bar{x}\|_X=\|\lambda x+r_n-u_n\|_X<2\delta\\ \|\tilde{x}_1-\bar{x}\|_X &= \|\lambda x+r_n\|_X=<2\delta\end{aligned} \tag{4.32}$$

De (4.32) concluimos que $\bar{x}+\lambda x+r_n-u_n$ y $\bar{x}+\lambda x+r_n$ pertenecen a $B(\bar{x},2\delta)$. Luego de (4.31), (4.25) y la hipótesis inductiva (c) obtenemos

$$\begin{aligned}\|h(\bar{x}+\lambda x+r_{n+1})\|_Z &\le \epsilon\,\|-u_n\|\\ &\le \frac{\epsilon\alpha}{\rho_0}d(\lambda)q^{n-1}=d(\lambda)q^n\end{aligned} \tag{4.33}$$

(c) De (4.33), tenemos que

$$\begin{aligned}\|u_{n+1}\|_X &\le \frac{\alpha}{\rho_0}\,\|h(\bar{x}+\lambda x+r_{n+1})\|_Z\\ &\le \frac{\alpha}{\rho_0}d(\lambda)q^n\end{aligned}$$

Quedando demostrada la Afirmación 3.

Afirmación 4 La sucesión (r_n) es de Cauchy en X.

En efecto, como $r_{n+1}=r_n-u_n$ para $n,k\in\mathbb{N}$ tenemos

$$\begin{aligned}\|r_{n+k}-r_n\|_X &= \|r_{n+k-1}-u_{n+k-1}-r_n\|_X\\ &= \|r_{n+k-2}-u_{n+k-2}-u_{n+k-1}-r_n\|_X\\ &= \|r_n-u_n-u_{n+1}-u_{n+2}-\cdots-u_{n+k-2}-u_{n+k-1}-r_n\|_X\\ &= \|u_n+u_{n+1}+u_{n+2}+\cdots+u_{n+k-2}+u_{n+k-1}\|_X\\ &\le \|u_n\|_X+\|u_{n+1}\|_X+\|u_{n+2}\|_X+\cdots+\|u_{n+k-2}\|_X+\|u_{n+k-1}\|_X\end{aligned} \tag{4.34}$$

De la Afirmación 3 y la desigualdad (4.34) tenemos

$$\begin{aligned}\|r_{n+k}-r_n\|_X &\leq \|u_n\|_X+\|u_{n+1}\|_X+\cdots+\|u_{n+k-2}\|_X+\|u_{n+k-1}\|_X\\ &\leq \frac{\alpha}{\rho_0}d(\lambda)q^{n-1}+\frac{\alpha}{\rho_0}d(\lambda)q^{n}+\cdots+\frac{\alpha}{\rho_0}d(\lambda)q^{n+k-3}+\frac{\alpha}{\rho_0}d(\lambda)q^{n+k-2}\\ &= \frac{\alpha}{\rho_0}d(\lambda)q^{n-1}\left\{1+q+\cdots+q^{k-2}+q^{k-1}\right\} \qquad (4.35)\\ &= \frac{\alpha}{\rho_0}d(\lambda)q^{n-1}\left\{\frac{1-q^k}{1-q}\right\}<\frac{\alpha}{\rho_0}d(\lambda)q^{n-1}\left\{\frac{1}{1-q}\right\}=\frac{\alpha}{\rho_0}d(\lambda)\frac{q^{n-1}}{1-q}\end{aligned}$$

puesto que $1-q^k<1$. Por lo tanto $\|r_{n+k}-r_n\|_X<\dfrac{\alpha}{\rho_0}d(\lambda)\dfrac{q^{n-1}}{1-q}$. Haciendo tender $n\to\infty$ el término q^{n-1} tiende a cero, en consecuencia $\|r_{n+k}-r_n\|_X\to 0$ si $n\to\infty$. De donde, la sucesión (r_n) es de Cauchy en X, mostrando así la afirmación 4.

Como X es un espacio de Banach existe $r(\lambda)\in X$ tal que $\lim\limits_{n\to\infty} r_n=r(\lambda)$. Observemos que dicho límite depende de λ. Asimismo, siendo $u_n=r_{n+1}-r_n$, la sucesión (u_n) es convergente y por la Afirmación 3, logramos

$$\begin{aligned}\|u_n\|_X &\leq \frac{\alpha}{\rho_0}d(\lambda)q^{n-1}\\ &\leq \frac{\alpha}{\rho_0}(\epsilon\delta)q^{n-1}\\ &\leq \delta q^n\end{aligned}$$

Obteniendo la desigualdad $\|u_n\|_X\leq\delta q^n$. Tomando límite cuando $n\to\infty$ en la desigualdad anterior, tenemos que $(u_n)\to 0_X$.

Afirmación 5 $\lim\limits_{n\to\infty}\dfrac{r(\lambda)}{\lambda}=0_X$.
En efecto, como $h'(\bar{x})\in L(X,Z)$, $u_n\longrightarrow 0_X$, $r_n\longrightarrow r(\lambda)$ y por la identidad (4.26) tenemos

$$h(\bar{x}+\lambda x+r(\lambda))=h'(\bar{x})(0_X)=0 \qquad (4.36)$$

puesto que $\bar{x}+\lambda x+r_n\in B(\bar{x},2\delta)$. Por la Afirmación 3, conseguimos

$$\|r_n\|_X\leq\frac{\alpha}{\rho_0}d(\lambda)\left(\frac{1-q^{n-1}}{1-q}\right)$$

Tomando límite cuando $n\to\infty$ a la desigualdad anterior y dividiendo por λ conseguimos

$$\frac{\|r(\lambda)\|_X}{\lambda}\leq\frac{\alpha}{\lambda\rho_0}d(\lambda)\left(\frac{1}{1-q}\right) \qquad (4.37)$$

Por la definición de $d(\lambda)$ y las hipótesis sobre $\bar{x}$, dividiendo por $\|x\|_X$ y de (4.37) logramos

$$\begin{aligned}
\frac{\|r(\lambda)\|_X}{\lambda} &\leq \frac{\alpha}{\rho_0(1-q)}\frac{d(\lambda)}{\lambda} \\
&= \frac{\alpha}{\rho_0(1-q)}\frac{\|h(\bar{x}+\lambda x)-h(\bar{x})-\lambda h'(\bar{x})(x)\|_Z\,\|x\|_X}{\lambda\,\|x\|_X} \\
&= \frac{\alpha}{\rho_0(1-q)}\frac{\|h(\bar{x}+\lambda x)-h(\bar{x})-h'(\bar{x})(\lambda x)\|_Z}{\|\lambda x\|_X}\,\|x\|_X
\end{aligned} \tag{4.38}$$

puesto que $h(\bar{x}) = 0$ y $h'(\bar{x})x = 0$.

