UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO

ESCUELA DE POSGRADO

MAESTRÍA EN MATEMÁTICAS

TESIS:

OPTIMIZACIÓN PARA FUNCIONES CONVEXAS NO DIFERENCIABLES

PRESENTADO POR:

BR. CARMEN LAIME LLOCLLA

PARA OPTAR AL GRADO ACADÉMICO DE MAESTRO EN MATEMÁTICAS

ASESOR: MGT. JOSE MOZO AYMA

CUSCO – PERÚ

2019

PRESENTACIÓN

Señor director de la Escuela de Posgrado de la Universidad Nacional de San Antonio Abad del Cusco, en cumplimiento del Reglamento para optar al grado académico de Maestro en Matemáticas, pongo en consideración el presente trabajo de tesis intitulado: “**OPTIMIZACIÓN PARA FUNCIONES CONVEXAS NO DIFERENCIABLES”,** para su revisión, sustentación y posterior inscripción en el libro correspondiente.

Para optimizar una función convexa no diferenciable es necesario escribirla como una composición de otras funciones, de tal forma que la función resultante sea una función convexa. Además determinar el subgradiente de la función en los puntos factibles y determinar el hiperplano de soporte en estos puntos. Se determina la condición de primer orden, asi como también una condición de regularidad y determinar la condición de segundo orden.

DEDICATORIA

A mis Padres

A mi Hermano

A mi Hijo

AGRADECIMIENTOS

Gracias a mi familia, que son mi soporte.

Gracias a mi asesor, por enseñarme y orientarme.

Gracias a todos los que hicieron posible que termine esta tesis, muchos con las palabras justas en los momentos necesarios.

ÍNDICE GENERAL

LISTA DE FIGURAS

OPTIMIZACIÓN PARA FUNCIONES CONVEXAS NO DIFERENCIABLES

RESUMEN

El objetivo de la presente tesis es determinar la existencia de valores óptimos, para funciones convexas no diferenciables, solo con la condición de continuidad.

En el primer captitulo se desarrollo el plantemiento del problema, se planteó la situsión problemática, formulación, justificación, objeivos e hipótesis. Para el segundo capitulo se desarrolló la teoria básica del análisis convexo, para obtener la generalización de la derivada para funciones convexas diferenciables y no diferenciables, esto determinando un hiperplano de soporte en el punto de interés, se introducen algunos conceptos y propiedades de subgradiente y la generalización de la diferenciabilidad. En el tercer capitulo se determinaron las condiciones de existencia de valores óptimos para funciones convexas no diferenciables, para ello se planteó el problema primario, asi como también problema dual. Se determinaron las condiciones de primer orden, segundo orden y una condición de regularidad.

Se concluyó que un problema de optimización convexa no diferenciable posee valores óptimos bajo las condiciones necesarias de primer y segundo orden, solo con la condición de continuidad.

PALABRAS CLAVES

Conjunto convexo, función convexa, Subdiferencial, gradiente generalizado, valores óptimos..

OPTIMIZATION FOR NONSMOOTH CONVEX FUNCTIONS

ABSTRACT

The objective of this thesis is to determine the existence of optimal values, for nonsmooth convex functions, only with the condition of continuity.

In the first chapter the problem statement was developed, the problematic situation, formulation, justification, objectives and hypotheses were raised. For the second chapter is the basic theory of convex analysis, to obtain the generalization of the derivative for differentiable and nonsmooth convex functions, this determining a support hyperplane at the point of interest, some concepts and properties of subgradient are introduced and the generalization of differentiability. In the third chapter the conditions of existence of optimal values for nonsmooth convex functions are determined, for this the primary problem was posed, as well as the dual problem. The conditions of first order, second order and a condition of regularity were determined.

It was concluded that a nonsmooth convex optimization problem has optimal values under the necessary first and second order conditions, only with the condition of continuity.

KEY WORDS

Convex set, convex function, Subdifferential, generalized gradient, optimal values.

INTRODUCCIÓN

La investigación de operaciones fue reconocida como una nueva área de investigación científica durante la segunda guerra mundial. Como todos los avances de la ciencia, estos surgen para solucionar problemas, en este caso el propósito fue reducir costos del ejército y aumentar las pérdidas del enemigo. La investigación de operaciones es una herramienta importante para la toma de decisiones.

Entre las técnicas de la investigación de operaciones más desarrolladas encontramos la programación lineal, que se ocupa de los modelos en los que la función objetivo, así como las restricciones, son lineales. Para solucionar este tipo de programas surgen varios métodos, entre los más conocidos está el método simplex.

Con el transcurso del tiempo se han desarrollado otro tipo de métodos que ha dado lugar a la programación No lineal, a la programación lineal entera, a la programación multiobjetivo, entre otras.

Para encontrar una solución óptima de los problemas de programación, la convexidad desempeña un papel fundamental, es sabido que sobre la hipótesis de convexidad las condiciones necesarias y suficientes de optimalidad son triviales, las cuales garantizan la existencia de los puntos óptimos.

Muchos problemas de optimización incluyen funciones no diferenciables. Estos problemas no se dejan resolver por los algoritmos clásicos desarrollados. La necesidad de crear nuevos métodos, para resolver estos problemas, condujo al desarrollo del campo de la optimización no diferenciable, que es muy reciente, su conocimiento se inicia entre 1975 y 1980.

Un concepto muy importante para el desarrollo de la teoría de optimización para funciones convexas no diferenciables, es el subdiferencial, que se puede considerar como un equivalente a la noción de derivada. Las reglas de cálculo clásicas se pudieron desarrollar con naturalidad. La primera extensión fue la clase de funciones lipschitzianas, este estudio lo realizó F. H. Clarke, introduciendo la noción de gradiente generalizado.

Para la solución de problemas de optimización no diferenciable, se requieren el cálculo del subdiferencial, se pueden distinguir dos clases básicas, las cuales son: Métodos de gradiente generalizado y el método de planos de corte. Otra técnica es utilizar el operador próximo. El operador próximo de una función convexa es una extensión natural de la noción de proyección sobre un conjunto convexo. (Nava Manzo, 2015)

En el presente trabajo desarrollaremos los conceptos básicos del análisis convexo, la optimización diferenciable y la optimización no diferenciable, para determinar la solución de este último tipo de problemas.

CAPÍTULO I: PLANTEAMIENTO DE INVESTIGACIÓN

1.1 PLANTEAMIENTO DEL PROBLEMA

1.1.1 SITUACIÓN PROBLEMÁTICA

A través de la historia la necesidad de optimizar ciertos procesos y recursos ha sido una preocupación permanente del hombre, para este fin se desarrolló la investigación de operaciones que tiene sus inicios y mayor énfasis en la segunda guerra mundial donde se desarrollaron muchos métodos para solucionar los llamados problemas de programación lineal, así por ejemplo surge el método simplex. Los problemas que envuelven un objetivo único de los modelos de optimización y sus técnicas puedan minimizar (o maximizar) una función, pero este no siempre es lineal, podría ser de varias variables o ser convexa o tener características propias del problema a programar.

Para encontrar una solución óptima de los problemas de programación, la convexidad desempeña un papel fundamental, es sabido que sobre la hipótesis de convexidad las condiciones necesarias y suficientes de optimalidad son triviales, las cuales garantizan la existencia de los puntos óptimos (eficientes), pero que pasa con las funciones convexas que no son diferenciables.

Las preguntas inmediatas serian si la función es continua y si es posible minimizar (maximizar) este tipo de funciones en un punto dado.

En el presente trabajo desarrollaremos la teoría de optimización de funciones convexas no diferenciables y para ello utilizaremos el concepto de subdiferenciabilidad en los puntos donde la función no es diferenciable.

1.1.2 FORMULACIÓN DEL PROBLEMA

En problemas de optimización se requiere que la función objetivo sea convexa, para asegurar la existencia de una solución óptima mínima o una solución óptima máxima. Si la función objetivo además es no diferenciable, se utilizará la subdiferenciabilidad, para lo que es necesario conocer la teoría de separabilidad de espacios vectoriales topológicos, así como también, la semicontinuidad.

1.1.2.1 PROBLEMA GENERAL

¿Es posible encontrar valores óptimos para funciones convexas no diferenciables, solo con la condición de continuidad?

1.1.2.2 PROBLEMA ESPECÍFICOS

a) ¿Es posible generalizar el concepto de derivada para funciones convexas diferenciables?
b) ¿Es posible determinar un equivalente de la derivada para funciones convexas no diferenciables?
c) ¿Es posible determinar un valor óptimo para funciones convexas no diferenciables?

1.1.3 JUSTIFICACIÓN DE LA INVESTIGACIÓN

En el modelamiento de problemas de optimización en computación, ingeniería y ciencias, se incluyen funciones convexas que no necesariamente son diferenciables. Esto genera un problema al momento de buscar la solución de dichos problemas, muchas veces no se logra encontrar esa solución, no porque el modelo este mal, sino que se deben incluir otras condiciones para poder encontrar el valor óptimo. En consecuencia, se han creado nuevos métodos, que condujeron al desarrollo del campo de la optimización no diferenciable, que se inicia entre 1975 y 1980.

1.1.4 OBJETIVOS DE LA INVESTIGACIÓN

1.1.4.1 OBJETIVO GENERAL

Determinar la existencia de valores óptimos para funciones convexas no diferenciables, solo con la condición de continuidad.

1.1.4.2 OBJETIVOS ESPECÍFICOS

a) Obtener la generalización de la derivada para funciones convexas diferenciables.

b) Determinar un equivalente de la derivada para funciones convexas no diferenciables.

c) Determinar las condiciones de existencia de valores óptimos para funciones convexas no diferenciables.

1.2 ANTECEDENTES DE LA INVESTIGACIÓN

1.2.1 ANTECEDENTES INTERNACIONALES

Nava, R. (2015), realizó la investigación: *"Funciones convexas no diferenciables",* para obtener el grado de Maestro en Ciencias en la Universidad Autónoma Metropolitana México. El objetivo de la investigación fue estudiar diversos métodos (subgradiente, bundle y métodos con operador próximo) utilizados para resolver problemas de optimización no diferenciable, aplicarlos a una serie de problemas y comparar los resultados con otros métodos de la literatura. La investigación llego a las siguientes conclusiones:

1. "El método de subgradiente es fácil de implementar, sin embargo, determinar el subdiferencial de cualquier función no es simple, por lo que el método de subgradiente, aunque es fácil de implementar, presenta complicaciones al momento de encontrar el subgradiente de la función en cada iteración" (Nava Manzo, 2015)

2. "El método bundle permite tener un método con muchas propiedades útiles, entre las cuáles está el tener un criterio de paro que no depende del conocimiento del punto óptimo.

3. Además, aunque el método bundle requiere del cálculo de subgradientes en cada iteración, otorga una manera de elegir un subgradiente adecuado en cada iteración" (Nava Manzo, 2015)

4. La regularización de Moreau-Yosida tiene la ventaja que puede obtenerse de manera explícita, lo cual es muy útil al momento de programar. (Nava Manzo, 2015)

5. Los métodos de punto próximo y gradiente próximo son sencillos de implementar computacionalmente ya que solo necesitan el cálculo del operador próximo de la función a minimizar. (Nava Manzo, 2015)

6. La convergencia de todos los métodos estudiados en este trabajo no dependen del punto inicial elegido, lo cual es una ventaja muy grande al momento de aplicarlos a problemas no diferenciables (incluso problemas diferenciables). (Nava Manzo, 2015)

Vera, C. (2006), realizó la investigación: *"Existencia y condiciones de optimización vectorial no convexa",* para obtener el grado de Doctor en Ciencias Aplicadas con mención en Ingeniería Matemática en la Universidad de concepción en Concepción Chile.

El objetivo de la investigación fue estudiar las propiedades de los mínimos vectoriales débilmente eficientes, bajo hipótesis de convexidad generalizada. La investigación llego a las siguientes conclusiones:

1. Se encontraron mínimos débiles eficientes, cuando el dominio de la función vectorial son números reales y que cumplen un tipo de diferenciabilidad generalizada, sin hipótesis de diferenciabilidad. (Vera Donoso, 2006)

2. Cuando el dominio de la función vectorial es un subconjunto de los números reales, de recorrido bidimensionesl y sus componentes son casiconvexas, sin hipótesis de diferencibilidad. (Vera Donoso, 2006)

3. Se encuentra una caracterización del caso estudiado que permite elaborar un algoritmo de tiempo finito, para calcular soluciones débiles eficientes y el supremo del conjunto de mínimos débiles eficientes. (Vera Donoso, 2006)

1.2.2 ANTECEDENTES NACIONALES

Villavicencio P. (2013), realizó la investigación: "*Condiciones de Kuhn-Tucker en la Optimización no lineal Convexa*", para optar al grado académico de Maestro en Matemática en la Escuela de Posgrado de la Universidad Nacional San Antonio Abad del Cusco.

El objetivo general de la investigación fue explicar y analizar un método de optimización aplicable a problemas de optimización no lineal, donde la función objetivo como las funciones de restricción son funciones de clase C^2.

. La investigación llego a las siguientes conclusiones:

1. Para que $X_0 \in \mathbb{E}^n$, sea una solución óptima del problema de optimización no lineal convexa diferenciable, es necesario y suficiente que satisfaga todas las condiciones de Kuhn – Tucker.

2. Las condiciones de Kuhn – Tucker son una generalización del método de Lagrange para problemas condicionados no lineales convexos y diferenciables, puesto que las funciones de restricciones del problema incluyen funciones de restricciones con desigualdades en los que una o más de las variables es no lineal.

3. Las condiciones de Kuhn – Tucker son importantes para encontrar la solución óptima de problemas de optimización no lineal convexa. En el sentido de que muestran que una solución local óptima es una solución óptima global del problema no lineal convexo dado.

Borda D. (2013), realizó la investigación: *"Convergencia del Método de punto proximal con distancia homogénea de orden r en optimización convexa"*, para optar el Título profesional de Licenciado en Matemática en la Universidad Nacional del Callao, Facultad de Ciencias Naturales y Matemática.

El objetivo general de la investigación fue analizar la convergencia del método del punto proximal. La investigación llego a las siguientes conclusiones:

1. En esta tesis hemos recopilado resultados de la convergencia del método del punto proximal con la distancia homogénea de orden dos presentado por Auslender, Ben tiba y Teboulle en 1999 y adaptamos para distancia homogénea de orden r conservando que las propiedades de homogeneidad se verifican al igual que la convergencia hacia un punto óptimo. Esta tesis puede mejorarse el algoritmo usando en el sentido computacional, como se puede corroborar en distintas investigaciones en la actualidad. (Borda Marcatinco, 2013)
2. Espero que este trabajo de tesis sirva como referencia para futuros trabajos de investigación que se puedan realizar, pues hay otros detalles por explorar, tanto en el aspecto teórico como computacional, dado su importancia en os distintos campos de la Ingeniería y Economía. (Borda Marcatinco, 2013)

Gimaray, H (1989), realizó la investigación: *"Optimización no diferenciable"*, para obtener el grado académico de Magister en Matemáticas en la Universidad Nacional de Ingeniería Facultad de Ciencias. El objetivo de la investigación fue estudiar la optimización no diferenciable. La investigación llego a las siguientes conclusiones:

1. El concepto de subdiferenciabilidad constituye una generalización de diferenciabilidad.
2. Las funciones conjugadas proporcionan una útil herramienta para investigar propiedades de conjuntos convexos y funciones convexas.

1.3 HIPÓTESIS Y VARIABLES

1.3.1 HIPÓTESIS

1.3.1.1 HIPÓTESIS GENERAL

Existen valores optímos para funciones convexas no diferenciables, solo con la condición de continuidad.

1.3.1.2 HIPÓTESIS ESPECÍFICA

a. Existe una generalización de la derivada para funciones convexas diferenciables.
b. Es posible determinar un equivalente derivada para funciones convexas no diferenciables.
c. Existen condiciones de existencia de valores óptimos para funciones convexas no diferenciables.

1.3.2 IDENTIFICACION DE VARIABLES

1.3.2.1 VARIABLE INDEPENDIENTE

Funciones continuas

1.3.2.2 VARIABLE DEPENDIENTE

Valores óptimos de funciones convexas no diferenciables.

1.4 METODOLOGÍA

1.4.1 TIPO Y NIVEL DE INVESTIGACION.

Según (Hernández, 2014), el enfoque cualitativo utiliza la recolección de datos sin medición numérica para descubrir o afirmar preguntas de investigación en el proceso de interpretación. Por lo tanto el presente trabajo tiene un enfoque cualitativo. Según el objetivo del estudio, el trabajo corresponde a una investigación pura o teórica, además se realizara a nivel Exploratorio (Crisóstomo).

1.4.2 UNIDAD DE ANALISIS

Se estudiaran a las funciones convexas no diferenciables.

1.4.3 TECNICAS DE RECOLECCION DE DATOS E INFORMACION

La información se recolectara de libros y trabajos de investigación relacionados al tema.

CAPÍTULO II: MARCO TEÓRICO

2.1. PRELIMINARES

DEFINICIÓN 2.1.1. Sea X un espacio vectorial real. Sean C y D subconjuntos de X, y sea $z \in X$. Se definen los siguientes conjuntos:

$$C + D := \{x + y / x \in C \ \wedge \ y \in D\}$$

$$C - D := \{x - y / x \in C \ \wedge \ y \in D\}$$

$$z + C \coloneqq \{z\} + C$$

$$C - z \coloneqq C - \{z\}$$

$$\lambda C \coloneqq \{\lambda x / x \in C\}, \ \ \forall \lambda \in \mathbb{R}$$

Si A es un subconjunto no vacío de $\mathbb{R}$ entonces:

$$AC = \bigcup_{\lambda \in A} \lambda C \quad \text{y} \quad Az = A\{z\} = \{\lambda z / \lambda \in A\}.$$

DEFINICIÓN 2.1.2.- Sea X un espacio vectorial real cualquiera. Una norma en X es una aplicación $\| \quad \|: X \to \mathbb{R}_0^+$ que cumple las siguientes condiciones:

1. $\|x\| \geq 0$, para cada x de X. Además $\|x\| = 0$ si y solo si $x = \vec{0}$.
2. Linealidad: $\|\alpha x\| = |\alpha| \|x\|$, para cada x de X y todo escalar α.

3. Desigualdad triangular: $\|x+y\| \leq \|x\| + \|y\|$, para cada x, y de X.

DEFINICIÓN 2.1.3.- Un espacio vectorial normado o simplemente espacio normado es un par $(X, \| \quad \|)$, donde X es un espacio vectorial real o complejo y $\| \quad \|$ es una norma en X.

DEFINICIÓN 2.1.4. (Producto interno).- Sea X un espacio vectorial (real o complejo) sobre el campo $\mathbb{R}$. Un producto escalar o producto interno definido sobre X es una aplicación

$$\langle \, , \rangle : X \times X \rightarrow \mathbb{R}_0^+$$

que verifica:

1. $\langle x, x \rangle \geq 0, \ \ \forall x \in X.$

   Además $\langle x, x \rangle = 0$ si y solo si $x = 0_X$

2. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle, \quad \forall x, y, z \in X \ , \alpha, \beta \in \mathbb{K}$

3 $\langle x, y \rangle = \langle y, x \rangle, \quad \forall x. y \in X$

Observemos que:

$$\|x\| = \sqrt{\langle x, x \rangle} \ \ , \forall x \in X$$

DEFINICIÓN 2.1.5.- Sean X e Y espacios vectoriales normados. Una aplicación $f: X \rightarrow Y$ es continua en $x_0 \in X$ si y solo si, para toda sucesión $\{x_n\}_{n \in \mathbb{N}}$ que converge a x_0 en X se tiene una sucesión $\{f(x_n)\}_{n \in \mathbb{N}}$ que converge a $f(x_0)$.

Esto es equivalente a:

f es continua en $x_0 \Leftrightarrow \|x_n - x_0\| \rightarrow 0$ implica $\|f(x_n) - f(x_0)\| \rightarrow 0$

DEFINICIÓN 2.1.6.- Sean X e Y espacios vectoriales normados y $f: X \to Y$ una aplicación. Se dice que f satisface la condición de Lipschitz (o f es una aplicación Lipschitzina), si existe una constante $k > 0$, tal que:

$$\|f(x) - f(y)\| \leq k\|x - y\| \,, \;\; \forall x, y \in X$$

donde k se llama constante de Lipschitz.

TEOREMA 2.1.1.- Sean X e Y espacios vectoriales normados y $f: X \to Y$ una aplicación lineal. Si f es lipschitzina en X entonces f es continua en X.

DEMOSTRACIÓN.

Supongamos que:

$$\|f(x)\| \leq k\|x\| \,, \;\; \forall x \in X$$

Si $(x_n) \to x_0$ entonces

$$\|f(x_n) - f(x_0)\| = \|f(x_n - x_0)\| \leq k\|x_n - x_0\| \to 0$$

Lo que implica que $f(x_n) \to f(x_0)$, esto prueba que f es continua en x_0.

DEFINICIÓN 2.1.7.- Sea $D \subset \mathbb{R}^n$. La aplicación $f: D \to \mathbb{R}^m$ es localmente lipschitzina, si existe $V_r(x)$, $r > 0$ tal que f es lipschitzina en $D \cap V_r(x)$, donde $V_r(x)$ es una vecindad de x.

DEFINICIÓN 2.1.8.- Sean X e Y espacios vectoriales normados. La aplicación lineal $f: X \to Y$ se llama acotada si existe un número real $k > 0$, tal que:

$$\|f(x)\| \leq k \;\; , \forall x \in X$$

DEFINICIÓN 2.1.9.- Sea $D \subset \mathbb{R}^n$. La aplicación $f: D \to \mathbb{R}^m$ es localmente acotada, si existe $V_r(x)$, $r > 0$ tal que f es acotada en $D \cap V_r(x)$.

DEFINICIÓN 2.1.10.- Sea $(X, \|\,.\,\|)$ espacio vectorial normado real su espacio dual topológico, esta definido por:

$$X^* := \{x^*: X \to \mathbb{R} / x^* \; es \; una \; aplicación \; lineal \; y \; continua\}$$

En el espacio dual topológico se define la norma como:

$$\|x^*\|_*:= \sup_{\substack{x \in X \\ ||x|| \leq 1}} \langle x, x^* \rangle \quad x^* \in X^*$$

donde $\langle \ , \ \rangle$ denota el producto interno interior de dualidad entre X y X^*; es decir, para una aplicación lineal continua $x^* \in X^*$ definida sobre el espacio X con valores en $\mathbb{R}$, se tiene x^* : $X \to \mathbb{R}$ y escribiremos:

$$\langle x, x^* \rangle = x^*(x) \qquad \forall x \in X$$

2.2. CONJUNTOS CONVEXOS

DEFINICIÓN 2.2.1.- Sea $A \subseteq \mathbb{R}^n$ un conjunto no vacío. Decimos que A es convexo si para cada $x, y \in A$

$$[\lambda x + (1 - \lambda) y] \in A, \quad \text{para cualquier } \lambda \in [0,1] \qquad (2.2.1)$$

LEMA 2.2.1.- Si A_1 y A_2 conjuntos convexos en $\mathbb{R}^n$ entonces:

3. $A_1 \cap A_2$ es convexo.
4. $A_1 + A_2$ es convexo.
5. $A_1 - A_2$ es convexo

DEFINICIÓN 2.2.2.- Sean $\mathbb{R}^n$ un espacio vectorial normado, ξ vector no nulo de $\mathbb{R}^n$ y $\alpha \in \mathbb{R}$. Se define el hiperplano H como la colección de la forma $\{x \in \mathbb{R}^n / \langle x, \xi \rangle = \alpha\}$

esto es:

$$H \coloneqq \{x \in \mathbb{R}^n / \langle x, \xi \rangle = \alpha\} \tag{2.2.2}$$

donde: $\xi \neq \vec{0}$ en $\mathbb{R}^n$ y $\alpha \in \mathbb{R}$.

Observemos que:

El vector ξ es llamado el vector normal al hiperplano.

En particular, el hiperplano H es un conjunto convexo.

En efecto:

Sean x, $y \in H$, por definición se tiene:

$$\langle x, \xi \rangle = \alpha$$

$$\langle y, \xi \rangle = \alpha$$

Luego para algún $\lambda \in [0,1]$, se tiene:

$$\begin{aligned}
\langle \lambda x + (1-\lambda)y, \xi \rangle &= \langle \lambda x, \xi \rangle + \langle (1-\lambda)y, \xi \rangle \\
&= \lambda \langle x, \xi \rangle + (1-\lambda)\langle y, \xi \rangle \\
&= \lambda\alpha + (1-\lambda)\alpha \\
&= \lambda\alpha + \alpha - \lambda\alpha = \alpha
\end{aligned}$$

entonces $\lambda x + (1-\lambda)y \in H$.

Por consiguiente H es convexo.

Un hiperplano divide al espacio $\mathbb{R}^n$ en dos semiespacios cerrados denotados por:

$$H^{+} = \{x \in \mathbb{R}^n / \langle x, \xi \rangle \geq \alpha\}$$

$$H^{-} = \{x \in \mathbb{R}^n / \langle x, \xi \rangle \leq \alpha\}$$

Claramente los semiespacios son conjuntos convexos.

Notemos que cualquier punto en $\mathbb{R}^n$ está en H^{+} ó H^{-} o en ambos.

También un hiperplano H y los correspondientes semiespacios se pueden escribir en referencia a un punto fijo $\bar{x} \in H$. Si $\bar{x} \in H$, $\langle \bar{x}, \xi \rangle = \alpha$ y para cualquier punto $x \in H$ debe satisfacer:

$$\langle x, \xi \rangle = \alpha$$

Entonces:

$$\langle x, \xi \rangle - \langle \bar{x}, \xi \rangle = \alpha - \alpha = 0$$

Esto implica: $\langle x - \bar{x}, \xi \rangle = 0$.

Los semiespacios H^{+} y H^{-} se pueden escribir como:

$$H^{+} = \{x \in \mathbb{R}^n / \langle x - \bar{x}, \xi \rangle \geq 0\}$$

$$H^{-} = \{x \in \mathbb{R}^n / \langle x - \bar{x}, \xi \rangle \leq 0\}$$

DEFINICIÓN 2.2.3.- Sea $A \subseteq \mathbb{R}^n$ un conjunto no vacío.

1. Se define la envoltura convexa de A, denotada por $co(A)$, al menor conjunto convexo que contiene a A, donde:

$$co(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i \middle/ \sum_{i=1}^{n} \lambda_i = 1 \ , \quad \lambda_i \in [0,1] \, , x_i \in A, \forall i = 1, \dots, n \, ; \quad n \in \mathbb{N} \right\}$$

donde:

$$\sum_{i=1}^{n} \lambda_i x_i \text{ , con } \sum_{i=1}^{n} \lambda_i = 1$$

se llama combinación convexa de $x_1, x_2, \ldots, x_n \in A$.

2. A es un cono, si y solo si, $tA \subseteq A, \quad \forall t \geq 0$.
3. Se define $cone(A)$, al menor cono que contiene a A, esto es:

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

4. Dado el conjunto $M \subset \mathbb{R}^n$. Denotaremos por $M^{\perp}$ al subespacio ortogonal de M en donde

$$M^{\perp} = \{x \in \mathbb{R}^n / \langle x, y \rangle = 0, \quad \forall y \in M\}$$

De la teoría básica de geometría en espacios vectoriales normados se tiene que un hiperplano cerrado H en X puede ser representado por:

$$H = \{x \in X : \langle x, x^* \rangle = \alpha\}$$

para algún $x^* \in X^*$ y $\alpha \in \mathbb{R}$. De igual modo, un semiespacio cerrado H de X puede ser representado por

$$H^- = \{x \in X : \langle x, x^* \rangle \leq \alpha\}$$

para algún $x^* \in X^*$ y $\alpha \in \mathbb{R}$.

PROPOSICIÓN 2.2.1.- Sea A un conjunto en $\mathbb{R}^n$. A es convexo si y solo si contiene todas las combinaciones convexas de sus elementos. *(Nava Manzo, 2015)*

La demostración se puede encontrar en Convex Analysis de Rockafellar. (1970, Pag. 34)

DEFINICIÓN 2.2.4.- Sean los espacios vectoriales normados $\mathbb{R}^m$, $\mathbb{R}^n$ y un vector $c \in \mathbb{R}^n$. La función $f: \mathbb{R}^m \to \mathbb{R}^n$ se denomina afín, cuando existe una aplicación lineal $T: \mathbb{R}^m \to \mathbb{R}^n$ y , tales que:

$$f(x) = Tx + c$$

En particular, si la aplicación lineal $T: \mathbb{R}^n \to \mathbb{R}$ se define como:

$$Tx = \langle x, \xi \rangle \text{ , con } \xi \in \mathbb{R}^n$$

se tiene que: $\{x \in \mathrm{R}^n / Tx + c = 0, \ \ c \in \mathbb{R}\}$ es un conjunto de nivel de la función afín f.

DEFINICIÓN 2.2.5.- Sea A un subconjunto de $\mathbb{R}^n$ y $x \in A$. Un hiperplano en $\mathbb{R}^n$ se dice hiperplano de soporte de A en x si se cumple que:

$$[x \in A \cap H \quad \wedge \quad (A \subset H^+)] \quad \vee \quad (A \subset H^-)$$

DEFINICIÓN 2.2.6 (Soporte de conjuntos en sus puntos de frontera).- Sea A un subconjunto no vacio de $\mathbb{R}^n$ y $\bar{x} \in \partial A$ (frontera de A). Un hiperplano

$$H = \{x \in \mathbb{R}^n / \langle x - \bar{x}, \xi \rangle = 0\} \text{ , } \ \xi \in \mathbb{R}^n$$

es llamado un hiperplano de soporte de A en $\bar{x}$ si $A \subseteq \mathrm{H}^+$, esto es:

$$\langle x - \bar{x}, \xi \rangle \geq 0, \ \ \forall x \in A \text{ ó } A \subseteq \mathrm{H}^-$$

esto es:

$$\langle x - \bar{x}, \xi \rangle \leq 0, \ \ \forall x \in A.$$

Además, si $A \subset H$, H es llamado hiperplano de soporte propio de A en $\bar{x}$.

Equivalentemente:

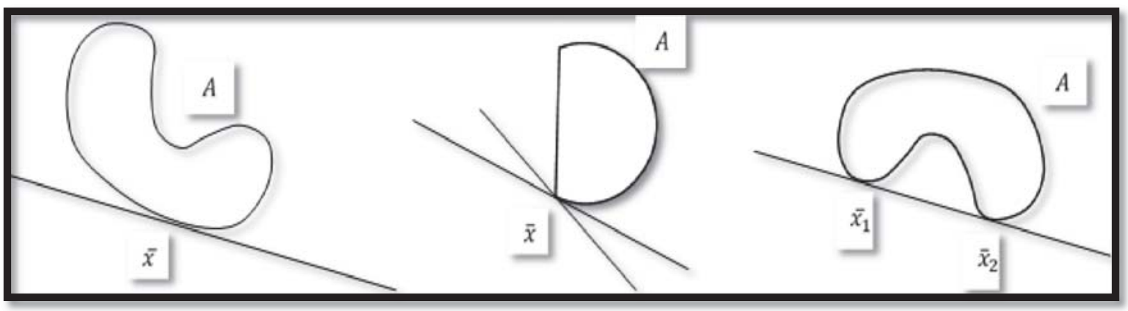
$H = \{x \in \mathbb{R}^n / \langle x - \bar{x}, \xi \rangle = 0\}$ es un hiperplano de soporte de A en $\bar{x} \in \partial A$, si:

$$\langle \bar{x}, \xi \rangle = \inf\{\langle x, \xi \rangle / x \in A\}$$

ó

$$\langle \bar{x}, \xi \rangle = \sup\{\langle x, \xi \rangle / x \in A\}$$

FIGURA 2.2.1. Hiperplanos de soporte



Fuente: M. Bazzara, 2006

TEOREMA 2.2.1.- Sean A un subconjunto no vacío convexo y cerrado de $\mathbb{R}^n$ y $z \notin A$. Existe un único punto $\bar{x} \in A$, tal que la distancia de $\bar{x}$ a z es mínima. Además, $\bar{x}$ es el punto minimal, si y solo si:

$$\langle x - \bar{x}, z - \bar{x} \rangle \leq 0, \quad \forall x \in A$$

La demostración de puede encontrar en Nonlinear Programming de Mokhtar S. Bazaraa y otros. (2006, Pag. 51)

TEOREMA 2.2.2. Sea A un conjunto convexo no vacio de $\mathbb{R}^n$ y $\bar{x} \in \partial A$. Existe un vector ξ no nulo de $\mathbb{R}^n$ tal que:

$$\langle x - \bar{x}, \xi \rangle \leq 0, \quad \forall x \in \bar{A}$$

La demostración se puede encontrar en Nonlinear Programming de Mokhtar S. Bazaraa y otros. (2006, Pag. 58)

DEFINICIÓN 2.2.7.- Sean A_1 y A_2 conjuntos no vacíos en $\mathbb{R}^n$. Un hiperplano separa a A_1 y A_2 si:

$$\langle x, \xi \rangle \geq \alpha; \quad \forall x \in \mathrm{A}_1$$

$$\langle x, \xi \rangle \leq \alpha; \quad \forall x \in \mathrm{A}_2$$

Si $A_1 \cup A_2 \not\subset H$, se dice que H esta propiamente separado por A_1 *y* A_2*.*