Por otra parte, como h es Fréchet diferenciable en $\bar{x}$ tenemos que

$$\lim_{\lambda\to 0^+} \frac{\|h(\bar{x}+\lambda x)-h(\bar{x})-h'(\bar{x})(\lambda x)\|_Z}{\|\lambda x\|_X} = 0 \tag{4.39}$$

En consecuencia, de (4.38) y (4.39) tenemos $\lim\limits_{n\to\infty}\dfrac{\|r(\lambda)\|}{\lambda} = 0_X$, mostrando de esta manera la Afirmación 5.

Consideremos una sucesión $(\lambda_n) \subseteq \mathbb{R}^+$ tal que $0 < \lambda_n < \hat{\lambda}$ para todo $n \in \mathbb{N}$ y $\lim\limits_{\lambda\to 0^+}\lambda_n = 0$. Sean las sucesiones $(\mu_n) \subseteq \mathbb{R}^+$ y $(x_n) \subseteq X$ definidas por

$$\mu_n = \frac{1}{\lambda_n} \quad y \quad x_n = \bar{x} + \lambda_n x + r(\lambda_n)$$

para cada $n \in \mathbb{N}$, donde $r(\lambda_n)$ es el límite de la sucesión (r_k) para el valor λ_n. Claramente $\mu_n > 0$ y por (4.36) $h(x_n) = h(\bar{x}+\lambda_n x + r(\lambda_n)) = 0_Z$. Así $x_n \in S$ para todo $n \in \mathbb{N}$. Por la afirmación 5 y la convergencia de (λ_n) tenemos

$$\begin{aligned}
\lim_{n\to\infty} x_n &= \lim_{n\to\infty} \bar{x} + \lambda_n x + r(\lambda_n) \\
&= \lim_{n\to\infty} \bar{x} + \lambda_n \left\{ x + \frac{r(\lambda_n)}{\lambda_n} \right\} \\
&= \bar{x}
\end{aligned}$$

y además

$$\begin{aligned}
\lim_{n\to\infty} \mu_n(x_n - \bar{x}) &= \lim_{n\to\infty} \frac{1}{\mu_n}(\bar{x} + \lambda_n x + r(\lambda_n) - \bar{x}) \\
&= \lim_{n\to\infty} x + \frac{r(\lambda_n)}{\lambda_n} = x
\end{aligned}$$

Por lo tanto $x \in T(S,\bar{x})$, es decir $\{x \in X / h'(\bar{x})x = 0_z\} \subseteq T(S,\bar{x})$, quedando demostrado el Teorema.

$\square$

Teorema 4.15. Sean $(X, \|\cdot\|_X)$ y $(Z, \|\cdot\|_Z)$ dos e.v.n y $h : X \longrightarrow Z$ una aplicación. Consideremos el conjunto $S = \{x \in X / h(x) = 0_Z\}$ y $\bar{x} \in S$. Si h es Fréchet diferenciable en $\bar{x}$ entonces

$$T(S, \bar{x}) \subseteq W$$

donde $W = \{x \in X / h'(\bar{x})x = 0_Z\}$.

Demostración.

Sea $y \in T(S, \bar{x})$. Si $y = 0_X$ entonces $y \in W$ pues $h'(\bar{x})$ es una aplicación lineal. Supongamos que $y \neq 0_X$, por la Definición 4.1 existen dos sucesiones $(x_n) \subseteq S$ y $(\lambda_n) \subseteq \mathbb{R}^+$ tales que

$\bar{x} = \lim\limits_{n\to\infty} x_n$ e $y = \lim\limits_{n\to\infty} \lambda_n(x_n - \bar{x})$.

Como h es Fréchet diferenciable en $\bar{x}$ tenemos

$$\begin{aligned}
h'(\bar{x})y &= h'(\bar{x}) \left\{ \lim_{n\to\infty} \lambda_n(x_n - \bar{x}) \right\} \\
&= \lim_{n\to\infty} h'(\bar{x})\{\lambda_n(x_n - \bar{x})\} \\
&= \lim_{n\to\infty} \lambda_n h'(\bar{x})(x_n - \bar{x}) \\
&= \lim_{n\to\infty} \lambda_n\{h(x_n) - h(\bar{x}) - (h(x_n) - h(\bar{x}) - h'(\bar{x})(x_n - \bar{x}))\} \\
&= -\lim_{n\to\infty} \frac{\lambda_n \|x_n - \bar{x}\|_X \{h(x_n) - h(\bar{x}) - h'(\bar{x})(x_n - \bar{x})\}}{\|x_n - \bar{x}\|_X} \quad (x_n, \bar{x} \in S) \\
&= -\lim_{n\to\infty} \|y_n\|_X \frac{\{h(x_n) - h(\bar{x}) - h'(\bar{x})(x_n - \bar{x})\}}{\|x_n - \bar{x}\|_X} = -\|y\|_X \{0_Z\} = 0_Z
\end{aligned}$$

con lo cual $h'(\bar{x})y = 0_Z$. En consecuencia $y \in W$, demostrando el resultado.

□

Observación.

Del teorema de Lyusternik y el teorema anterior bajo sus respectivas hipótesis, tenemos la igualdad

$$T(S, \bar{x}) = \{x \in X / h'(\bar{x})x = 0_Z\}$$

Capítulo 5

Multiplicadores de Lagrange en espacios de Banach

En este capítulo desarrollaremos una extensión de la teoría de optimización sujeta a restricciones o condiciones calificadas, dentro de un espacio de Banach. En principio daremos las siguientes definiciones referidas a la generalización de las condiciones calificadas para problemas de optimización restringida en espacios normados.