El hiperplano H esta estrictamente separado por A_1 *y* A_2 si:

$$\langle x, \xi \rangle > \alpha; \quad \forall x \in \mathrm{A}_1$$

$$\langle x, \xi \rangle < \alpha; \quad \forall x \in \mathrm{A}_2$$

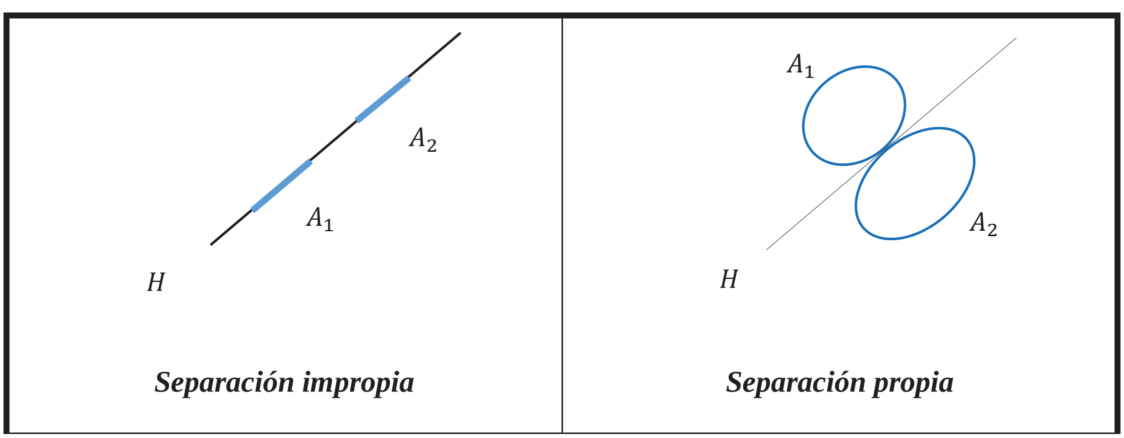
El hiperplano H está fuertemente separado por A_1 *y* A_2 si

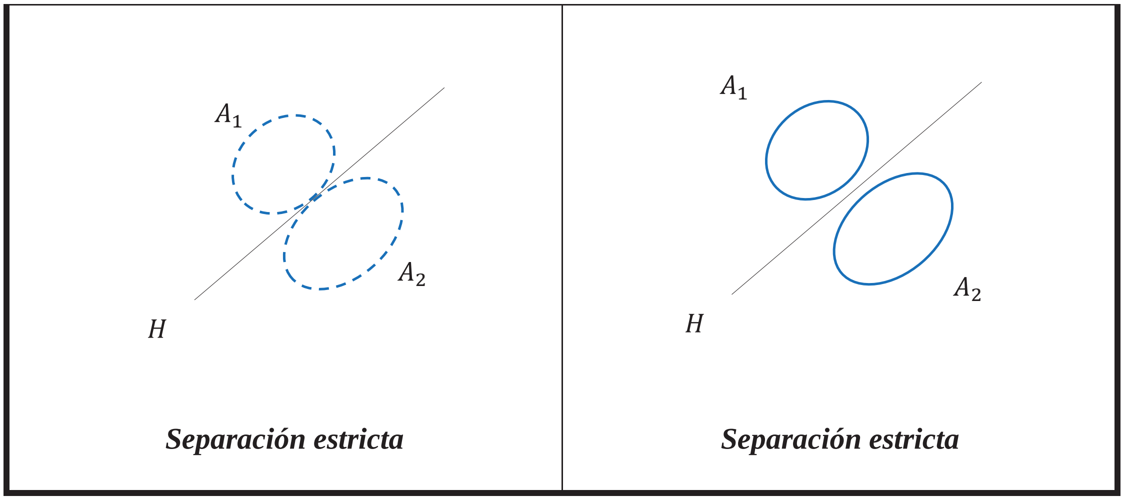
$$\langle x, \xi \rangle \geq \alpha + \varepsilon; \quad \forall x \in \mathrm{A}_1, \ \varepsilon \text{ es un escalar positivo}$$

$$\langle x, \xi \rangle \leq \alpha; \quad \forall x \in A_2$$

La separación impropia usualmente no es de mucho valor, esta corresponde a un hiperplano que contiene a A_1 y A_2.

FIGURA 2.2.2: Tipos de separación de conjuntos convexos





Fuente: M. Bazzara, 2006

TEOREMA 2.2.3. (Teorema de separación de Hahn-Banach). Sea A un subconjunto no vacío, convexo y cerrado de $\mathbb{R}^n$ y $z \notin A$. Entonces existe un vector $\xi \in \mathbb{R}^n$ no nulo y un escalar α tal que:

$$\langle z, \xi \rangle > \alpha \quad \text{y} \quad \langle x, \xi \rangle \leq \alpha \quad \text{para cada } x \in A$$

es decir, existe un hiperplano

$$H = \{x \in \mathbb{R}^n : \langle x, \xi \rangle = \alpha\}$$

tal que $z \in H^+$ y $A \subset H^-$.

DEMOSTRACIÓN

Supóngase que A es un conjunto, no vacío, convexo y cerrado en $\mathbb{R}^n$ y $z \notin A$. Por el Teorema (2.2.1), existe un único punto $\bar{x}$ en A, tal que:

$$\langle x - \bar{x}, z - \bar{x} \rangle \leq 0, \quad \forall x \in A$$

Entonces

$$\langle x, z - \bar{x} \rangle - \langle \bar{x}, z - \bar{x} \rangle \leq 0$$

$$\langle z - \bar{x}, x \rangle - \langle \bar{x}, z - \bar{x} \rangle \leq 0$$

$$\langle z-\bar{x}, x\rangle \leq \langle \bar{x}, z-\bar{x}\rangle \tag{2.2.1}$$

considerando:

$$\xi = z-\bar{x} \neq 0 \quad \text{y} \quad \alpha = \langle \bar{x}, z-\bar{x}\rangle = \langle \bar{x}, \xi\rangle \tag{2.2.2}$$

Sustituyendo *(2.2.2)* en *(2.2.1)* se obtiene:

$$\langle x, \xi\rangle \leq \alpha; \quad \text{para cada } x \text{ en } A$$

Por otro lado, si $\alpha = \langle \bar{x}, z-\bar{x}\rangle$, se puede escribir:

$$\langle z, z-\bar{x}\rangle - \alpha = \langle z, z-\bar{x}\rangle - \langle \bar{x}, z-\bar{x}\rangle = \langle z-\bar{x}, z-\bar{x}\rangle = \|z-\bar{x}\|^2 > 0, \;\; z-\bar{x} \neq 0$$

$$\langle z, z-\bar{x}\rangle - \alpha > 0$$

De donde se tiene que:

$$\langle z, z-\bar{x}\rangle > \alpha \tag{2.2.3}$$

Reemplazando *(2.2.2)* en *(2.2.3)*, tenemos lo que queríamos demostrar:

$$\langle z, \xi\rangle > \alpha; \quad \text{para cada } x \text{ en } A.$$

Como consecuencia de este resultado, podemos afirmar que cada $z \notin A$ puede ser fuertemente separado de A por un hiperplano cerrado.

COROLARIO 2.2.1. Sea A un conjunto convexo cerrado de $\mathbb{R}^n$. A es la intersección de todos los semiespacios cerrados que lo contienen.

Esto es:

$$A = \bigcap_i H_i \,, \qquad A \subseteq H_i$$

donde, H_i es un semiespacio cerrado que contiene a A

DEMOSTRACIÓN

Sea A un conjunto convexo cerrado de $\mathbb{R}^n$ y H_i es un semiespacio cerrado que contiene a A.

Debemos mostrar que: $A \subseteq \bigcap_i H_i \quad \wedge \quad \bigcap_i H_i \subseteq A$

i. Como A está contenido en cada semiespacio entonces también está en la intersección de todos los semiespacios que lo contienen. Es decir:

$$A \subseteq \bigcap_i H_i$$

ii. Para demostrar $\bigcap_i H_i \subseteq A$, Se utilizara la demostración por el absurdo. Es decir, existe un elemento:

$$z \in \bigcap_i H_i \quad \text{y} \quad z \notin A$$

Por el teorema de separación de Hahn – Banach, existe un semiespacio cerrado que contiene a A pero no a z, lo que es una contradicción, ya que estamos asumiendo que z está en $\bigcap_i H_i$, con lo que queda demostrado el corolario.

Por i. y ii. se tiene $A = \bigcap_i H_i \,, \qquad A \subseteq H_i$

Observemos que, los conjuntos cerrados H_i del corolario anterior también resultan ser cerrados con la topología débil y por lo tanto el conjunto convexo cerrado A es cerrado para la topologa débil (intersección de cerrados débiles).

Así, una conclusión importante, y que será utilizada más adelante, del Corolario 2.2.1, es que todo conjunto convexo cerrado es también cerrado para la topología débil.

TEOREMA 2.2.4 (Separación de dos conjuntos convexos).- Sean A_1 y A_2 dos subconjuntos disjuntos, convexos y no vacíos de $\mathbb{R}^n$. Entonces existe un hiperplano que separa A_1 y A_2; esto es, existe un vector $\xi \in \mathbb{R}^n$ diferente de cero tal que*:*

$$\inf\{\langle \xi, x\rangle : x \in \mathrm{A}_1\} \geq \sup\{\langle \xi, x\rangle : x \in \mathrm{A}_2\}$$

La demostración de puede encontrar en Nonlinear Programming de Mokhtar S. Bazaraa y otros. (2006, Pag. 60)

2.3. ANÁLISIS CONVEXO

En problemas de optimización, se consideraran funciones que toman valores en la recta real extendida, esto es:

$$[-\infty, +\infty] = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\} = \overline{\mathbb{R}}$$

Las operaciones usuales en aritmética se extienden de forma natural para los elementos de $[-\infty, +\infty]$.

En particular:

El límite inferior de una sucesión $\{x_n\}_{n\in\mathbb{N}}$ en $[-\infty, +\infty]$ está definido como:

$$lím \inf x_n = \sup_{n\in\mathbb{N}} \inf_{\substack{m\in\mathbb{N} \\ m\geq n}} x_m$$

y el límite superior

$$lím \sup x_n = \inf_{n\in\mathbb{N}} \sup_{\substack{m\in\mathbb{N} \\ m\geq n}} x_m$$

Es claro que: $lim\,inf\,x_n \leq lim\,sup\,x_n$.

2.3.1. FUNCIONES CONVEXAS

DEFINICIÓN 2.3.1.- Dados un conjunto convexo, no vacío, A de $\mathbb{R}^n$ y una función $f: A \to \overline{\mathbb{R}}$. Se definen los siguientes conjuntos:

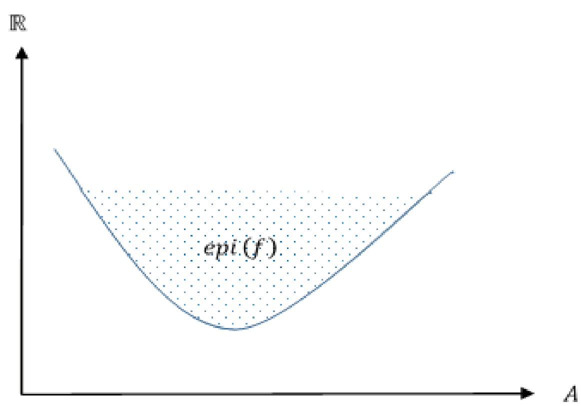
i. $dom\,(f) = \{x \in A / f(x) < +\infty\}$, llamado dominio esencial de f.

ii. $gráf(f) = \{(x, y) \in \mathbb{R}^{n+1} / f(x) = y\}$, llamado grafico de f.

iii. El epígrafo (epigrafía) de f, denotado por $epi(f)$, es el conjunto $\{(x, y) \in A \times \mathbb{R} / f(x) \leq y\}$; es decir:

$$epi(f) = \{(x, y) \in A \times \mathbb{R} / f(x) \leq y\} \subset \mathbb{R}^{n+1}$$

FIGURA 2.3. 1: Epigrafía de una función



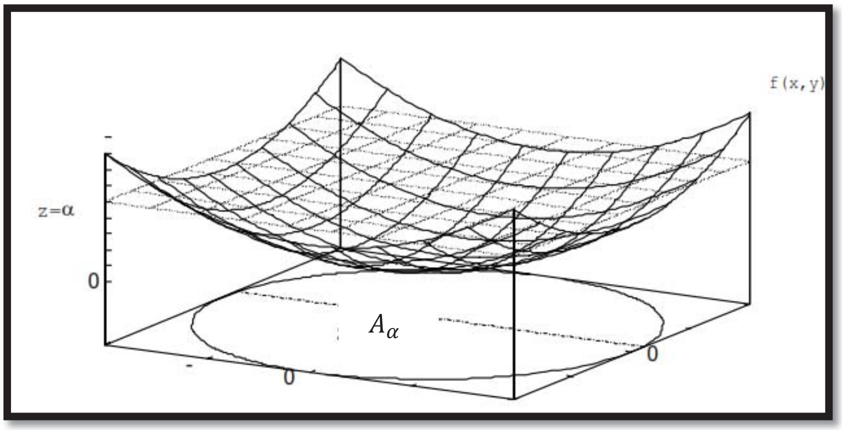
Fuente: (Jhan, 1996)

iv. La hipografía de f, denotado por $hyp\,(f)$, es el conjunto $\{(x, y) \in A \times \mathbb{R} / f(x) \geq y\}$; esto es:

$$hyp\,(f) = \{(x, y) \in A \times \mathbb{R} / f(x) \geq y\} \subset \mathbb{R}^{n+1}$$

v. El conjunto $A_\alpha = \{x \in A / f(x) \leq \alpha\}$, se llama conjunto de nivel de f de altura α $(\alpha \in \mathbb{R})$.

FIGURA 2.3.2: Conjunto de nivel de una función convexa



Fuente: (http://programacionlinealunadibague.blogspot.com/2011/05/funciones-convexas-y-concavas.html, 2011)

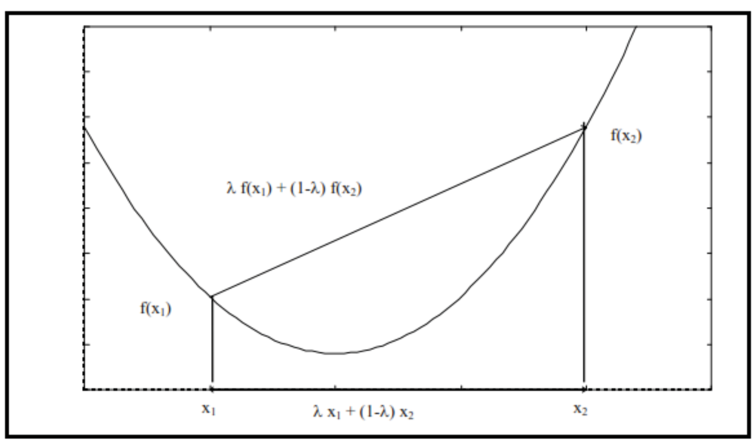
DEFINICIÓN 2.3.2. Sea A conjunto convexo no vacío en $\mathbb{R}^n$, y sea

$f: A \to (-\infty; +\infty]$ una función. Se dice que f es una función convexa si

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda) f(x_2) \qquad (2.3.1)$$

Para todo $x_1, x_2 \in A$ y $\lambda \in]0,1[$

FIGURA 2.3.3: Gráfica de una función convexa



Fuente: (http://programacionlinealunadibague.blogspot.com/2011/05/funciones-convexas-y-concavas.html, 2011)

DEFINICIÓN 2.3.3.- Sea A un subconjunto convexo no vacío de $\mathbb{R}^n$, y sea

$f: A \to (-\infty; +\infty]$ una función. Se dice que f es una función estrictamente convexa si

$$f[\lambda x_1 + (1-\lambda)x_2] < \lambda f(x_1) + (1-\lambda)f(x_2) \qquad \forall\lambda \in]0,1[\ , \ \ \forall x_1, x_2 \in A$$

con $x_1 \neq x_2$.

DEFINICIÓN 2.3.4.- Sea A conjunto convexo no vacío en $\mathbb{R}^n$, y sea

$f: A \to (-\infty; +\infty]$ una función. Se dice que f es una función cóncava si:

$$f[\lambda x_1 + (1-\lambda)x_2] \geq \lambda f(x_1) + (1-\lambda)f(x_2) \qquad \forall\lambda \in]0,1[\ , \ \ \forall x_1, x_2 \in A$$

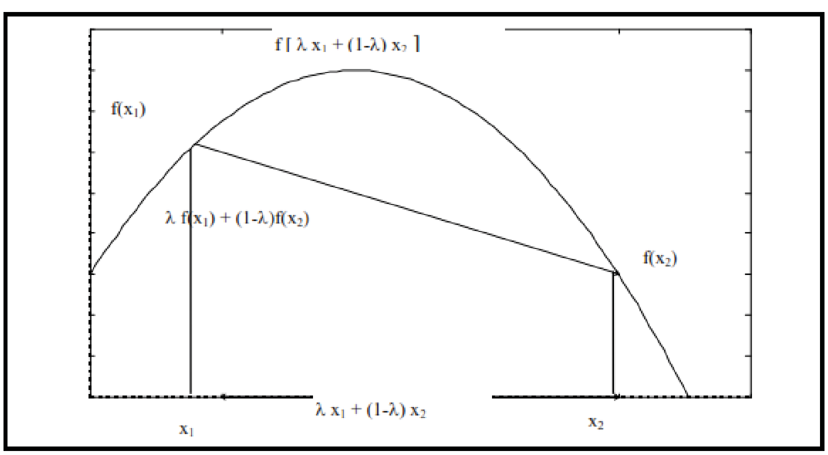
DEFINICIÓN 2.3.5.- Sea A conjunto convexo no vacío en $\mathbb{R}^n$, y sea

$f: A \to (-\infty; +\infty]$ una función. Se dice que f es una función estrictamente cóncava si:

$$f[\lambda x_1 + (1-\lambda)x_2] > \lambda f(x_1) + (1-\lambda)f(x_2) \qquad \forall\lambda \in [0,1]\ y\ \forall x_1, x_2 \in A$$

con $x_1 \neq x_2$.

FIGURA 2.3.4: Gráfica de una función cóncava



Fuente: (http://programacionlinealunadibague.blogspot.com/2011/05/funciones-convexas-y-concavas.html, 2011)

Observemos que, las definiciones dadas anteriormente no exigen necesariamente que la función sea continua y/o diferenciable. Así tenemos, lo siguiente:

$$f: \mathbb{R} \to \mathbb{R}$$

$$x \to f(x) = x^2$$

y

$$g: \mathbb{R} \to \mathbb{R}_0^+$$

$$x \to f(x) = |x|$$

Son funciones convexas y continuas. Lo que no ocurre en el caso de la diferenciabilidad, f es diferenciable para todo x de su dominio y g no es diferenciable en $x = 0$.

PROPOSICIÓN 2.3.1.- Sean A un subconjunto convexo, no vacío de $\mathbb{R}^n$, y sea $f: A \to (-\infty; +\infty]$ una función. Si f es una función convexa, entonces la función $-f$ es una función cóncava.

DEMOSTRACIÓN.

Sea f es una función convexa en A, por definición $\forall \lambda \in]0,1[, \quad \forall x_1, x_2 \in A$, se tiene:

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

$$-f\ [\lambda x_1 + (1-\lambda)x_2] \geq -\lambda f(x_1) - (1-\lambda)f(x_2)$$

$$(-f)\ [\ \lambda x_1 + (1-\lambda)x_2\ ] \geq \lambda(-f)(x_1) + (1-\lambda)(-f)(x_2)$$

Por consiguiente $-f$ es una función cóncava.

TEOREMA 2.3.1. Sean A un conjunto convexo no vacío en $\mathbb{R}^n$ y $f: A \to (-\infty; +\infty]$ una función convexa. Entonces el conjunto de nivel a la altura α, $\alpha \in \mathbb{R}$

$$A_\alpha = \{x \in A / f(x) \leq \alpha\}$$

es un conjunto convexo.

DEMOSTRACIÓN

Sean $x_1, x_2 \in A_\alpha$, luego:

$$f(x_1) \leq \alpha$$

$$f(x_2) \leq \alpha$$

y sea $f: A \to (-\infty; +\infty]$ una función convexa, por definición se tiene:

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

dado que: $f(x_1) \leq \alpha$ y $f(x_2) \leq \alpha$, entonces:

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda(x_1) + (1-\lambda)f(x_2)$$

$$\leq \lambda\alpha + (1-\lambda)\alpha = \alpha$$

con lo que

$$\lambda x_1 + (1-\lambda)x_2 \in A_\alpha$$

Por consiguiente, A_α es un conjunto convexo.

DEFINICIÓN 2.3.6.- Sea A un subconjunto no vacío de $\mathbb{R}^n$*.* La función $f: A \to \overline{\mathbb{R}}$ es propia si $-\infty \notin f(A)\ y\ \ dom\ (f) \neq \emptyset$.

DEFINICIÓN 2.3.7.- Sean A un conjunto en $\mathbb{R}^n$ y $f\ : A \to \overline{\mathbb{R}}$ una función propia.

i. El ínfimo de f sobre A, denotado por $inf_A(f)$. Diremos que f alcanza su mínimo sobre A, si existe $\bar{x}$ en A tal que $f(\bar{x}) = inf_A(f)$. Esto es:

$$f(\bar{x}) = \min_{x \in A} f(x)$$

ii. El supremo de f sobre A, denotado por $sup_A(f)$. Diremos que f alcanza su supremo sobre A, si existe $\bar{x}$ en A tal que $f(\bar{x}) = sup_A(f)$. Esto es:

$$f(\bar{x}) = \max_{x \in C} f(x)$$

PROPOSICIÓN 2.3. 2.- Sean A un conjunto convexo, no vacío en $\mathbb{R}^n$ y

$f : A \to (-\infty, +\infty]$. Entonces, f es convexa si y solamente si, el conjunto $epi\ (f)$ es convexo.

DEMOSTRACIÓN.-

Primero demostremos la condición necesaria:

(⇒)

Supóngase que $f : A \to (-\infty, +\infty]$ es una función convexa. Entonces:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2), \quad \forall \lambda \in]0,1[\qquad (2.3.2)$$

Sean (x_1, y_1) , $(x_2, y_2) \in epi\ (f)$

$$f(x_1) \leq y_1 \qquad \wedge \qquad f(x_2) \leq y_2 \qquad (2.3.3)$$

De (2.3.2) y (2.3.3), se tiene:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda y_1 + (1-\lambda)y_2 \qquad (2.3.4)$$

Debemos demostrar:

$$\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) \in epi\, f, \quad \forall\lambda \in]0,1[$$

En efecto:

$$\lambda(x_1, y_1) + (1-\lambda)\,(x_2, y_2) = (\lambda x_1, \lambda y_1) + \big((1-\lambda)\,x_2, (1-\lambda)\,y_2\big)$$

$$= (\lambda x_1 + (1-\lambda)\,x_2, \lambda y_1 + (1-\lambda)\,y_2)$$

Como A es un conjunto convexo $\lambda x_1 + (1-\lambda)x_2 \in A$ y $\lambda y_1 + (1-\lambda)y_2 \in \mathbb{R}$, entonces:

$$(x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \in A \times \mathbb{R} \qquad (2.3.5)$$

De (2.3.4) y (2.3.5), se tiene que:

$$\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) \in epi\, f, \quad \forall\lambda \in]0,1[$$

Finalmente, mostremos la condición suficiente:

($\Longleftarrow$)

Asúmase que el conjunto $epi\ (f)$ es convexo. Luego para $(x_1, f(x_1))$, $(x_2, f(x_2))$ puntos del epígrafo, se tiene:

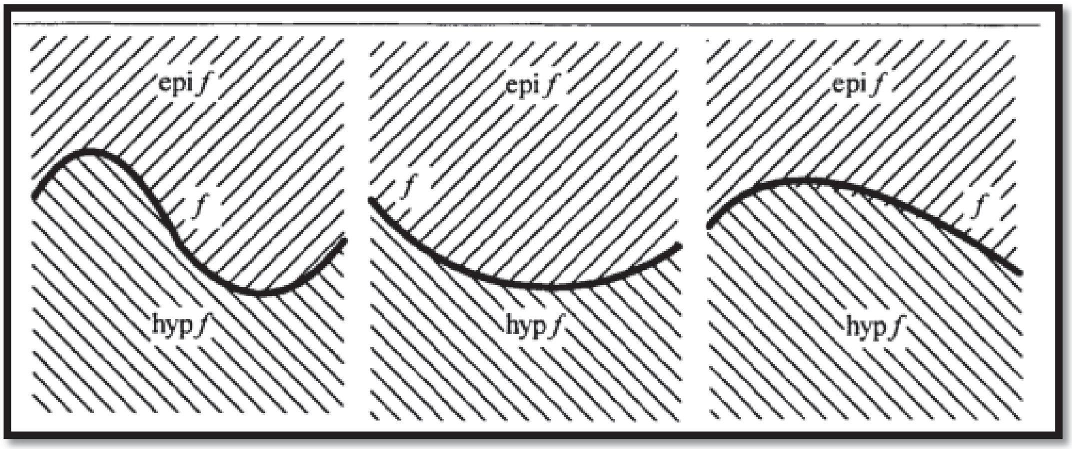
$$\big(x_1 + (1-\lambda)x_2, \lambda f(x_1) + (1-\lambda)f(x2)\big) \in\ epi\, f$$

Por definición de epigrafo:

$$f(x_1 + (1-\lambda)x_2)\ \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

para cada $\lambda \in]0,1[$. Por consiguiente, f es una función convexa.

FIGURA 2.3. 5: Epigrafia e Hipografia



Fuente: Nonlinear Programming (M. Bazzara, 2006)

Consecuentemente, podemos afirmar que: las funciones convexas son continuas en el interior de su dominio.

2.3.2. CONTINUIDAD DE FUNCIONES CONVEXAS

Las funciones convexas y cóncavas son continuas en el interior de su dominio.

TEOREMA 2.3.2.- Sean A un conjunto convexo, no nulo, en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa. Entonces f es continua en el interior de A.

DEMOSTRACIÓN (M. Bazzara, 2006)

Sea $\bar{x} \in int\ (A)$.

Existe un $\delta' > 0$, tal que $\|x - \bar{x}\| \leq \delta'$ con $x \in A$.

Por otro lado, sea

$$\theta = \max_{1 \leq i \leq n} [f(\bar{x} + \delta' e_i) - f(\bar{x}), f(\bar{x} - \delta' e_i) - f(\bar{x})]$$

Donde: e_i: es el vector canónico i-ésimo. Notemos que $0 \leq \theta < \infty$

Tomando:

$$\delta = \min\left(\frac{\delta'}{n}, \frac{\varepsilon\delta'}{n\theta}\right)$$

Podemos escoger un x con $\|x - \bar{x}\| < \delta$.

Si $x_i - \bar{x}_i \geq 0$ con $z_i = \delta' e_i$, de otra manera $z_i = -\delta' e_i$. Entonces:

$$x - \bar{x} = \sum_{i=1}^{n} \alpha_i z_i, \quad \alpha_i \geq 0$$

Además

$$\|x - \bar{x}\| = \delta' \left(\sum_{i=1}^{n} {\alpha_i}^2\right)^{\frac{1}{2}}$$

Y como $\|x - \bar{x}\| \leq \delta$ se sigue que $\alpha_i \leq \frac{1}{n}$; $\quad i = \overline{1, n}$.

Por la convexidad de f y desde $0 \leq n\alpha_i \leq 1$, se tiene:

$$f(x) = f\left(\bar{x} + \sum_{i=1}^{n} \alpha_i z_i\right) = f\left(\frac{1}{n}\sum_{i=1}^{n}(\bar{x} + n\alpha_i z_i)\right)$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} f(\bar{x} + n\alpha_i z_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n} f[(1 - n\alpha_i)\bar{x} + n\alpha_i(\bar{x} + z_i)]$$

$$\leq \frac{1}{n}\sum_{i=1}^{n}[(1 - n\alpha_i)f(\bar{x}) + n\alpha_i f(\bar{x} + z_i)]$$

Por consiguiente:

$$f(x) - f(\bar{x}) \leq \sum_{i=1}^{n} \alpha_i [f(\bar{x} + z_i) - f(\bar{x})]$$

Con $(\bar{x} + z_i) - f(\bar{x}) \leq 0, \ \ \forall i = \overline{1, n}$ y $\alpha_i \geq 0$.

Se sigue que:

$$f(x) - f(\bar{x}) \leq \theta \sum_{i=1}^{n} \alpha_i$$

Y como $\alpha_i \leq \frac{\varepsilon}{n\theta}$, se tiene:

$$f(x) - f(\bar{x}) \leq \varepsilon \tag{2.3.6}$$

Hasta aquí hemos demostrado que:

$$\|x - \bar{x}\| < \delta \text{ implica } f(x) - f(\bar{x}) \leq \varepsilon$$

Para demostrar la prueba necesitamos demostrar la semicontinuidad inferior de f en $\bar{x}$.

Sea $y = 2\bar{x} - x$ entonces $\|y - \bar{x}\| < \delta$, de donde:

$$f(y) - f(x) \leq \varepsilon \tag{2.3.7}$$

Pero $\bar{x} = \frac{1}{2}y + \frac{1}{2}x$ y como f es convexa, se tiene:

$$f(\bar{x}) \leq \frac{1}{2}f(y) + \frac{1}{2}f(x) \tag{2.3.8}$$

De (2.3.7) y (2.3.8), se tiene:

$$f(\bar{x}) - f(x) \leq \varepsilon \tag{2.3.9}$$

De (2.3.6) y (2.3.9), se tiene:

$$-\varepsilon \leq f(x) - f(\bar{x}) \leq \varepsilon$$

Es decir: $|f(x) - f(\bar{x})| \leq \varepsilon$.

TEOREMA 2.3.3.- Sean A un conjunto convexo no acotado en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa continua, cuyo conjunto de nivel A_α es acotado. Entonces f tiene un mínimo global.

La demostración de este resultado se puede encontrar en (Persson, 2006)

Notemos que todos los conjuntos de nivel A_α de f son cerrados y acotados, es decir, conjuntos compactos en $\mathbb{R}^n$. Entonces cada sucesión convergente, posee una subsucesión convergente, de donde inmediatamente existe un mínimo global.

PROPOSICIÓN 2.3.3.- Sea A un conjunto convexo de $\mathbb{R}^n$. Un conjunto de nivel A_α es acotado si y solo si:

$$\lim_{\|x\| \to \infty} \inf \frac{f(x)}{\|x\|} > 0$$

DEMOSTRACIÓN.- *(Persson, 2006)*

($\Rightarrow$) Es obvio.

($\Leftarrow$)

Por reducción al absurdo

Sea $(x_k)_k$ una sucesión en A tal que:

$$\|x_k\| \to \infty \qquad \text{y} \qquad f(x_k) \leq \frac{\|x_k\|}{k}$$

Lo que contradice que los conjuntos de nivel sean acotados, pues: $\frac{\|x_k\|}{k} \to \infty$

Dado que $x \in A$, la sucesión $x + \frac{k}{\|x_k\|}(x_k - x)$ no es acotada para algún conjunto de nivel $A_{f(x)+\varepsilon}$; $\varepsilon > 0$.

DEFINICIÓN 2.3.8.- Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. La función f es coerciva si

$$\lim_{\|x\| \to \infty} f(x) = +\infty$$

LEMA 2.3.1.- Sean A un conjunto convexo, abierto, no acotado en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa. Entonces f es localmente acotada.

DEMOSTRACIÓN (Persson, 2006)

Sea un punto fijo a en A. Escojamos un cubo K en A con centro en a y vértices $v_1, \dots, v_{2^n}$. Cada x que escojamos en K es una combinación convexa de sus vértices.

Luego:

$$f(x) = f\left(\sum_{k=1}^{2^n} \lambda_k \, v_k\right) \le M = \sup_{1 \le k \le 2^n} f(v_k)$$

Puesto que K es simético y dado que $x \in int\ (K)$, existe un $y \in K$, tal que:

$$a = (x + y)/2$$

y como f es una función convexa, se tiene:

$$2f(a) - M \le 2f(a) - f(y) \ \le f(x) \le M$$

Por consiguiente, f es acotada en K.

PROPOSICIÓN 2.3.4.- Sean A un conjunto convexo abierto, no vacío, en $\mathbb{R}^n$ y $f: A \to (-\infty, +\infty]$ una función convexa. Si f es localmente acotada en A entonces es localmente Lipschitziana en A. Particularmente f es continua en el interior de A.

DEMOSTRACIÓN (Persson, 2006)

Sea A un conjunto convexo abierto, no vacío, en $\mathbb{R}^n$.

Sean $a \in A$ y $B_{2r}(a) \subset A$ una vecindad de a.

Sean x e y en B_r con $x \neq y$. Sea

$z = y + \left(\frac{r}{\alpha}\right)(y - x)$

Con $\alpha = ||y - x||$.

Obviamente $z \in B_{2r}$. Despejando y se tiene:

$$y = \frac{r}{r+\alpha}x + \frac{\alpha}{r+\alpha}z$$

Y como f es convexa tenemos:

$$f(y) \leq \frac{r}{r+\alpha}f(x) + \frac{\alpha}{r+\alpha}f(z) .$$

Entonces

$$f(y) - f(x) \leq \frac{\alpha}{r+\alpha}(f(z) - f(x))$$

$$\leq \frac{\alpha}{r}(f(z) - f(x)) \leq \frac{2M}{r}||y - x||$$

Reemplazando y por x, se completa la prueba. ■

DEFINICIÓN 2.3.9.- Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ se dice que es inferiormente semicontinua en x, si:

$$f(x) \leq \liminf_{i\to\infty} f(x_i)$$

donde $(x_i)_{i\in\mathbb{N}}$ es cualquier sucesión convergente a x en $\mathbb{R}^n$.

DEFINICIÓN 2.3.10.- Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ se dice que es superiormente semicontinua en x, si:

$$f(x) \geq \lim_{n\to\infty} \sup f(x_n)$$

donde $(x_i)_{i\in\mathbb{N}}$ es cualquier sucesión convergente a x en $\mathbb{R}^n$.

Observemos que, si f es inferiormente semicontinua y superiormente semicontinua en x entonces f es continua en x.