5.1. Conos duales y de orden en espacios normados

Definición 5.1. Sea X un espacio vectorial real.

(i) Decimos que R es una *relación binaria* en X si $R \subseteq X \times X$. Si $(x, y) \in R$ escribiremos xRy. Usaremos la notación $\leq$ para indicar una relación en X.

(ii) Sea $\leq$ una relación binaria en X. Diremos que $\leq$ es un *orden parcial* en X si:

- $a)$ $x \leq x$ para todo $x \in X$.
- $b)$ $x \leq y$ e $y \leq z$ entonces $x \leq z$ para cada $x, y, z \in X$.
- $c)$ $x \leq y$ y $w \leq z$ entonces $x + w \leq y + z$ para todo $x, y, z, w \in X$.
- $d)$ $x \leq y$ y $\alpha \in \mathbb{R}^+$ entonces $\alpha x \leq \alpha y$ para todo $x, y \in X$.

(iii) Sea $\leq$ un orden parcial. Decimos que $\leq$ es *antisimétrica* si para todo $x, y \in X$ tales que $x \leq y$ e $y \leq x$ se tiene que $x = y$.

(iv) Un espacio vectorial X equipado con un orden parcial es llamado un *espacio vectorial parcialmente ordenado*.

Ejemplos.

- En $\mathbb{R}^n$ definimos la relación

$$\leq := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n / x_i \leq y_i \text{ para cada } i = 1, 2, \cdots n\}$$

entonces $\mathbb{R}^n$ es parcialmente ordenado por $\leq$.

- Sea el espacio vectorial $X = \{f : [a, b] \longrightarrow \mathbb{R}/f$ es continua en $[a, b]\}$. Definimos el orden parcial $\leq$ en X como

$$\leq := \{(x, y) \in X \times X/x(t) \leq y(t) \text{ para todo } t \in [a, b]\}$$

se cumple que X es un espacio vectorial parcialmente ordenado por $\leq$.

Teorema 5.2. Sea X un espacio vectorial.

(i) Si $\leq$ es un orden parcial en X entonces el conjunto $C = \{x \in X/0_X \leq x\}$ es un cono convexo. Además si $\leq$ es antisimétrica entonces C es cono puntual.

(ii) Si C un cono convexo en X entonces $\leq := \{(x, y) \in X \times X/y - x \in C\}$ es un orden parcial en X. Ademas si C es puntual entonces el orden parcial es antisimétrico.

Demostración.

(i) Por hipótesis $\leq$ es reflexivo, así $0_X \leq 0_X$ de donde $0_X \in C$ en consecuencia C es no vacío.

Sean $x, y \in C$ entonces $0_X \leq x$ y $0_X \leq y$ por tanto $0_X \leq x + y$. Así $x + y \in C$.

Por otra parte si $\alpha > 0$ y $0_X \leq x$ tenemos que $0_X\alpha \leq \alpha x$ en consecuencia $0_X \leq \alpha x$, por lo tanto $\alpha x \in C$. Por consiguiente, C es un cono convexo.

Supongamos que $\leq$ es antisimétrica y además $x, -x \in C$ es decir $0_X \leq x$ y $0_X \leq -x$. De esta última relación y por la reflexividad de $\leq$ tenemos $x \leq 0_X$. Por lo tanto $x = 0_X$. En consecuencia por la Definición 1.28 (ii) tenemos que C es cono puntual.

(ii) Sea C un cono convexo. Como $0_X \in C$ entonces $x - x \in C$ por lo tanto $x \leq x$. Sean $x \leq y$ e $y \leq z$ entonces $y - x \in C$ y $z - y \in C$ como C es cono convexo $z - x \in C$, con lo cual $x \leq z$.

Por otra parte si $x \leq y$ y $w \leq z$ entonces $y - x \in C$ y $z - w \in C$ y por ser C cono convexo tenemos que $(y + z) - (x + w) = y - z + z - w \in C$, por lo tanto $x + w \leq y + z$.

Además si $x \leq y$ y $\alpha \geq 0$ tenemos que $\alpha(y - x) \in C$, en consecuencia $\alpha x \leq \alpha y$. Mostrando de esta forma que $\leq$ es un orden parcial.

Finalmente, si C es puntual y $y \leq x$ y $x \leq y$ tenemos que $y - x \in X$ y $-(y - x) = x - y \in C$ lo cual implica que $x - y = 0$,es decir $x = y$ con lo cual $\leq$ es antisimétrica.

□

El teorema 5.2 da origen a las siguientes definiciones.

Definición 5.3. Un cono convexo C que define un orden parcial en un espacio vectorial real X es llamado un *cono de orden* (o cono positivo) en X.

Ejemplos.

- En $\mathbb{R}^n$ el cono $C = \{x \in \mathbb{R}^n / x_i \geq 0 \, \text{para todo} \; i\}$ es un cono de orden.
- En el espacio de funciones continuas $C = \{x \in C([a,b]) / x(t) \geq 0 \; \forall t \in [a,b]\}$ es un cono de orden.

Definición 5.4. Sea X un espacio vectorial real con un cono de orden C. Definimos el cono

$$C' = \{l \in X' / l(x) \geq 0 \, \text{para todo} \; x \in C\}$$

denominado *cono dual algebraico* de C en X'.

Observación.

Si $(X, \| \, \|)$ es un espacio normado, definimos el *cono dual topológico* como el conjunto

$$C^{\star} = \{l \in X^{\star} / l(x) \geq 0 \, \text{para todo} \; x \in C\}$$

5.2. Formulación del problema de optimización con restricciones en un espacio de Banach

En esta sección se formulan las hipótesis y se plantea el problema de optimización condicionada de funcionales definidas en un espacio de Banach.