TEOREMA 2.3.4.- Sea $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ una función. Entonces las siguientes proposiciones son equivalentes:

1. f es semicontinua en todo $\mathbb{R}^n$.
2. A_α es cerrado para todo $\alpha \in \mathbb{R}$.
3. $epi\ f$ es un conjunto cerrado en $\mathbb{R}^{n+1}$

La demostración de este resultado se encontrar en *(Rockafellar, 1970)* página 53.

Consecuentemente, las funciones semicontinuas se les llaman funciones cerradas.

2.4. DERIVADAS DIRECCIONALES DE FUNCIONES CONVEXAS

DEFINICIÓN 2.4.1.- Sean A un conjunto convexo, no vacío, en $\mathbb{R}^n$ y $f\colon A \to \mathbb{R}$ una función. Sean $\bar{x} \in A$ y d un vector no nulo de$\mathbb{R}^n$, tal que $\bar{x} + td \in A$ con $t > 0$ suficientemente pequeño. La derivada direccional de f en $\bar{x}$ en la dirección d, denotada por $f'_+(\bar{x}; d)$, es:

$$f'_+(\bar{x}; d) = \lim_{t\to 0^+} \frac{f(\bar{x} + td) - f(\bar{x})}{t}; \bar{x} \in A$$

Siempre que limite existe.

LEMA 2.4. 1.- Sean $f:\mathbb{R}^n \to \mathbb{R}$ una función convexa, $\bar{x} \in \mathbb{R}^n$ y d una dirección, no nula, de $\mathbb{R}^n$. Entonces la derivada direccional $f'_+(\bar{x};d)$ existe.

DEMOSTRACIÓN (Nava Manzo, 2015)

Sean t_1 y t_2 escalares positivos de $\mathbb{R}$, con $t_1 > t_2$ y $\bar{x} \in \mathbb{R}^n$. Además $\bar{x} + t_2 d$ esta en el dominio de f.

Por hipótesis, f es convexa, entonces:

$$f(\bar{x}+t_1 d) = f\left[\frac{t_1}{t_2}(\bar{x}+t_2 d) + \left(1-\frac{t_1}{t_2}\right)x\right] \leq \frac{t_1}{t_2} f(\bar{x}+t_2 d) + \left(1-\frac{t_1}{t_2}\right) f(\bar{x})$$

$$\frac{f(\bar{x}+t_1 d)-f(\bar{x})}{t_1} \leq \frac{f(x+t_2 d)-f(\bar{x})}{t_2}$$

y por lo tanto, el cociente:

$$\frac{f(\bar{x}+\mathrm{t}d)-f(\bar{x})}{t}$$

es monótona decreciente para $t \to 0^+$.

Puesto que, $t \geq 0$ es suficientemente pequeño y por la convexidad de f, obtenemos lo siguiente:

$$f(\bar{x}) = f\left[\frac{t}{1+t}(\bar{x}-d) + \frac{1}{1+t}(\bar{x}+td)\right] \leq \frac{t}{1+t}(\bar{x}-d) + \frac{1}{1+t}(\bar{x}+td)$$

$$\frac{f(\bar{x}+td)-f(\bar{x})}{t} \geq f(\bar{x}) - f(\bar{x}-d)$$

Por lo tanto, dado que el cociente anterior está acotado inferiormente se tiene que:

$$\lim_{t\to 0^+} \frac{f(\bar{x}+td)-f(\bar{x})}{t} = \underset{t>0}{ínf} \frac{f(\bar{x}+td)-f(\bar{x})}{t}$$

2.5. SUBDIFERENCIABILIDAD PARA UNA FUNCIÓN CONVEXA

DEFINICIÓN 2.5.1.- Sean A un conjunto convexo, no vacío, de $\mathbb{R}^n$ y $f: A \to (-\infty, +\infty]$ una función convexa. El subgradiente de f en x es ξ, tal que:

$$f(y) \geq f(x) + \langle y - x, \xi \rangle \ , \forall y \in A$$

El subgradiente ξ define un hiperplano de soporte para el conjunto $epi\ (f)$ en un punto frontera $(x, f(x))$. Esto es, si H es un hiperplano definido por la función afín $h(y) = f(x) + \langle \xi, y - x \rangle$ entonces $epi\ (f) \subseteq epi\ (h)$.

El subdiferencial de f en x es el conjunto:

$$\partial f(x) \coloneqq \{\xi / f(y) \geq f(x) + \langle \xi, y - x \rangle , \forall y \in A\} \qquad (2.5.1)$$

PROPOSICIÓN 2.5.1.- Sean A un conjunto convexo, no vacío, de $\mathbb{R}^n$ y $f: A \to (-\infty, +\infty]$ una función convexa. El subdiferencial de f en x es cerrado y convexo.

DEMOSTRACIÓN

Sean $\xi_1, \xi_2 \in \partial f(x)$ por definición, se tiene:

$$f(y) \geq f(x) + \langle \xi_1, y - x \rangle \ , \forall y \in A$$

$$f(y) \geq f(x) + \langle \xi_2, y - x \rangle , \forall y \in A$$

Sean $y \in A$ y $d_1, d_2 \in]0,1[$ tales que $d_2 = 1 - d_1$

$$f(x) + \langle d_1\xi_1 + d_2\xi_2, y - x \rangle$$

$$= d_1 f(x) + (1 - d_1) f(x) + d_1 \langle \xi_1, y - x \rangle + d_2 \langle \xi_2, y - x \rangle$$

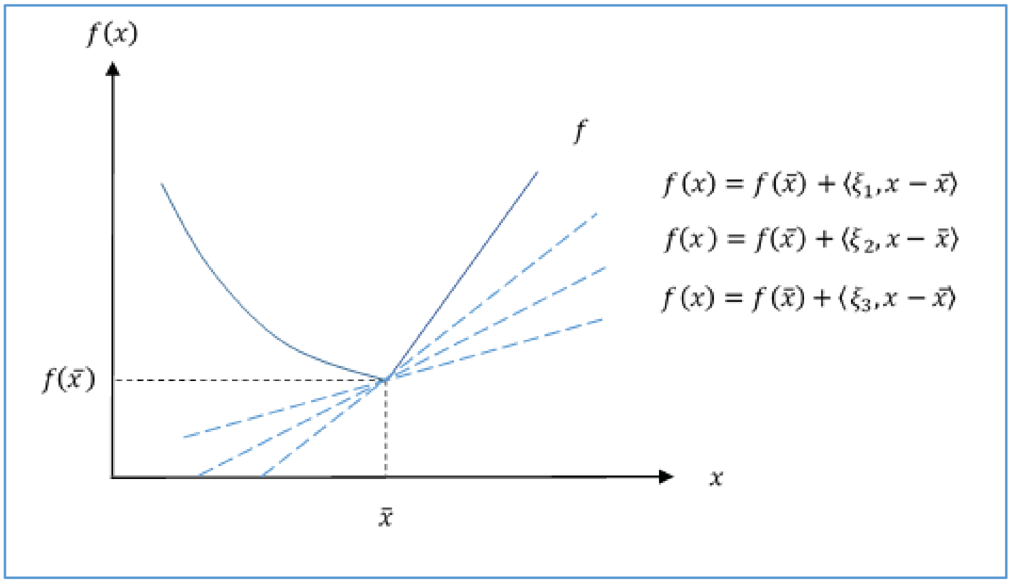
$$= d_1 f(x) + d_1 \langle \xi_1, y - x \rangle + d_2 f(x) + d_2 \langle \xi_2, y - x \rangle$$

$$\leq d_1 f(y) + d_2 f(y) = f(y)$$

Por lo tanto, el conjunto $\partial f(x)$ es convexo.

$\partial f(x)$ es cerrado, puesto que el producto interior es una función continua.

FIGURA 2.5.1: Subgradientes de un función convexa



Fuente: (Jhan, 1996)

LEMA 2.5.1.- Sean A un conjunto convexo, no vacio, de $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa propia. Sea $x \in int(A)$ y $B_r(x) \subseteq int(A)$ una vecindad de x, con $r > 0$. Si f es Lipschitziana, con constante L, en $B_r(x)$ entonces $\partial f(x)$ es acotado.

DEMOSTRACIÓN (Nava Manzo, 2015)

Sean $\xi \in \partial f(x)$ un vector no nulo de $\mathbb{R}^n$ y $x + \gamma\xi \in B_r(x)$ con $\gamma > 0$. Como f es convexa, se tiene:

$$f(x + \gamma\xi) \geq f(x) + \langle \xi, \gamma\xi \rangle = f(x) + \gamma\|\xi\|^2$$

Como f es Lipschitziana, se deduce que:

$$\gamma L\|\xi\| \geq f(x + \gamma\xi) - f(x) \geq \gamma\|\xi\|^2$$

De donde

$$\|\xi\| \leq L$$

Por consiguiente $\partial f(x)$ s acotado.

Consecuentemente, como $\partial f(x)$ es cerrado y acotado, entonces $\partial f(x)$ es compacto.

TEOREMA 2.5.1.- Sean A un conjunto convexo, no vacío, en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa. Entonces para $\bar{x} \in int\ A$, existe un vector ξ, tal que el hiperplano

$$H = \{(x, y)/y = f(\bar{x}) + \langle \xi, x - \bar{x} \rangle \}$$

Soporta al $epi\ (f)$ en $(\bar{x}, f(\bar{x}))$.

En particular:

$$f(x) \geq f(\bar{x}) + \langle \xi, x - \bar{x} \rangle, \quad \forall x \in A$$

Esto es: ξ es un subgradiente de f en $\bar{x}$.

DEMOSTRACIÓN (M. Bazzara, 2006)

Sea $f: A \to \mathbb{R}$ una función convexa. Entonces su $epi\ (f)$ es un conjunto convexo. Notemos que $(\bar{x}, f(\bar{x}))$ pertenece a la frontera del $epi\ (f)$.

Por el teorema 2.2.1, existe un vector, no nulo,$(\xi_0, \mu) \in \mathbb{R}^n \times \mathbb{R}$ tal que:

$$\langle \xi_0, x - \overline{x} \rangle + \mu(y - f(\overline{x})) \leq 0\ , \quad \forall (x, y) \in epi\ (f)$$

Donde μ no es positivo, de otra forma la desigualdad no sería verdadera para $y \to \infty$.

Si $\mu = 0$, entonces:

$$\langle \xi_0, x - \overline{x} \rangle \leq 0, \ \ \forall x \in A$$

Como $\overline{x} \in int\ (A)$, existe un escalar $\alpha > 0$, tal que:

$$\overline{x} + \alpha\xi_0 \in A$$

Entonces:

$$\langle \xi_0, \overline{x} + \alpha\xi_0 - \overline{x} \rangle \leq 0$$

$$\langle \xi_0, \alpha\xi_0 \rangle \leq 0$$

$$\alpha\langle \xi_0, \xi_0 \rangle \leq 0$$

$$\alpha\|\xi_0\|^2 \leq 0$$

De donde $\xi_0 = 0$, entonce $(\xi_0, \mu) = (0,0)$, lo que es una contradicción.

La única posibilidad es que μ sea estrictamente negativo $(\mu < 0)$. Sea $\xi = \frac{\xi_0}{|\mu|}$,

entonces

$$\langle \xi_0, x - \overline{x} \rangle + \mu\big(y - f(\overline{x})\big) \leq 0, \quad \forall (x, y) \in epi\ (f)$$

$$\langle \frac{\xi_0}{|\mu|}, x - \overline{x} \rangle - \big(y - f(\overline{x})\big) \leq 0$$

$$\langle \xi, x - \overline{x} \rangle - y + f(\overline{x}) \ \leq 0$$

$$\langle \xi, x - \overline{x} \rangle + f(\overline{x}) \ \leq y$$

En particular, tomando $y = f(x)$ obtenemos:

$$f(x) \geq f(\overline{x}) + \langle \xi, x - \overline{x} \rangle\ , \ \ \forall x \in A$$

Es decir, existe un hiperplano de soporte H del $epi(f)$ en el punto $\big(\overline{x}, f(\overline{x})\big)$.

Como consecuencia de este resultado, se tiene que toda función convexa tiene al menos un subgradiente en cada uno de los puntos del interior de su dominio.

COROLARIO 2.5.1.- Sean A un conjunto convexo no vacio en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función estrictamente convexa. Si $\overline{x} \in intA$ entonces existe un vector ξ tal que:

$$f(x) > f(\overline{x}) + \langle\, \xi, x - \overline{x}\rangle \quad , \forall x \in A, x \neq \overline{x}$$

DEMOSTRACION (M. Bazzara, 2006)

Por el teorema anterior, existe un vector ξ tal que:

$$f(x) \geq f(\overline{x}) + \langle \xi, x - \overline{x}\rangle \,, \forall x \in A \tag{2.5.2}$$

Usando el método de reducción al absurdo. Supongamos que existe un $x_0 \neq \overline{x}$ tal que $f(x_0) = f(\overline{x}) + \langle \xi, x_0 - \overline{x}\rangle$. Como $\mathrm{f}: \mathrm{A} \to \mathbb{R}$ es estrictamente convexa, se tiene:

$$f(\lambda\overline{x} + (1-\lambda)x_0) < \lambda f(\overline{x}) + (1-\lambda)f(x_0), \quad \lambda \in]0{,}1[$$

$$< f(\overline{x}) + (1-\lambda)\langle \xi, x_0 - \overline{x}\rangle \tag{2.5.3}$$

Sea $x = \lambda\overline{x} + (1-\lambda)x_0$ en (2.5.2) se tiene:

$$f(\lambda\overline{x} + (1-\lambda)x_0) \geq f(\overline{x}) + (1-\lambda)\langle \xi, x_0 - \overline{x}\rangle$$

Lo que contradice (2.5.3)

PROPOSICIÓN 2.5.2.- Sea A un conjunto convexo de $\mathbb{R}^n$ y sea una función convexa $f: A \to \mathbb{R}$. Si para todo $x \in int\, A$ existe ξ, tal que:

$$f(x) \geq f(\bar{x}) + \langle\, \xi, x - \bar{x}\rangle \;, \qquad \forall x \in A$$

entonces f es convexa en $intA$.

DEMOSTRACION (M. Bazzara, 2006)

Sean x_1, x_2 en el interior de A y $\lambda \in]0{,}1[$. Como A es un conjunto convexo,

$$\lambda x_1 + (1-\lambda) \in A$$

Por hipótesis, existe un subgradiente ξ de f en $\lambda x_1 + (1-\lambda)x_2$. En particular, las siguientes desigualdades se cumplen:

$$f(x_1) \geq f(\lambda x_1 + (1-\lambda)x_2) + (1-\lambda)\langle \xi, x_1 - x_2 \rangle$$

$$f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2) + (1-\lambda)\langle \xi, x_2 - x_1 \rangle.$$

Multiplicando las desigualdades por λ y $(1-\lambda)$ respectivamente y sumándolas, se tiene:

$$\lambda f(x_1) + (1-\lambda)f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2)$$

Por consiguiente f es una función convexa .

LEMA 2.5.2.- Sea $(x_n)_{n\in\mathbb{N}}$ una sucesión convergente a x en A, con $x \in A$ y sea $(\xi_n)_{x_n}$ una sucesión, tal que: $\xi_n \in \partial f(x_n)$, para todo x_n. Entonces cualquier punto de acumulación de $(\xi_n)_{x_n}$ pertenece a $\partial f(x)$.

DEMOSTRACION (Fletcher, 2013)

Dado $y \in A$, por la definición de subgradiente se tiene:

$$f(y) \geq f(x_n) + \langle \xi_n, y - x_n \rangle.$$

Sea ξ un punto de acumulación de la sucesión $(\xi_n)_{x_n}$, entonces existe una subsucesión (ξ_{n_i}) tal que $\xi_{n_i} \to \xi$. Entonces $x_{n_i} \to x$ y como f es continua se tiene:

$$f(y) \geq f(x) + \langle \xi, y - x \rangle$$

para todo $y \in A$. Por lo tanto, $\xi \in \partial f(x)$.

TEOREMA 2.5. 2.- Sea $f: \mathbb{R}^n \to (-\infty, +\infty]$ una función convexa y cerrada. Si $x \in int(domf)$ entonces $f'(x;d) = max\{\langle \xi; d \rangle / \xi \in \partial f(x)\}$.

DEMOSTRACION (Nava Manzo, 2015)

Sea $\xi_n \in \partial f(x+\frac{1}{n}d)$, donde $d \in \mathbb{R}^n$ y $n \in \mathbb{N}$, tal que: $x+\frac{1}{n}d$ esta en $int(domf)$

Entonces

$$f(x) \geq f(x+\frac{1}{n}d) - \frac{1}{n}\langle \xi_n, d\rangle$$

Y

$$f(x+\frac{1}{n}d) \geq f(x) + \frac{1}{n}\langle \xi, d\rangle, \quad \forall \xi \in \partial f(x) .$$

Consecuentemente:

$$\langle \xi_n, d\rangle \geq \frac{f(x+\frac{1}{n}d) - f(x)}{\frac{1}{n}} \geq \max_{\xi}\langle \xi, d\rangle .$$

Se sabe que $\partial f(x+\frac{1}{n}d)$ es acotada en alguna vecindad de x. Existe una subsucesión $(\xi_{n_i})_{n_i \in \mathbb{N}}$ convergente a ξ, con $\xi \in \partial f(x)$. Por un resultado anterior y tomando el límite cuando $n_i \to \infty$, se tiene:

$$f'(x;d) = \max_{\xi}\langle \xi, d\rangle.$$

2.5.1. APROXIMACIÓN DEL SUBDIFERENCIAL A UNA FUNCION CONVEXA

En esta sección se define el conjunto $\partial_\varepsilon f(x)$. Se demuestra que $\partial_\varepsilon f(x)$ es un conjunto compacto**.**

DEFINICIÓN 2.5. 2.- Sea $\varepsilon > 0$, el ε-diferencial de una función convexa $f: \mathbb{R}^n \to (-\infty, +\infty]$ en un punto x, es el conjunto

$$\partial_\varepsilon f(x) \coloneqq \{\xi \in \mathbb{R}^n / f(y) \geq f(x) + \langle \xi, y - x\rangle, \forall y \in \mathbb{R}^n\}$$

Propiedades:

1. $\partial f(x) \subseteq \partial_\varepsilon f(x)$. Más aún, $\partial f(x) = \bigcup_{\varepsilon>0} \partial_\varepsilon f(x)$.

2. $\partial_\varepsilon f(x)$ es un conjunto convexo y cerrado.

LEMA 2.5.3.- Sean $f: \mathbb{R}^n \to \mathbb{R}$ una función, $x, x' \in dom(f)$ y $\xi' \in \partial f(x')$. Entonces:

$$x' \in \partial_\varepsilon f(x) \Leftrightarrow f(x') + \langle \xi', x' - x \rangle - \varepsilon.$$

DEMOSTRACIÓN (S. Boyd, 2006)

($\Rightarrow$)

Bastara reemplazar $y = x'$ en la definición de $\partial_\varepsilon f(x)$.

($\Leftarrow$)

Sea $\xi' \in \partial \mathrm{f}(\mathrm{x}')$.

$$f(y) \geq f(x') + \langle \xi', y - x' \rangle$$

$$= f(x) + \langle \xi', y - x \rangle + [f(x') - f(x) + \langle \xi', x - x' \rangle]$$

$$\geq f(x) + \langle \xi', y - x \rangle - \varepsilon$$

Por lo tanto $x' \in \partial_\varepsilon f(x)$.

En adelante, se denotara por $\mathcal{B}(\mathbb{R}^n)$ a la familia de conjuntos borelianos de $\mathbb{R}^n$.

DEFINICIÓN 2.5.3.- Sea $M: \mathbb{R}^n \to \mathcal{B}(\mathbb{R}^n)$ una función.

1. M es superiormente semicontinua, si para cualquier sucesión $(x_k, s_k) \to (x, s)$, tal que $s_k \in M(x_k)$ y $k \in \mathbb{N}$, se tiene que $s \in M(x)$.

2. M es inferiormente semicontinua, si dado $x \in M(x)$ y cualquier sucesión $x_k \to x$, existe una sucesión $\{s_k\}_{k \geq 1}$, tal que $s_k \in M(x_k)$ y

$$s_k \to s.$$

3. M es continua si es inferiormente semicontinua y superiormente semi continua.

LEMA 2.5. 4.- Sean A un conjunto convexo, no vacio, en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa propia. Sea $x \in int(A)$ y $\delta > 0$ tal que $B_\delta(x) \subseteq int(A)$. Si f es Lipschitziana, con constante L, en $B_\delta(x)$, entonces dado $\delta' < \delta$, $w \in B_\delta(x)$ y

$\xi \in \partial_\varepsilon f(w)$, se tiene que:

$$\|\xi\| \leq L + \frac{\epsilon}{\delta - \delta'}$$

DEMOSTRACION (Nava Manzo, 2015)

Sean ξ un vector no nulo y $z = w + (\delta - \delta')\frac{\xi}{\|\xi\|}$. Como $\xi \in \partial_\varepsilon f(w)$ y f es convexa

se tiene:

$$f(z) \geq f(w) + \langle \xi, z - w \rangle - \varepsilon.$$

Dado $z, w \in B_\delta(x)$ y como f es Lipschitziana, con constante L ,en $B_\delta(x)$, se tiene:

$$L\|z - w\| \geq f(z) - f(w) \geq \langle \xi, z - w \rangle - \varepsilon.$$

Se sabe que $\|z - w\| = \|(\delta - \delta')\frac{\xi}{\|\xi\|}\| = \delta - \delta'$ y por la desigualdad anterior

$$(\delta - \delta')L \geq \langle \xi, (\delta - \delta')\frac{\xi}{\|\xi\|} \rangle - \varepsilon = \|\xi\|(\delta - \delta') - \varepsilon.$$

Del anterior resultado se tiene que el conjunto $\partial_\varepsilon f(x)$ es acotado y por continuidad del producto interno, $\partial_\varepsilon f(x)$ es cerrado. Más aun $\partial_\varepsilon f(x)$ es un conjunto compacto**.**

TEOREMA 2.5.3.- Sea $f: \mathbb{R}^n \to \mathbb{R}$ una función convexa Lipschitziana con constante L en$\mathbb{R}$. Entonces, existe una constante $K > 0$ tal que para toda $x, x' \in \mathbb{R}^n$,

$\epsilon, \epsilon' > 0$, $\xi \in \partial_\varepsilon f(x)$, existe $\xi' \in \partial_{\varepsilon'} f(x')$, tal que:

$$\|\xi - \xi'\| \leq \frac{K}{\min\{\epsilon, \epsilon'\}}(\|x - x'\| + \|\epsilon - \epsilon'\|)$$

DEMOSTRACION (Nava Manzo, 2015)

Como $\partial_\varepsilon f(x)$ y $\partial_{\varepsilon'} f(x')$ son conjuntos convexos, será suficiente probar que

$|\max\{\langle \xi, d\rangle / \xi \in \partial_\varepsilon f(x)\} - \max\{\langle \xi', d\rangle / \xi' \in \partial_{\varepsilon'} f(x')\}| \leq \frac{K}{min\{\varepsilon, \varepsilon'\}}(\|x - x'\| + |\varepsilon - \varepsilon'|)$.

Sea d un vector unitario, se define la ε-derivada direccional de f en x en dirección d como:

$$f'_\varepsilon(x; d) = \inf_{t > 0} \frac{f(x+td) - f(x) + \varepsilon}{t}.$$

El cociente del lado derecho se denotará como:

$$q_\varepsilon(x, t, d) = \frac{f(x + td) - f(x) + \varepsilon}{t}.$$

Por el teorema anterior se tiene el siguiente resultado:

$$f'_\varepsilon(x; d) = \max\{\langle s, d\rangle \ / \ s \in \partial_\varepsilon f(x)\}$$

Luego, para cualquier $\nu > 0$, existe $t_\nu > 0$ tal que

$$q_\varepsilon(x, t_\nu, d) < f'_\varepsilon(x; d) + \nu$$

usando el lema anterior haciendo $\delta \to \infty$, tenemos que $f'_\varepsilon(x; d) \leq L$, y así $q_\varepsilon(x, t) \leq L + \nu$. Por otro lado

$$q_\varepsilon(x, t_\nu, d) = \frac{f(x + t_\nu d) - f(x) + \varepsilon}{t_\nu} \geq -L + \frac{\varepsilon}{t_\nu},$$

Entonces:

$$\frac{1}{t_\nu} \leq \frac{2L + \nu}{\varepsilon}$$

Así

$$f'_{\varepsilon'}(x'; d) - f'_\varepsilon(x; d) - \nu \leq q_{\varepsilon^0}(x', t_\nu, d) - q_\varepsilon(x; t_\nu, d)$$

$$= \frac{f(x' + t_\nu d) - f(x + t_\nu d) + f(x) - f(x') + \varepsilon' - \varepsilon}{t_\nu}$$

$$\leq \frac{2L\|x - x'\| + |\varepsilon - \varepsilon'|}{t_\nu}$$

$$\leq \frac{2L+\nu}{\varepsilon}(2L\|x - x'\| + |\varepsilon - \varepsilon'|) ,$$

como ν es arbitrario y se pueden intercambiar x y x' de la siguiente manera

$$|f'_{\varepsilon'}(x'; d) - f'_\varepsilon(x; d)| \leq \frac{2L}{m' \ln\{\varepsilon, \varepsilon\}}(2L\|x - x'\| + |\varepsilon - \varepsilon'|) .$$

TEOREMA 2.5.4.- Sea $f: \mathbb{R}^n \to \mathbb{R}$ una función convexa e inferiormente semi continua, entonces la correspondecia $M: \mathbb{R}^n \times \mathbb{R}_+ \to \mathcal{B}(\mathbb{R}^n)$ definida por:

$$(\mathrm{e}, \epsilon) \to \partial_\varepsilon \mathrm{f}(\mathrm{x})$$

es continua.

DEMOSTRACIÓN (Nava Manzo, 2015)

Del teorema anterior prueba que la aplicación M es inferiormente semicontinua para cualquier $\varepsilon > 0$ y cualquier $x \in \mathbb{R}^n$.

Para demostrar que M es superiormente semicontinua:

Sea la sucesión $\left(x^k, s^k, \varepsilon_k\right) \to (\bar{x}, \bar{s}, \bar{\varepsilon})$, para todo $y \in \mathbb{R}^n$, se tiene:

$$f(y) \geq f\left(x^k\right) + \langle s^k, y - x^k \rangle - \varepsilon_k$$

Dado que f es inferiormente continua, tomando el limite $k \to \infty$,

$$f(y) \geq f(\bar{x}) + \langle \bar{s}, y - \bar{x} \rangle - \bar{\varepsilon},\ \forall\ y \in \mathbb{R}^{\mathrm{n}}$$

En consecuencia, $\bar{s} \in \partial_{\bar{\varepsilon}} f(\bar{x})$ y M es superiormente semicontinua.

2.6. FUNCIONES CONVEXAS DIFERFENCIABLES

DEFINICIÓN 2.6.1.- Sean A un conjunto convexo, no vacío, en $\mathbb{R}^n$ y $f\colon A \to \mathbb{R}$ una función. Entonces f es diferenciable en $\bar{x} \in int\ A$, si existe $v \in \mathbb{R}^n$ tal que:

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} = 0\ \ , \forall x \in A$$

v es único, donde $v = \nabla f(\bar{x})$ es el vector gradiente.

$$\nabla f(\bar{x}) \coloneqq \left(\frac{\partial f(\bar{x})}{\partial x_1}, \dots, \frac{\partial f(\bar{x})}{\partial x_n} \right)$$

LEMA 2.6.1.- Sea A un conjunto convexo abierto, no vacío, en $\mathbb{R}^n$, y sea

$f\colon A \to \mathbb{R}$ convexa. Si f es diferenciable en $\bar{x} \in intA$ entonces $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

DEMOSTRACION (Nava Manzo, 2015)

Ya que f es convexa entonces $\partial f(\overline{x}) \neq \emptyset$.

Sean $\xi \in \partial f(\overline{x})$, $\boldsymbol{d} \in \mathbb{R}^n$ y $\lambda > 0$ suficientemente pequeño. Por definición de subgradiente, se tiene:

$$f(\overline{x} + \lambda d) > f(\overline{x}) + \lambda \langle \xi, d \rangle$$

$$\frac{f(\overline{x} + \lambda d) - f(\overline{x})}{\lambda} \geq \langle \xi, d \rangle$$

Luego, haciendo $\lambda \to 0$, el lado izquierdo converge dado que f es diferenciable. Por lo tanto:

$$\langle \nabla f(\overline{x}) \,, d\rangle \ \geq \ \langle s, d\rangle \,, \quad \text{para todo } \boldsymbol{d} \in \mathbb{R}^n.$$

En consecuencia, $\nabla f(\overline{x}) = \xi$.

TEOREMA 2.6.1.- Sean A un conjunto abiero y convexo en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ diferenciable en A. f es convexa, si y solo si, para cualquier $\bar{x} \in A$ se tiene:

$$f(x) \geq f(\bar{x}) + \langle \nabla f(\bar{x}) \,, x - \bar{x}\rangle$$

Para cada $x \in A$.

La demostración de puede encontrar en Nonlinear Programming de Mokhtar S. Bazaraa y otros. (2006, Pag. 51)

TEOREMA 2.6.2.- Sea A un conjunto abierto y convexo de $\mathbb{R}^n$, y sea $f: A \to \mathbb{R}$ diferenciable en A. Entonces f es convexa si y solo si, para cada $x_1, x_2 \in A$

$$\langle \nabla f(x_2) - \nabla f(x_1) \,, x_1 - x_2\rangle \geq 0$$

DEMOSTRACIÓN.

($\Longrightarrow$)

Sea $\mathrm{f}: \mathrm{A} \to \mathbb{R}$ una función convexa y sean $x_1, x_2 \in A$, entonces:

$$f(x_1) \ \geq f(x_2) + \langle \nabla f(x_2), x_1 - x_2\rangle$$

$$f(x_2) \geq f(x_1) + \langle \nabla f(x_1) \,, x_2 - x_1\rangle.$$

Sumando las dos desigualdades, obtenemos que:

$$f(x_1) + f(x_2) \geq f(x_2) + f(x_1) + \langle \nabla f(x_2), x_1 - x_2 \rangle + \langle \nabla f(x_1), x_2 - x_1 \rangle$$

$$\langle \nabla f(x_2) - \nabla f(x_1), x_2 - x_1 \rangle \geq 0.$$

(⇐)

Sean $x_1, x_2 \in A$. Por el teorema del valor medio, existe $\lambda \in]0,1[$, tal que:

$$f(x_2) - f(x_1) = \langle \nabla f(x), x_1 - x_2 \rangle \qquad (2.6.1)$$

donde $x = \lambda x_1 + (1-\lambda)x_2$.

Por hipótesis, se tiene:

$$\langle \nabla f(x) - \nabla f(x_1), x - x_1 \rangle \geq 0$$

esto implica que $\langle \nabla f(x), x_2 - x_1 \rangle \geq \langle \nabla f(x_1), x_2 - x_1 \rangle$.

Por (2.6.1),se tiene:

$$f(x_2) \geq f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle$$

por lo tanto f es una función convexa.

CAPÍTULO III: OPTIMIZACIÓN PARA FUNCIONES CONVEXAS NO DIFERENCIABLES

3.1. OPTIMIZACIÓN CONVEXA

Un problema de optimización matemática o simplemente problema de optimización es de la forma:

$$\min f_0(x)$$

$$\text{Sujeto a: } f_i(x) \leq b_i\ ,\ i = \overline{1,n}$$

Donde:

$x \in \mathbb{R}^n$ es la variable de optimización del problema.

$f_0: \mathbb{R}^n \to \mathbb{R}$ es la función objetivo

$f_i: \mathbb{R}^n \to \mathbb{R}$ son las funciones de restricciones

Un vector $\bar{x}$ es llamado óptimo o solución del problema, si para cualquier $z \in \mathbb{R}^n$ con $f_1(z) \leq b_1, f_2(z) \leq b_2, \dots, f_n(z) \leq b_n$, se tiene:

$$f_0(z) \geq f_0(\bar{x})$$

Por lo general consideramos familias o clases de problemas de optimización, caracterizados por la forma de la función objetivo o las funciones de la restricción.

Un problema de optimización se llama lineal, si su función objetivo y sus restricciones son funciones lineales. En caso que, algunas o todas las funciones no sean lineales, se considera que el problema es un problema de optimización no lineal.

Un problema de optimización convexa es aquel donde la función objetivo y las funciones de la restricción son funciones convexas.

Una generalización de los problemas de optimización, son los problemas de optimización convexa.

La optimización convexa trata de resolver el problema general de minimizar una función convexa, sobre un conjunto convexo.

Sea $f\colon \mathrm{A} \subset \mathbb{R}^n \to \mathbb{R}$ una función convexa. Entonces el problema P es definido como:

$$P = \min_{x \in A} f(x) \tag{2.7.1}$$

con A conjunto convexo en $\mathbb{R}^n$.

En general para resolver problemas de optimización convexa no existe una fórmula analítica, pero existen métodos efectivos para resolverlos. Los métodos relacionados al punto interior funcionan bien en la práctica.

DEFINICIÓN 3.1.1.- Sea $f\colon \mathbb{R}^n \to \mathbb{R}$ y consideremos el problema P. Un punto $x^* \in A$ es llamado solución óptima, mínimo global o solución factible, si:

$f(x) \geq f(x^*), \quad \forall x \in A.$

Si $x^* \in A$ y existe una vecindad $V(x^*)$, para cada $x \in A \cap V(x^*)$, entonces x^* es llamado mínimo local del problema. Por otro lado, si $x^* \in A$ es el único mínimo local en $A \cap V(x^*)$ para alguna vecindad de x^*, entonces x^* es llamado mínimo local aislado del problema.

TEOREMA 3.1.1.-. Sean A un subconjunto convexo, no vacío, de $\mathbb{R}^n$ y

$f: A \to \mathbb{R}$ convexa en A. Supongamos que $x^* \in A$ es un mínimo del problema P. Entonces x^* es un minimo global.