Hipótesis Generales (HG)

Sean $(X, \| \, \|_X)$ y $(Z, \| \, \|_Z)$ dos espacios de Banach reales. Consideremos el espacio real normado $(Y, \| \, \|_Y)$ parcialmente ordenado bajo el cono de orden $C \subseteq Y$ que satisface $\overset{\circ}{C} \neq \emptyset$.

Sea $\hat{S}$ un subconjunto convexo de X con interior no vacío. Consideremos $f : X \longrightarrow \mathbb{R}$ una funcional, las aplicaciones $g : X \longrightarrow Y$ y $h : X \longrightarrow Z$ y el conjunto

$$S = \{x \in \hat{S} / g(x) \in -C \; \wedge \; h(x) = 0_Z\}$$

que será denominado *conjunto de restricciones o conjunto de condiciones calificadas.*

Bajo estos supuestos e hipótesis generales, consideremos el problema de optimización con restricciones generalizadas $\min\limits_{x \in S} f(x)$ en el espacio de Banach X.

5.3. Condiciones necesarias de optimalidad

Bajo las hipótesis generales (HG) presentamos las condiciones necesarias para los puntos minimizantes de f en S. Estas condiciones se denominan regla de multiplicadores de Lagrange generalizadas. En principio, usando los resultados del Teorema 4.14 tenemos el siguiente lema.

Lema 5.5. Supongamos válidas las hipótesis generales (HG). Consideremos $\bar{x} = \arg\min\limits_{x \in S} f(x)$. Supongamos que:

(i) f y g son Fréchet diferenciales en $\bar{x}$.

(ii) h es Fréchet diferenciable en una vecindad de $\bar{x}$.

(iii) $h'(\cdot)$ es continua en $\bar{x}$.

(iv) $h'(\bar{x})$ es una aplicación sobreyectiva.

Entonces no existe $x \in \mathring{\hat{S}}$ tal que $f'(\bar{x})(x-\bar{x}) < 0$; $g(\bar{x})+g'(\bar{x})(x-\bar{x}) \in -\mathring{C}$ y $h'(\bar{x})(x-\bar{x}) = 0_Z$.

Demostración.

Sea $\bar{x} = \arg\min\limits_{x \in S} f(x)$ y tomemos $x \in \mathring{\hat{S}}$ tal que:

$$x \neq \bar{x},\ g(\bar{x}) + g'(\bar{x})(x - \bar{x}) \in -\mathring{C}\ y\ h'(\bar{x})(x - \bar{x}) = 0_Z.$$

Veamos que $f'(\bar{x})(x - \bar{x}) \geq 0$. En efecto, consideremos el conjunto $\tilde{S} = \{x \in X / h(x) = 0_Z\}$. Como $h'(\bar{x})(x - \bar{x}) = 0_Z$, por el teorema de Lyusternik tenemos que $x - \bar{x} \in T(\hat{S}, \bar{x})$. Por la Definición 4.1 existen dos sucesiones $(x_n) \subseteq \hat{S}$ y $(\lambda_n) \subseteq \mathbb{R}^+$ tales que $\bar{x} = \lim\limits_{n\to\infty} x_n$ y $x - \bar{x} = \lim\limits_{n\to\infty} \lambda_n(x_n - \bar{x})$. Sea la sucesión $(y_n) \subseteq X$ definida por $y_n = \lambda_n(x_n - \bar{x})$. Claramente por la convergencia anterior, tenemos que $x = \lim\limits_{n\to\infty} x_n + y_n$. Dado que $x \in \mathring{\hat{S}}$, por el Lema 1.5-(i), existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$ se cumple que $\bar{x} + y_n \in \hat{S}$.

Como $x - \bar{x} \neq 0_X$, la sucesión $\dfrac{1}{\lambda_n} \to 0$ si $n \to \infty$, en consecuencia existe $n_1 \in \mathbb{N}$ tal que $0 < \dfrac{1}{\lambda_n} < 1$ para todo $n \geq n_1$. Sea $n_2 = \max\{n_0, n_1\}$, si $n \geq n_2$ por la convexidad de $\hat{S}$ y la definición de y_n tenemos

$$\begin{aligned}
x_n &= \bar{x} + \frac{1}{\lambda_n} y_n \\
&= \bar{x} + \frac{1}{\lambda_n}(y_n + \bar{x} - \bar{x}) \\
&= (1 - \frac{1}{\lambda_n})\bar{x} + \frac{1}{\lambda_n}(y_n + \bar{x}) \in \hat{S}
\end{aligned}$$

Por lo tanto, si $n \geq n_2$ tenemos que $x_n \in \hat{S} \cap \tilde{S}$. Por otra parte como $x_n - \bar{x} = \dfrac{1}{\lambda_n} y_n$, obtenemos por la linealidad de $g'(\bar{x})$

$$\begin{aligned}
g(x_n) &= g(x_n) - g'(\bar{x})(x_n - \bar{x}) + \frac{1}{\lambda_n} g'(\bar{x}) y_n \\
&= g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x}) + \frac{1}{\lambda_n} g'(\bar{x}) y_n - \frac{1}{\lambda_n} g'(\bar{x})(x - \bar{x}) + \frac{1}{\lambda_n} g(\bar{x}) \\
&\quad + \frac{1}{\lambda_n} g'(\bar{x})(x - \bar{x}) + g(\bar{x}) - \frac{1}{\lambda_n} g(\bar{x}).
\end{aligned} \tag{5.1}$$

para cada $n \in \mathbb{N}$.

De esta última expresión, reordenando los términos convenientemente, logramos la identidad

$$\begin{aligned} g(x_n) = \frac{1}{\lambda_n}\{\lambda_n(g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x})) + g'(\bar{x})(y_n - (x - \bar{x})) \\ + g(\bar{x}) + g'(\bar{x})(x - \bar{x})\} + (1 - \frac{1}{\lambda_n})g(\bar{x}), \end{aligned} \tag{5.2}$$

para todo $n \in \mathbb{N}$.

Como $\lambda_n = \dfrac{\|y_n\|}{\|x_n - \bar{x}\|}$ y por la hipótesis (i) sobre g obtenemos

$$\begin{aligned} \lim_{n\to\infty} \lambda_n(g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x})) &= \lim_{n\to\infty} \frac{\|y_n\|\,(g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x}))}{\|x_n - \bar{x}\|} \\ &= 0 \end{aligned} \tag{5.3}$$

puesto que g es Fréchet diferenciable en $\bar{x}$. Como $g'(\bar{x}) \in L(X, Y)$ y $y_n \longrightarrow x - \bar{x}$ si $n \to \infty$, tenemos que

$$\lim_{n\to\infty} g'(\bar{x})(y_n - (x - \bar{x})) = 0 \tag{5.4}$$

De las convergencias (5.3) y (5.4) logramos

$$\lim_{n\to\infty} \lambda_n(g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x})) + g'(\bar{x})(y_n - (x - \bar{x})) = 0 \tag{5.5}$$

De (5.5) tenemos que la sucesión

$$w_n = \lambda_n(g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x})) + g'(\bar{x})(y_n - (x - \bar{x})) + g(\bar{x}) + g'(\bar{x})(x - \bar{x})$$

converge a $g(\bar{x}) + g'(\bar{x})(x - \bar{x})$ cuando $n \to \infty$.

Por hipótesis $g(\bar{x}) + g'(\bar{x})(x - \bar{x}) \in -\mathring{C}$, nuevamente por el Lema 1.5 existe $n_3 \in \mathbb{N}$ tal que $w_n \in -C$ para todo $n \geq n_3$. Como $g(\bar{x}) \in -C$ si $n \geq \text{máx}\{n_3, n_2\}$ obtenemos por la convexidad de $-C$ y la identidad (5.2) que

$$\begin{aligned} g(x_n) = \frac{1}{\lambda_n}\{\lambda_n(g(x_n) - g(\bar{x}) - g'(\bar{x})(x_n - \bar{x})) + g'(\bar{x})(y_n - (x - \bar{x})) \\ + g(\bar{x}) + g'(\bar{x})(x - \bar{x})\} + (1 - \frac{1}{\lambda_n})g(\bar{x}) \in -C \end{aligned}$$

Así, si $n \geq \text{máx}\{n_3, n_2\}$ tenemos que $x_n \in \{x \in \hat{S} / g(x) \in -C \wedge h(x) = 0_Z\} = S$. Por lo tanto, existen $(x_n) \subseteq S$ y $(\lambda_n) \subseteq \mathbb{R}^+$ tales que

$$\bar{x} = \lim_{n\to\infty} x_n \quad \text{y} \quad x - \bar{x} = \lim_{n\to\infty} \lambda_n(x_n - \bar{x})$$

En consecuencia $x - \bar{x} \in T(S, \bar{x})$. Luego por el Teorema 4.8 tenemos $f'(\bar{x})(x - \bar{x}) \geq 0$, mostrando así el Lema.

□

Teorema 5.6. (Teorema de multiplicadores de Lagrange en espacios de Banach)

Supongamos válidas las hipótesis generales (HG). Sea $\bar{x} = \arg\min_{x\in S} f(x)$. Si se cumplen:

(i) f y g son Fréchet diferenciables en $\bar{x}$.

(ii) h es Fréchet diferenciable en una vecindad de $\bar{x}$.

(iii) $h'(\cdot)$ es continua en $\bar{x}$.

(iv) $h'(\bar{x})X$ es cerrado en Z.

Entonces existen $\mu \geq 0$, $l_1 \in C^\star$, $l_2 \in Z^\star$ con $(\mu, l_1, l_2) \neq (0, 0_{Y'}, 0_{Z'})$ tales que

$$(\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x}))(x - \bar{x}) \geq 0$$

para todo $x \in \hat{S}$ y $l_1 \circ g(\bar{x}) = 0$, siendo $C^\star$ el cono dual de C. Además, si

$$\begin{pmatrix} g'(\bar{x}) \\ h'(\bar{x}) \end{pmatrix} cono\left(\hat{S} - \{\bar{x}\}\right) + cono\begin{pmatrix} C + \{g(\bar{x})\} \\ \{0_Z\} \end{pmatrix} = Y \times Z$$

entonces $\mu > 0$.

Demostración.

En principio veamos la primera afirmación. Consideremos dos casos:

Caso 1: $h'(\bar{x})$ no es sobreyectiva.

En efecto, por la suposición del caso, existe $\bar{z} \in Z$ tal que $\bar{z} \notin h'(\bar{x})X$ y siendo este un conjunto cerrado y convexo, por el Teorema 1.48 existe $l_2 \in Z^\star \setminus \{0_Z\}$ tal que

$$l_2(\bar{z}) < \inf_{z\in h'(\bar{x})X} l_2(z)$$

Sea $z \in h'(\bar{x})X$ entonces por la linealidad de l_2 se obtiene

$$\begin{aligned} l_2(\bar{z}) &< l_2(\lambda z) \\ &= \lambda l_2(z) \end{aligned}$$

para todo $\lambda \in \mathbb{R}$. Tomando $\lambda = 0$ en la desigualdad anterior, tenemos $l_2(\bar{z}) < 0$ y $\dfrac{l_2(\bar{z})}{\lambda} < l_2(z)$. De esta última desigualdad tomando los límites cuando $\lambda \to +\infty$ y $\lambda \to -\infty$ resulta que $l_2(z) \leq 0$ y $l_2(z) \geq 0$, logrando de esta manera que $l_2(z) = 0$ para cada $z \in h'(\bar{x})X$. De este modo tenemos $l_2(h'(\bar{x})x) = 0$ para todo $x \in X$, en consecuencia $l_2 \circ h'(\bar{x}) = 0_{X^\star}$. Eligiendo $\mu = 0$, $l_1 = 0_{Y^\star}$ y $l_2 \in Z^\star$ tenemos que $(\mu, l_1, l_2) \neq (0, 0_{Y^\star}, 0_{Z^\star})$ y

$$
\begin{aligned}
(\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x}))(x - \bar{x}) &= l_2 \circ h'(\bar{x})(x - \bar{x}) \\
&= 0
\end{aligned}
$$

para todo $x \in \hat{S}$. Además $l_1(g(\bar{x})) = 0$.