DEMOSTRACIÓN (M. Bazzara, 2006)

Dado que x^* es una solución local, existe una vecindad $V(x^*)$ de x^* tal que:

$$f(x) \geq f(x^*)$$

para cada $x \in A \cap V(x^*)$.

Supongamos, por el contrario que x^* no es un mínimo global del problema, es decir, existe $\hat{x} \in S$ tal que $f(\hat{x}) < f(x^*)$. Por la convexidad de f, para cada $\lambda \in [0,1]$ se tiene que:

$f(\lambda\hat{x} + (1-\lambda)x^*) \leq \lambda f(\hat{x}) + (1-\lambda)f(x^*) < \lambda f(x^*) + (1-\lambda)f(x^*) = f(x^*)$.

Así

$$f(\lambda\hat{x} + (1-\lambda)x^*) < f(x^*)$$

y se llega a una contradicción cuando λ es suficientemente pequeño. Por lo tanto, x^* es un mínimo global.

Un punto muy importante que se prueba fácilmente es que si f es convexa, $x^* \in A$ es un mínimo global del problema P si y solo si $0 \in \partial f(x^*)$. Más aún, si A es abierto y f es convexa y diferenciable en A, entonces $x^* \in A$ es un mínimo global del problema P si y solo si $\nabla f(x^*) = 0$.

TEOREMA 3.1.2. Sean A un conjunto convexo, no vacio, en $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa . El punto $x^* \in A$ es un minimo global del problema P si y solo si f tiene un subgradiente ξ en x^* tal que:

$$\langle \xi , x - x^* \rangle \geq 0 \text{ para todo } x \in A.$$

DEMOSTRACIÓN *(Nava Manzo, 2015)*

($\Longrightarrow$)

Supongamos que $\langle \xi, x - x^* \rangle \geq 0$ para todo $x \in A$, donde $\xi \in \partial f(x^*)$. Entonces, por hipótesis

$$f(x) \geq f(x^*) + \langle \xi, x - x^* \rangle \geq f(x^*)$$

para todo $x \in A$, por lo tanto x^* es una solución óptima del problema dado.

Para mostrar la necesidad, supongamos que x^* es una solución óptima del problema, se construyen dos conjuntos en $\mathbb{R}^{n+1}$ de la siguiente manera

$$\Gamma_1 = \{(x - x^*, y) : x \in \mathbb{R}^n, y > f(x) - f(x^*)\}$$

$$\Gamma_2 = \{(x - x^*, y) : x \in A, y \leq 0\}.$$

Es fácil verificar que tanto Γ_1 y Γ_2 son conjuntos convexos. Además, $\Gamma_1 \cap \Gamma_2 = \emptyset$ porque de otra manera existiría un punto (x, y) tal que

$$x \in A, 0 \geq y > f(x) - f(x^*)$$

(⇐)

La suposición de que x^* es un mínimo global del problema. Existe un hiperplano que separa Γ_1 y Γ_2; esto es, existe

un vector diferente de cero (ξ_0, μ) y un escalar α tal que:

$$\langle \xi_0, x - x^* \rangle + \mu y \leq \alpha, \forall x \in \mathbb{R}^n, y > f(x) - f(x^*) \tag{2.7.2}$$

$$\langle \xi_0, x - x^* \rangle + \mu y \geq \alpha, \forall x \in A, y \leq 0 \tag{2.7.3}$$

Si se toma $x = x^*$ e $y = 0$ en la igualdad anterior se sigue que $\alpha \leq 0$. Ahora bien, al tomar $x = x^*$ e $y = \varepsilon > 0$ en (2.7.2) se sigue que $\mu\varepsilon \leq \alpha$. Dado que esto es verdad para cualquier $\varepsilon > 0$, se tiene que $\mu \leq 0$ y $\alpha \geq 0$.

Así, se ha demostrado que $\mu \leq 0$ y $\alpha = 0$. Si $\mu = 0$, $\langle \xi_0, x - x^* \rangle \leq 0$ para cada $x \in \mathbb{R}^n$. Si se toma $x = x^* + \xi_0$, se sigue que:

$$0 \geq \langle \xi_0, x - x^* \rangle = \|\xi_0\|^2$$

y por lo tanto $\xi_0 = 0$. Dado que $(\xi_0, \mu) \neq (0,0)$, se debe tener que $\mu < 0$. Dividiendo (2.7.2) y (2.7.4) por $-\mu$ y denotando $-\xi_0/\mu$ por ξ, se tienen las siguientes desigualdades:

$$y \geq \langle \xi, x - x^* \rangle, \forall x \in \mathbb{R}^n, y > f(x) - f(x^*) \tag{2.7.5}$$

$$\langle \xi, x - x^* \rangle - y \geq 0, \forall x \in A, y \leq 0. \tag{2.7.6}$$

Si $y = 0$ en la ecuación anterior, se obtiene que $\langle \xi, x - x^* \rangle \geq 0$ para todo $x \in A$. De (2.7.7) se sigue que:

$$f(x) \geq f(x^*) + \langle \xi, x - x^* \rangle$$

para toda $x \in R^n$. Por lo tanto, ξ es un subgradiente de f en x^* .

3.1.1. DUALIDAD Y FUNCIÓN LAGRANGIANA

La dualidad entre una función convexa f y su conjugado de Fenchel f^*, radica en su poder para describir la teoría de la dualidad para problemas convexos, una de las ideas de mayor alcance en el estudio de optimización

Sea el problema convexo primario (P):

$$\min f(x)$$

Sujeto a:

$$g(x) \leq 0$$

$$h(x) = 0$$

$$x \in A$$

donde:

$f: A \subset \mathbb{R}^n \to \mathbb{R}$ función convexa.

$g: \mathbb{R}^n \to \mathbb{R}^m$ con funciones componentes $g_1, g_2, \ldots, g_m: A \to (-\infty; +\infty]$ funciones convexas.

$h: A \subset \mathbb{R}^n \to \mathbb{R}^l$ función afín.

Además $dom\,(f) \neq \phi$, con $dom\,(f) \subset \bigcup_1^m dom\,(g_i)$.

Entre las diversas formulaciones de dualidad, la formulación usando la función Lagrangiana es una de las que más ha llamado la atención. A continuación se define el problema dual (D):

$$\max \theta(u, v)$$

Sujeto a: $u \geq 0$

donde $\theta(u,v) = inf\left\{f(x) + \sum_{i=1}^{m} u_i g_i(x) \ + \sum_{j=1}^{l} v_j h_j(x) \ / x \in \mathbb{R}^n\right\}$.

Mediante el proceso de dualización, las restricciones de nuestro problema primario han sido agregadas a nuestra función objetivo del problema dual.

Notemos que la función dual Lagrangiana θ puede asumir el valor $-\infty$ para algunos vectores (u,v).

TEOREMA 3.1.3.- Sea $\bar{x}$ una solución factible del problema primario (P) y (u,v) una solución del problema dual (D). Entonces $f(\bar{x}) \geq \theta(u,v)$.

DEMOSTRACIÓN.-

Sea $\bar{x} \in A$, con A un subconjunto convexo de $\mathbb{R}^n$. Por la definición de θ, se tiene:

$$\theta(u,v) = inf\{f(y) + \langle u, g(y)\rangle + \langle v, h(y)\rangle / \ x \in A\}$$

En particular

$$\theta(u,v) \leq f(\bar{x}) + \langle u, g(\bar{x})\rangle + \langle v, h(\bar{x})\rangle$$

Como $u \geq 0$, $g(\bar{x}) \leq 0$ y $h(\bar{x}) = 0$, se tiene:

$$\theta(u,v) \leq f(x).$$

Claramente, se tiene:

$$inf\{f(x)/x \in C\} \geq sup\{\theta(u,v)/u \geq 0\}$$

COROLARIO 3.1.1.- Si $f(\bar{x}) = \theta(\bar{u}, \bar{v})$, donde $\bar{u} \geq 0$ y $\bar{x} \in A$, entonces $\bar{x}$ y $(\bar{u}, \bar{v})$ son soluciones del problema primario (P) y el problema dual (D) respectivamente.

Diremos que existe una brecha de dualidad cuando:

$$inf\{f(x)/x \in C\} > sup\{\theta(u,v)/u \geq 0\}$$

TEOREMA 3.1.4.- Sean A un conjunto convexo, no vacio, en $\mathbb{R}^n$, $f: \mathbb{R}^n \to \mathbb{R}$ y $g: \mathbb{R}^n \to \mathbb{R}^m$ funciones convexas, y sea y $h: \mathbb{R}^n \to \mathbb{R}^l$ una función afín; esto es: $h(x) = Tx - b$. Si existe $\hat{x} \in A$ tal que $g(\hat{x}) < 0$, $h(\hat{x}) = 0$ y $0 \in int(h(A))$, entonces:

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

Además, si el inf es finito, entonces el $sup\{\theta(u,v)/u \geq 0\}$ es lograda en $(\bar{u}, \bar{v})$ con $\bar{u} \geq 0$. Si el inf es logrado en $\bar{x}$, luego $\langle \bar{u}, g(\bar{x}) \rangle = 0$.

DEMOSTRACIÓN.-

Sea $\gamma = inf\{f(x)/x \in A, g(x) \leq 0, h(x) = 0\}$ y $\gamma < \infty$. Si $\gamma = -\infty$ y por el teorema (3.1.3)

$$\sup\{\theta(u,v)/u \geq 0\} = -\infty$$

de donde lo que asumimos es verdadero.

Supongamos que γ es finito, consideremos el siguiente sistema:

$$f(x) - \gamma < 0,\ g(x) \leq 0,\ h(x) = 0,\ x \in A$$

De la definiion de γ el sistema no tiene solución. Del teorema (3.1.3), exite un vector no nulo (u_0, u, v), con $(u_0, u) \geq 0$, tal que:

$$u_0[f(x) - \gamma] + \langle u, g(x) \rangle + \langle v, h(x) \rangle \geq 0, \quad \forall x \in A \tag{3.1.1}$$

Primero mostraremos que $u_0 > 0$. Por cntradicción, supongamos que $u_0 = 0$, luego existe un $\hat{x} \in A$ tal que $\mathrm{g}(\hat{\mathrm{x}}) < 0$ y $\mathrm{h}(\hat{\mathrm{x}}) = 0$. Sustituyendo en (3.1.1) se tiene $\langle u, g(\hat{\mathrm{x}}) \rangle \geq 0$. Como $\mathrm{g}(\hat{\mathrm{x}}) < 0$ y $u \geq 0$, $\langle u, g(\hat{\mathrm{x}}) \rangle \geq 0$ es posible solo si $u = 0$. Pero desde (3.1.1), $u_0 = 0$ y $u = 0$, implica que $\langle u, g(\hat{\mathrm{x}}) \rangle \geq 0$ para todo $x \in A$. Desde que $0 \in int\, h(A)$, podemos elegir un $x \in A$ tal que $h(x) = -\lambda v$, $\lambda > 0$. En consecuencia $0 \leq \langle v, h(x) \rangle = -\lambda \|v\|^2$ implica que $v = 0$. Entonces. Se ha probado que $u_0 = 0$ implica que $(u_0, u, v) = 0$, lo que es una contradicción. Por consiguiente $u_0 > 0$.

Dividiendo (3.1.1) por u_0, se tiene:

$$f(x) + \langle \bar{u}, g(x) \rangle + \langle \bar{v}, h(x) \rangle \geq \gamma, \quad \forall x \in A \tag{3.1.2}$$

donde $\bar{u} = \frac{u}{u_0}$ y $\bar{v} = \frac{v}{u_0}$.

Esto muestra que $\theta(\bar{u}, \bar{v}) = \inf\{f(x) + \langle \bar{u}, g(x) \rangle + \langle \bar{v}, h(x) \rangle \geq \gamma / \forall x \in A\}$. Además esta claro que $\theta(\bar{u}, \bar{v}) = \gamma$ y $(\bar{u}, \bar{v})$ soluciona el problema dual.

Por otro lado, supongamos que $\bar{x}$ es una solución óptima del problema primario; esto es, $\bar{x} \in A$, $g(\bar{x}) \leq 0$, $h(\bar{x}) = 0$ y $f(\bar{x}) = \gamma$. En (3.1.2), tomando $x = \bar{x}$, se tiene: $g(\bar{x}) \geq 0$. Como $\bar{u} \geq 0$ y $g(\bar{x}) \leq 0$, se tiene:

$$\langle \bar{u}, g(\bar{x}) \rangle = 0$$

con lo que prueba concluye.

DEFINICIÓN 3.1.2.- Sean $f: A \subset \mathbb{R}^n \to \mathbb{R}$ función convexa, $g: \mathbb{R}^n \to \mathbb{R}^m$ una función con funciones componentes $g_1, g_2, \dots, g_m: A \to (-\infty; +\infty]$ funciones convexas y $h: A \subset \mathbb{R}^n \to \mathbb{R}^l$ función afín. La función Lagrangiana $\mathcal{L}: A \times \mathbb{R}_+^m \times \mathbb{R}^l \to (-\infty; +\infty]$,

se define como:

$$\mathcal{L}(x, u, v) = f(x) + \langle u, g(x) \rangle + \langle v, h(x) \rangle$$

donde $u \geq 0$.

Una condición necesaria y suficiente para la existencia de una solución óptima para el problema primario (P) es la existencia de un punto de silla.

DEFINICIÓN 3.1.3.- Sean $f: A \subset \mathbb{R}^n \to \mathbb{R}$ función convexa, $g: \mathbb{R}^n \to \mathbb{R}^m$ una función con funciones componentes $g_1, g_2, \dots, g_m: A \to (-\infty; +\infty]$ funciones convexas, $h: A \subset \mathbb{R}^n \to \mathbb{R}^l$ función afín y la función Lagrangiana $\mathcal{L}: A \times \mathbb{R}_+^m \times \mathbb{R}^l \to (-\infty; +\infty]$. $(\bar{x}, \bar{u}, \bar{v})$ es llamado punto de silla de la función Lagrangiana $\mathcal{L}$, si $\bar{x} \in \mathbb{R}^n$, $\bar{u} \geq 0$, y $\mathcal{L}(\bar{x}, u, v) \leq \mathcal{L}(\bar{x}, \bar{u}, \bar{v}) \leq \mathcal{L}(x, \bar{u}, \bar{v})$, para todo $x \in \mathbb{R}^n$ y un (u, v) tal que $u \geq 0$.

Claramente, $\bar{x}$ minimiza a $\mathcal{L}$ en $\mathbb{R}^n$ cuando (u, v) se fija en $(\bar{u}, \bar{v})$,y $(\bar{u}, \bar{v})$ maximiza $\mathcal{L}$ en $\mathbb{R}^m \times \mathbb{R}^l$, con $\bar{u} \geq 0$, cuando x se fija en $\bar{x}$.

TEOREMA 3.1.5.- Sea el punto $(\bar{x}, \bar{u}, \bar{v})$ tal que $\bar{x} \in \mathbb{R}^n$ y $\bar{u} \geq 0$. Entonces $(\bar{x}, \bar{u}, \bar{v})$ es punto de silla para la función Lagrangiana $\mathcal{L}$ si y solo si

i. $\mathcal{L}(\bar{x}, \bar{u}, \bar{v}) = min\{\mathcal{L}(x, \bar{u}, \bar{v}) / g(x) \leq 0, h(x) = 0, x \in \mathbb{R}^n\}$

ii. $g(\bar{x}) \leq 0, h(\bar{x}) = 0$

iii. $\langle \bar{u}, g(\bar{x}) \rangle = 0$

Cabe notar que*:* $(\bar{x}, \bar{u}, \bar{v})$ es punto de silla si y solo si $\bar{x}$ y $(\bar{u}, \bar{v})$ soluciones óptimas del problema primario P y del problema dual (D) respectivamente. Además es una condición necesaria y suficiente para la no existencia de la brecha de dualidad. Es decir, se cumple que $f(\bar{x}) = \theta(\bar{u}, \bar{v})$.

DEMOSTRACIÓN.-

Supongamos que $(\bar{x}, \bar{u}, \bar{v})$ es un punto de silla para la función Lagrangiana $\mathcal{L}$.

Por definición (i.) es verdadera. Además de (3.2.1) se tiene:

$$f(\bar{x}) + \langle \bar{u}, g(\bar{x}) \rangle + \langle \bar{v}, h(\bar{x}) \rangle \geq f(\bar{x}) + \langle u, g(\bar{x}) \rangle + \langle v, h(\bar{x}) \rangle$$

Para todo (u, v), con $u \geq 0$. De donde $g(\bar{x}) \leq 0$ y $h(\bar{x}) = 0$.

Tomando $u = 0$ en (3.1.3), se obtiene que $\langle \bar{u}, g(\bar{x}) \rangle \geq 0$. Cabe notar que si $\bar{u} \geq 0$ y $g(\bar{x}) \leq 0$ implica que $\langle \bar{u}, g(\bar{x}) \rangle \leq 0$, por lo que necesarimente $\langle \bar{u}, g(\bar{x}) \rangle = 0$.