Caso 2: $h'(\bar{x})$ es sobreyectiva.

En este caso tenemos que $h'(\bar{x})X = X$. En el espacio vectorial producto $\mathbb{R} \times Y \times Z$ consideremos el conjunto

$$
M = \{(f'(\bar{x})(x-\bar{x}) + \alpha, g(\bar{x}) + g'(\bar{x})(x-\bar{x}) + y, h'(\bar{x})(x-\bar{x})) \in \mathbb{R} \times Y \times Z / x \in \mathring{\hat{S}}, \alpha > 0, y \in \mathring{C}\}
$$

Por hipótesis f, g y h son diferenciables en $\bar{x}$ y además $\mathring{\hat{S}} \subseteq X$, $\mathring{C} \subseteq Y$ son conjuntos no vacíos. En consecuencia M es no vacío. Definimos $W = \mathring{\hat{S}} - \{\bar{x}\}$ y considerando las aplicaciones lineales continuas $f'(\bar{x})$, $g'(\bar{x})$ y $h'(\bar{x})$ podemos escribir

$$
M = R_1 \times Y_1 \times Z_1
$$

donde $R_1 = f'(\bar{x})(W) + \mathbb{R}^+$, $Y_1 = g(\bar{x}) + g'(\bar{x})(W) + \mathring{C}$ y $Z_1 = h'(\bar{x})(W)$. Afirmamos que M es abierto y convexo en $\mathbb{R} \times Y \times Z$. En efecto, por la convexidad de $\mathring{\hat{S}}$ y el conjunto unitario $\{\bar{x}\}$ por el Lema 1.25-(ii) (tomando $\alpha = 1$ y $\beta = -1$), tenemos que W es convexo. Así por la convexidad de $\mathbb{R}^+$, $\mathring{C}$ y W, del Lema 1.25 obtenemos que R_1, Y_1 y Z_1 son convexos en $\mathbb{R}$, Y y Z respectivamente. En consecuencia $M = R_1 \times Y_1 \times Z_1$ es convexo en el espacio producto $\mathbb{R} \times Y \times Z$.

Veamos que M es abierto. Como $\mathbb{R}^+$ y $\mathring{C}$ son abiertos, tenemos que R_1 y Y_1 son abiertos en $\mathbb{R}$ e Y respectivamente. Por otra parte como W es abierto en X y $h'(\bar{x})$ es continua y sobreyectiva, por el teorema de la aplicación abierta, el conjunto $Z_1 = h'(\bar{x})(W)$ es abierto en Z, así concluimos que M es abierto en el espacio producto $\mathbb{R} \times Y \times Z$, mostrando así nuestra afirmación.

Afirmamos que $(0, 0_Y, 0_Z)$ no pertenece a M. En efecto, supongamos que $(0, 0_Y, 0_Z)$ pertenezca a M (hipótesis auxiliar), entonces existen $x_1 \in \hat{S}$, $\alpha > 0$ e $y \in \mathring{C}$ tales que

$$
(f'(\bar{x})(x_1 - \bar{x}) + \alpha, g(\bar{x}) + g'(\bar{x})(x_1 - \bar{x}) + y, h'(\bar{x})(x_1 - \bar{x})) = (0, 0_Y, 0_Z) \tag{5.6}
$$

De la identidad (5.6) conseguimos

$$f'(\bar{x})(x_1 - \bar{x}) + \alpha = 0 \tag{5.7}$$
$$g(\bar{x}) + g'(\bar{x})(x_1 - \bar{x}) + y = 0_Y \tag{5.8}$$
$$h'(\bar{x})(x_1 - \bar{x}) = 0_Z \tag{5.9}$$

De las identidades (5.7) y (5.8) obtenemos las desigualdades $f'(\bar{x})(x_1 - \bar{x}) = -\alpha < 0$ y $g(\bar{x}) + g'(\bar{x})(x_1 - \bar{x}) = -y \in -\mathring{C}$, esto sumado a la identidad (5.9) contradice el Lema 5.5. En consecuencia $M \cap \{(0, 0_Y, 0_Z)\} = \emptyset$.

Por el teorema de separación de Eidelheit (Teorema 1.47) existen $\mu \in \mathbb{R}$, $l_1 \in Y^\star, l_2 \in Z^\star$ y $\gamma \in \mathbb{R}$ tales que $(\mu, l_1, l_2) \neq (0, 0_{Y^\star}, 0_{Z^\star})$ y

$$\begin{aligned} \mu(f'(\bar{x})(x - \bar{x}) + \alpha) + l_1(g(\bar{x}) + g'(\bar{x})(x - \bar{x}) + y) + l_2(h'(\bar{x})(x - \bar{x})) &> \gamma \\ &\geq \mu(0) + l_1(0_{Y^\star}) + l_2(0_{Z^\star}) = 0 \end{aligned}$$

para todo $x \in \mathring{\hat{S}}$, $\alpha > 0$ e $y \in \mathring{C}$. Por el Lema 1.26 aplicado a los conjuntos $\mathring{\hat{S}}, \mathring{C}$ y la continuidad de los operadores lineales tenemos que

$$\begin{aligned} \mu(f'(\bar{x})(x - \bar{x}) + \alpha) + l_1(g(\bar{x}) + g'(\bar{x})(x - \bar{x}) + y) + l_2(h'(\bar{x})(x - \bar{x})) &\geq \gamma \\ &\geq 0 \end{aligned} \tag{5.10}$$

para todo $x \in \hat{S}$, $\alpha \geq 0$ e $y \in C$.

Tomando $x = \bar{x}$, obtenemos de (5.10)

$$\mu\alpha + l_1(g(\bar{x}) + y) \geq 0 \tag{5.11}$$

para todo $\alpha \geq 0$ e $y \in C$. Haciendo $\alpha = 1$ e $y = -g(\bar{x}) \in C$ $(\bar{x} \in S)$ en (5.11) tenemos que $\mu \geq 0$. Asimismo si $\alpha = 0$ en (5.11) tenemos la desigualdad $l_1(g(\bar{x}) + y) \geq 0$, con lo cual

$$l_1(g(\bar{x})) \geq -l_1(y) \tag{5.12}$$

para todo $y \in C$. Afirmamos que $l_1(y) \geq 0$ para todo $y \in C$. En efecto, supongamos que existe $y_1 \in C$ tal que $l_1(y_1) < 0$. Por la desigualdad anterior tenemos que $l_1(g(\bar{x})) \geq -l_1(y_1) > 0$. Por hipótesis, C es cono, en consecuencia $\lambda y_1 \in C$ para todo $\lambda > 0$ resultando la desigualdad

$$l_1(g(\bar{x})) \geq -\lambda l_1(y_1) \tag{5.13}$$

para todo $\lambda > 0$. Tomando $\lambda \to +\infty$ en (5.13) tenemos que $l_1(g(\bar{x})) \geq +\infty$ contradiciendo la acotación de l_1. Por lo tanto $l_1(y) \geq 0$ para todo $y \in C$.

Por la definición de cono dual topológico tenemos que $l_1 \in C^\star$. Tomando $y = 0$ en (5.12) se tiene que $l_1(g(\bar{x})) \geq 0$. Puesto que $\bar{x} \in S$ y por la desigualdad anterior resulta que $g(\bar{x}) \in -C$, es decir $-g(\bar{x}) \in C$ y por la afirmación anterior $l_1(-g(\bar{x})) \geq 0$ obteniendo $l_1(g(\bar{x})) \leq 0$. Por lo tanto $l_1(g(\bar{x})) = 0$.

Haciendo $\alpha = 0$ e $y = -g(\bar{x})$ en (5.10) resulta la desigualdad

$$\mu(f'(\bar{x})(x - \bar{x})) + l_1(g'(\bar{x}))(x - \bar{x})) + l_2(h'(\bar{x})(x - \bar{x})) \geq 0$$

para todo $x \in \hat{S}$. En consecuencia

$$\left(\mu f'(\bar{x}) + l_1(g'(\bar{x})) + l_2(h'(\bar{x}))\right)(x - \bar{x})) \geq 0 \tag{5.14}$$

para todo $x \in \hat{S}$. Finalmente, supongamos que

$$\begin{pmatrix} g'(\bar{x}) \\ h'(\bar{x}) \end{pmatrix} cono\left(\hat{S} - \{\bar{x}\}\right) + cono\begin{pmatrix} C + \{g(\bar{x})\} \\ \{0_Z\} \end{pmatrix} = Y \times Z$$

entonces dados $y \in Y$ y $z \in Z$ existen $\alpha, \beta \geq 0$, $x \in \hat{S}$ y $c \in C$ tales que

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} g'(\bar{x})(\alpha(x - \bar{x})) + \beta(c + g(\bar{x})) \\ h'(\bar{x})(\alpha(x - \bar{x})) \end{pmatrix}$$

es decir

$$\begin{aligned} y &= g'(\bar{x})(\alpha(x - \bar{x})) + \beta(c + g(\bar{x})) \\ z &= h'(\bar{x})((\alpha x - \bar{x})) \end{aligned}$$

Supongamos que $\mu = 0$ (hipótesis auxiliar). De (5.14) tenemos $(l_1(g'(\bar{x})) + l_2(h'(\bar{x})))\,(x - \bar{x})) \geq 0$, para todo $x \in \hat{S}$. Como $l_1(c) \geq 0$ y $l_1(g(\bar{x})) = 0$ tenemos

$$\begin{aligned} l_1(y) + l_2(z) &= l_1(g'(\bar{x})(\alpha(x - \bar{x})) + \beta(c + g(\bar{x})) + l_2(h'(\bar{x})\alpha(x - \bar{x})) \\ &= l_1 \circ g'(\bar{x})(\alpha(x - \bar{x})) + \beta l_1(c) + \beta l_1(g(\bar{x})) + l_2 \circ h'(\bar{x})(\alpha(x - \bar{x})) \\ &= \alpha\left\{l_1 \circ g'(\bar{x})(x - \bar{x}) + l_2 \circ h'(\bar{x})(x - \bar{x})\right\} + \beta l_1(c) \geq 0 \end{aligned}$$

Obteniéndose de esta manera $l_1(y) + l_2(z) \geq 0$ para cada $y \in Y$ y $z \in Z$. Tomando $z = 0$ tenemos que $l_1(y) \geq 0$ para todo $y \in Y$. Haciendo el mismo cálculo con $-y \in Y$ tenemos que $l_1(y) \leq 0$ para todo $y \in Y$, en consecuencia $l_1(y) = 0$ para todo $y \in Y$ es decir $l_1 = 0_{Y^\star}$. Análogamente, tomando $y = 0$ en dicha desigualdad y efectuando los mismos pasos anteriores, concluimos que $l_2 = 0_{Z^\star}$. Por lo tanto $(\mu, l_1, l_2) = (0, 0_{Y^\star}, 0_{Z^\star})$ lo que contradice la primera parte de la demostración. De esta forma $\mu > 0$.

$\square$

Observaciones.

- La condición

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} g'(\bar{x})(\alpha(x-\bar{x})) + \beta(c + g(\bar{x})) \\ h'(\bar{x})(\alpha(x-\bar{x})) \end{pmatrix} = Y \times Z$$

  es llamada condición de regularidad o de Kurcyuz-Robinson-Zowe.

- Observemos que la condición de regularidad, es solo una condición sobre el conjunto S (conjunto de restricciones) y es independiente de la funcional objetivo f. Si $\mu = 0$ tenemos que

$$\left(l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x})\right)(x - \bar{x}) \geq 0$$

  para todo $x \in \hat{S}$. En este caso, la regla de los multiplicadores de Lagrange no contiene información sobre la funcional objetivo, caso que no es deseable.

En general, cuando $\mu > 0$, la condición necesaria de optimalidad puede ser expresada como

$$\left(f'(\bar{x}) + \frac{1}{\mu}(l_1 \circ g'(\bar{x})) + \frac{1}{\mu}(l_2 \circ h'(\bar{x}))\right)(x - \bar{x}) \geq 0$$

para todo $x \in \hat{S}$. Además como $l_1(g(\bar{x})) = 0$ tenemos que $\dfrac{l_1}{\mu}(g(\bar{x})) = 0$. Denotando por $u = \dfrac{l_1}{\mu} \in C^\star$ y $v = \dfrac{l_2}{\mu} \in Z^\star$, obtenemos de la desigualdad anterior

$$\left(f'(\bar{x}) + u \circ g'(\bar{x}) + v \circ h'(\bar{x})\right)(x - \bar{x}) \geq 0 \tag{5.15}$$

para todo $x \in \hat{S}$ y $u(g(\bar{x})) = 0$. Definimos el funcional $L = f + u \circ g + v \circ h$, llamado el *funcional de Lagrange generalizado*. Como g y h son Fréchet diferenciables en $\bar{x}$ tenemos que

$$L'(\bar{x}) = f'(\bar{x}) + u \circ g'(\bar{x}) + v \circ h'(\bar{x})$$

En consecuencia, podemos expresar (5.15) como

$$L'(\bar{x})(x - \bar{x}) \geq 0$$

para todo $x \in \hat{S}$. Cuando $\hat{S} = X$ el teorema de los multiplicadores de Lagrange queda resumido por el siguiente corolario.

Corolario 5.7. Supongamos válidas las hipótesis (H.G). Si $\hat{S} = X$, $\bar{x} = \arg\min_{x \in S} f(x)$ y se satisfacen:

(*i*) f y g son Fréchet diferenciables.

(*ii*) h es Fréchet diferenciable en una vecindad de $\bar{x}$.

(*iii*) $h'(\cdot)$ es continua en X y $h'(\bar{x})X$ es un conjunto cerrado.

Entonces, existen $\mu \geq 0$, $l_1 \in C^\star$, $l_2 \in Z^\star$ con $(\mu, l_1, l_2) \neq (0, 0_{Y^\star}, 0_{Z^\star})$ tales que

$$\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x}) = 0_{X^\star} \quad y \quad l_1(g(\bar{x})) = 0$$

Además, si las condiciones de regularidad de Kurcyuz-Robinson-Zowe son satisfechas, entonces $\mu > 0$.

Demostración.

Por el Teorema 5.6, tomando $\hat{S} = X$ tenemos que existen $\mu \geq 0$, $l_1 \in C^\star$, $l_2 \in Z^\star$ con $(\mu, l_1, l_2) \neq (0, 0_{Y^\star}, 0_{Z^\star})$ tales que

$$\left(\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x})\right)(x - \bar{x})) \geq 0 \tag{5.16}$$

para todo $x \in X$. Como las aplicaciones son lineales y $\bar{x} \in X$, tomando el vector $2\bar{x} - x$ $(x \in X)$, obtenemos de (5.16)

$$\left(\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x})\right)(x - \bar{x})) \leq 0 \tag{5.17}$$

para todo $x \in X$. De (5.16) y (5.17) conseguimos

$$\left(\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x})\right)(x - \bar{x}) = 0$$

para cada $x \in X$. Por lo tanto $\mu f'(\bar{x}) + l_1 \circ g'(\bar{x}) + l_2 \circ h'(\bar{x}) = 0_{X^\star}$. Si las condiciones de regularidad son satisfechas, tenemos que $\mu > 0$, mostrando el resultado.

$\square$

Bibliografía

[1] Barbu V., Precupanu T., *Convexity and Optimization in Banach Spaces.* Springer Science+Bussines Media B.V., Dordrecht, (2012).

[2] Berkovitz L., *Convexity and Optimization in* $\mathbb{R}^n$. John Wiley & Sons, New York, (2002).

[3] Botelho F., *Topics on Functional Analysis, Calculus of Variations and Duality.* Academic Publications Ltd, Londres, (2011).

[4] Eidelman Y., Milman V., Tsolomitis A., *Functional Analysis An Introduction.* AMS, Rhode Island, (2004).

[5] Guignard M., Generalized Kuhn-Tucker conditions for Mathematical Programming Problems in Banach Spaces. *SIAM J. Control*, 7(2): 232-241, (1969).

[6] Holmes R., *Geometric Functional Analysis and its Applications.* Springer-Verlag, New York, (1975).

[7] Jahn J., Existence Theorems in Vector Optimization. *JOTA*, 50(3):397-406, (1986).

[8] Jahn J., *Introduction to the Theory of Nonlinear Optimization.* Springer-Verlag, Berlín, (2007).

[9] Jost J. *Postmodern Analysis.* Springer-Verlag, Berlín, (2005).

[10] Kazmi K., Some remarks on Vector Optimization Problems. *JOTA*, 96(1): 133-138, (1998).

[11] Kreyszig E., *Introductory Functional Analysis with Applications.* John Wiley & Sons, New York, (1989).

[12] Lima E., *Curso de Analisis Matematico, Vol 1.* Ed. Edunsa, Barcelona, (1991).

[13] Luenberger D., *Optimization by Vector Spaces Methods.* John Wiley & Sons, New York, (1998).

[14] Ritter K., Optimization Theory in Linear Spaces. Part III. Mathematical Programming in Partially Ordered. *Math. Ann.*, 184: 133-154, (1970).

[15] Singer I., *Abstract Convex Analysis.* John Wiley & Sons, New York, (1997).

[16] Varaiya P., Nonlinear Programming in Banach Spaces. *SIAM J. Appl. Math.*, 15(2): 284-293, (1967).

[17] Werner J., *Optimization Theory and Applications*. Vieweg, Braunschweig, (1984).

[18] Wheatherwax R. General Lagrange Multiplier Theorems. *JOTA* , 14(1): 51 -72, (1974).

[19] Yu G., Liu S., Some Vector Optimization Problems in Banach Spaces with Generalized Convexity. *Computers and Mathematics with Applications*, 54: 1403-1410, (2007).

[20] Zalinescu C., *Convex Analysis in General Vector Spaces*. World Scientific Publishing Co, Singapur, (2002).