Por otro lado, de (i.) se tiene:

$$\mathcal{L}(\bar{x}, \bar{u}, \bar{v}) \leq \mathcal{L}(x, \bar{u}, \bar{v}) \text{ , } \forall x \in \mathbb{R}^n$$

Además:

$$\mathcal{L}(\bar{x}, \bar{u}, \bar{v}) = f(\bar{x}) + \langle \bar{u}, g(\bar{x}) \rangle + \langle \bar{v}, h(\bar{x}) \rangle \geq f(\bar{x}) + \langle u, g(\bar{x}) \rangle + \langle v, h(\bar{x}) \rangle = \mathcal{L}(\bar{x}, u, v)$$

Para todo (x, v) con $u \geq 0$, desde que $g(\bar{x}) \leq 0$ y $h(\bar{x}) = 0$. Por lo tanto $(\bar{x}, \bar{u}, \bar{v})$ es un punto de silla. Lo que prueba la primera parte del teorema.

Ahora supongamos que $(\bar{x}, \bar{u}, \bar{v})$ es un punto de silla. Por (ii.) $\bar{x}$ es una solución factible del problema primario P. Luego como $\bar{u} \geq 0$, $(\bar{u}, \bar{v})$ es una solución factible para el problema dual D. Además por (i.), (ii.) y (iii.), se tiene:

$$\theta(\bar{u}, \bar{v}) = \mathcal{L}(\bar{x}, \bar{u}, \bar{v}) = f(\bar{x}) + \langle \bar{u}, g(\bar{x}) \rangle + \langle v, h(\bar{x}) \rangle = f(\bar{x})$$

Por el corolario (3.1.1), $\bar{x}$ y $(\bar{u}, \bar{v})$ solucionan P y D respectivamente, sin brecha de dualidad.

Finalmente, supongamos que $\bar{x}$ y $(\bar{u}, \bar{v})$ son soluciones de problema primario P y el problema dual D, respectivamente, con $f(\bar{x}) = \theta(\bar{u}, \bar{v})$. Por lo tanto, tenemos $\bar{x} \in \mathbb{R}^n$, $g(\bar{x}) \leq 0$, $h(\bar{x}) = 0$ y $\bar{u} \geq 0$. Además por la factibilidad del primario y el dual, se tiene:

$$\theta(\bar{u}, \bar{v}) = \min\{f(x) + \langle \bar{u}, g(x)\rangle + \langle \bar{v}, h(x)\rangle / \ \bar{x} \in \mathbb{R}^n\}$$

$$\leq f(\bar{x}) + \langle \bar{u}, g(\bar{x})\rangle + \langle \bar{v}, h(\bar{x})\rangle = f(\bar{x}) + \langle \bar{u}, g(\bar{x})\rangle \leq f(\bar{x})$$

Por hipotesis $\theta(\bar{u}, \bar{v}) = f(\bar{x})$, por lo tanto la desigualdad es verdadera.

En particular, $\langle \bar{u}, g(\bar{x})\rangle = 0$, entonces:

$$\mathcal{L}(\bar{x}, \bar{u}, \bar{v}) = f(\bar{x}) = \theta(\bar{u}, \bar{v}) = min\{\mathcal{L}(x, \bar{u}, \bar{v}) / \bar{x} \in \mathbb{R}^n\}$$

Por lo tanto, de (i.), (ii.) y (iii.) $(\bar{x}, \bar{u}, \bar{v})$ es un punto de silla.

COROLARIO 3.1.2.- Sean $f: \mathbb{R}^n \to \mathbb{R}$, $g: \mathbb{R}^n \to \mathbb{R}^m$ funciones convexas y

$h: \mathbb{R}^n \to \mathbb{R}^l$ función afín. Además supongamos que $0 \in int\ h(\mathbb{R}^n)$ y que existe $\hat{x} \in \mathbb{R}^n$ tal que $g(\hat{x}) < 0$ y $h(\hat{x}) = 0$. Si $\bar{x}$ es una solución óptima de problema primario (P), entonces existe un vector $(\bar{u}, \bar{v})$, con $\bar{u} \geq 0$, tal que $(\bar{x}, \bar{u}, \bar{v})$ es un punto de silla.

La demostración de puede encontrar en Nonlinear Programming de Mokhtar S. Bazaraa y otros. (2006, Pag. 271)

Es de observar que el valor óptimo del problema primario coincide con el del problema dual si y solo si se cumple $sup\ inf\ \mathcal{L} = inf\ sup\ \mathcal{L}$.

3.2. CONDICIONES NECESARIAS DE PRIMER Y SEGUNDO ORDEN PARA FUNCIONES CONVEXAS NO DIFERENCIABLES

Se llama optimización no diferenciable, cuando la función objetivo no es diferenciable en todos los puntos del interior de su dominio. En estos puntos puede presentar dobleces o puntos de esquina. Más es necesario que la función sea continúa en el interior de su dominio.

Consideremos una función $f: A \to (-\infty, +\infty]$, entonces el problema P se define como:

$$P = \min_{x \in A} f(x) \qquad (3.2.1)$$

donde: A es un conjunto convexo en $\mathbb{R}^n$ y es la región de factibilidad.

La solución de este tipo de problemas se hallara mediante el uso de algoritmos iterativos.

Se sabe que el gradiente no existe en todos lados, lo que implica que la función puede tener dobleces o puntos de esquina. Por lo tanto no se puede aproximar localmente por un hiperplano tangente o por alguna aproximación cuadrática.

DEFINICIÓN 3.2.1.- Sean A un conjunto $\mathbb{R}^n$, $x \in \mathbb{R}^n$ y un vector, no nulo, $d \in \mathbb{R}^n$. Se dice que, d es una dirección factible con respecto a A en x si, existe una sucesión $\{x_n\}_{n\in\mathbb{N}}$ convergente a x en A y la sucesión:

$$d_n = \frac{x_n - x}{\|x_n - x\|} \to d \qquad (3.2.2)$$

LEMA 3.2.1.- Sean A un conjunto convexo, no vacío, de $\mathbb{R}^n$ y $f: A \to \mathbb{R}$ una función convexa. Sea d una dirección factible con respecto a A en x, donde $\{x_n\}_{n\in\mathbb{N}}$ es una sucesión como en la definición 3.2.1.

Si $x_n \to x$ con $\|x_n - x\| > 0$, entonces:

$$\lim_{n\to\infty} \frac{f(x_n) - f(x)}{\|x_n - x\|} = \max_{\xi\in\partial f(x)} \langle \xi, d\rangle$$

DEMOSTRACIÓN.-

Sea $\bar{x} \in int(A)$, para cada hiperplano de soporte en $\bar{x}$, se tiene:

$$f(x) \geq f(\bar{x}) + \langle \xi, x - \bar{x}\rangle;\ \forall x \in A$$

Si $\xi_n \in \partial f(x_n)$ entonces, de la ecuación anterior para n suficientemente grande, se tiene:

$$f(x_n) \geq f(x) + \langle \xi, x_n - x\rangle$$

Como $\|x_n - x\| > 0$, se tiene:

$$\frac{f(x_n) - f(x)}{\|x_n - x\|} \geq \langle \xi, d\rangle \geq \max_{\xi\in\partial f(x)} \langle \xi, d\rangle \tag{3.2.3}$$

Y por definición de máximo

$$\max_{\xi\in\partial f(x)} \langle \xi, d\rangle \geq \frac{f(x_n) - f(x)}{\|x_n - x\|} \tag{3.2.4}$$

de (3.2.3) y (3.2.4), se tiene:

$$\frac{f(x_n) - f(x)}{\|x_n - x\|} = \max_{\xi\in\partial f(x)} \langle \xi, d\rangle$$

Para un n suficientemente grande.

El lema 3.2.1 muestra que la derivada direccional está determinada por un hiperplano de soporte extremo cuyo subgradiente da el mayor descenso. Por lo tanto una condición necesaria para la existencia de un mínimo local es que la derivada direccional es no negativa en todas sus direcciones.

El concepto de subdiferencial y los resultados de la derivada direccional nos dirigen a las condiciones de primer orden.

Para el estudio de la optimización convexa no diferenciable, se deben estudiar funciones más generales, que son construidas de la siguiente manera:

$$\phi(x) = f(x) + h\big(c(x)\big)$$

donde:

$f: \mathbb{R}^n \to \mathbb{R}$ función de clase C^1.

$c: \mathbb{R}^n \to \mathbb{R}^m$ función de clase C^1.

$h \circ c: \mathbb{R}^n \to \mathbb{R}$ función convexa de clase C^0, no diferenciable.

Las condiciones de optimalidad se conseguirán como una extensión sencilla por funciones convexas.

Si d es una dirección factible en $\mathbb{R}^n$, por (3.2.2), la definición de derivada direccional y la utilizando la serie de Taylor, se tiene:

$$f(x_n) = f(x) + \|x_n - x\|\langle \nabla f(x), d_n\rangle + o(\|x_n - x\|)$$

Donde: $o: A \to \mathbb{R}$ una función con $o(x) = \lim_{x_n \to x} o(x_n) = 0$

Asi

$$\frac{f(x_n) - f(x)}{\|x_n - x\|} \to \langle \nabla f(x), d \rangle$$

De forma similar, se obtiene para c

$$\frac{c(x_n) - c(x)}{\|x_n - x\|} \to D^* c(x). d$$

donde: $c(x_n) \to c(x)$ es una sucesión direccional en $\mathbb{R}^m$.

Por la composición de h con c, se tiene:

$$\lim_{\|x_n - x\| \to 0} \frac{\phi(x_n) - \phi(x)}{\|x_n - x\|} = \max_{\xi \in \partial f(x)} \langle \nabla f(x) + D^*\big(c(x)\big)\xi, d \rangle \qquad (3.2.5)$$

esto es la derivada direccional de x en la dirección factible d para la función $\phi(x)$.

Si x^* es un minimo local de ϕ, entonces $\phi(x_n) \geq \phi(x^*)$, con n suficientemente grade. Esto es:

$$\max_{\xi \in \partial f(x)} \langle \nabla f(x^*) + D^*\big(c(x^*)\big)\xi, d \rangle \geq 0\,, \qquad \forall d$$

El lema 3.2.1 muestra que la derivada direccional está determinada por un hiperplano de soporte extremo cuyo subgradiente da el mayor descenso. Por lo tanto, una condición necesaria para la existencia de un mínimo local es que la derivada direccional sea no negativa en todas sus direcciones. Este resultado, se puede escribir como:

$$0 \in \partial\phi(x^*) = \{\nabla f(x^*) + D^* c(x^*)\xi : \xi \in \partial h(x^*)\} \qquad (3.2.6)$$

Las ecuación (3.2.3) y el conjunto (3.2.4) son el resultado de una generalización de la regla de la cadena para funciones diferenciables. (Clarke, 1990)

El conjunto $\partial\phi(x^*)$ es convexo y compacto, no es la subdiferencial, puesto que ϕ no necesariamente es una función convexa, aunque se usa la misma notación por conveniencia.

Otra forma de establecer la condición de primer orden es usando la función Lagrangiana, de la siguiente manera:

$$\mathcal{L}(x,\lambda) = f(x) + \langle \lambda, c(x) \rangle$$

Donde $\lambda \in \mathbb{R}^n$. Esta ecuación es equivalente a equivalente a (3.2.6).

TEOREMA 3.2.1 (condiciones necesarias de primer orden).- Sean $x^* \in \mathbb{R}^n$ y $\phi: \mathbb{R}^n \to \mathbb{R}$ una función. Si x^* es un mínimo de ϕ, entonces existe un vector $\lambda^* \in \partial h(x^*)$ tal que:

$$\nabla\mathcal{L}(x^*,\lambda^*) = \nabla f(x^*) + D^*c(x)\lambda^* = 0$$

La demostración de puede encontrar en Practical Methods of Optimization de Roger Fletcher. (2000, Pag. 370).

Este resultado muestra la relación cercana que hay entre el vector $\lambda^* \in \partial h(x^*)$ y los multiplicadores de Lagrange. Ya que la función ϕ no necesariamente es convexa, es importante considerar condiciones de segundo orden, para garantizar que x^* sea mínimo local de ϕ. Estas condiciones muestran la relación entre nuestro problema y el problema en el caso que ϕ sea diferenciable. (Nava Manzo, 2015)

Se define el conjunto X por:

$$X = \{x / h(c(x)) = h(c(x^*)) + \langle c(x) - c(x^*), \lambda^* \rangle\}$$

donde λ^* cumple con las condiciones del teorema anterior.

Se define $\mathcal{G}^*$ como el conjunto de las direcciones factibles normalizadas con respecto a X en x^*. Esto es, $d \in \mathcal{G}^*$ implica que existe una sucesión de direcciones factibles de $x_n \to x^*$ tal que $d_n \to d$ con $\|d\| = 1$. $\mathcal{G}^*$ puede ser considerado como el conjunto de las direcciones estacionarias para $h\big(c(x)\big)$, habiendo linealizado $h(c)$ pero no $c(x)$. Estas direcciones están relacionadas con el conjunto G^* de las direcciones normalizadas de pendiente cero, con:

$$G^* := \left\{ d \Big/ \max_{\xi \in \partial h(x^*)} \langle \nabla f(x^*) + D^*\big(c(x^*)\big)\xi, d \rangle = 0, \|d\| = 1 \right\}$$

Esto es: G^* es el conjunto de direcciones factibles que hacen $\phi'(x^*, d) = 0$.

G^* puede ser considerado como el conjunto de las direcciones estacionarias para $h\big(c(x)\big)$ habiendo linealizado $h(c)$ y $c(x)$.

LEMA 3.2.2.- Sean $\mathcal{G}^*$ el conjunto de las direcciones factibles normalizadas respecto a X en x^* y G^* es el conjunto de direcciones normalizadas de pendiente cero. Entonces $\mathcal{G}^* \subset G^*$.

DEMOSTRACION.- (Fletcher, 2013)

Sea $d \in \mathcal{G}^*$. Entonces d es una dirección factible con respecto a X en x^* con $d_n \to d$, $\|d\| = 1$. Luego, existe una sucesión $\{\mathrm{x_n}\}_{\mathrm{n}\in\mathbb{N}}$ en X convergente a x^*. Utilizando la serie de Taylor, el lema (3.2.1) y la definición de ϕ, se tiene:

$$\max_{\lambda^* \in \partial h(x^*)} \langle \nabla f(\acute{x^*}) + Dc(x^*)\lambda^*, d \rangle = \lim_{n\to\infty} \frac{\phi(x_n) - \phi(x^*)}{||x_n - x^*\|}$$

$$= \lim_{n\to\infty} \frac{f(x_n) - f(x^*) + h(c(x_n)) - h(c(x^*))}{\|x_n - x^*||}$$

$$= \lim_{n\to\infty} \frac{f(x_n) - f(x^*) + \langle c(x_n) - c(x^*), \lambda^* \rangle}{||x_n - x^*||}$$

$$= \langle \nabla f(x^*) + D^* c(x^*)\lambda^*, d\rangle.$$

Por el teorema (3.2.1) , se cumple que $d \in G^*$.

Se tiene una condición de regularidad como en el caso diferenciable para poder deducir condiciones de segundo orden. La condición de regularidad para nuestro problema será: $\mathcal{G}^* = G^*$.

DEFINICIÓN 3.2.2.- Si existe una vecindad abierta $V\big(c(x^*)\big)$ tal que:

$$h\big(c(x)\big) = h\big(c(x^*)\big) + \max_{\lambda \in \partial h(x^*)} \langle c(x), c(x^*), \lambda\rangle$$

para todo $c(x) \in V\big(c(x^*)\big)$, se dice que $h\big(c(x)\big)$ es localmente lineal en $c(x^*)$.

Se denotara por l^* a la dimensión de $\partial\phi(x^*)$ con $(l^* \leq m)$. Se define a la matriz

$$D \subset \mathbb{R}^{m \times l^*}$$

Cuyas columnas son los elementos $d_i^*,\ i = \overline{1, l^*}$.

LEMA 3.2.3 (Condiciones suficientes de regularidad).- Si x^* satisface las condiciones de primer orden, si $h(c)$ es localmente lineal en c^* y si

$$rango(\nabla f(x^*) D^*) = l^*$$

Entonces $\mathcal{G}^* = G^*$. (Nava Manzo, 2015)

DEMOSTRACIÓN (Fletcher, 2013)

Por el lema (3.2.2), se tiene $\mathcal{G}^* \subseteq G^*$.

Por otro lado, sea $d \in G^*$

$$\partial h_d^* = \big\{\lambda \in \partial h^* / \langle \nabla f(x^*) + D^*\big(c(x^*)\big)\lambda, d\rangle = 0\big\}$$

Como ∂h_d^* depende de d, por las condiciones de primer orden, se tiene:

$$\langle D^*\big(c(x^*)\big)(\lambda - \lambda^*), d\rangle = 0\,, \quad \forall \lambda \in \partial h_d^*$$

Además

$$G^* := \left\{ d / \max_{\lambda \in \partial h^*} \langle D^*\big(c(x^*)\big)(\lambda - \lambda^*), d\rangle = 0, \|d\| = 1 \right\}$$

Entonces:

$$\langle D^*\big(c(x^*)\big)(\lambda - \lambda^*), d\rangle < 0, \forall \lambda \in \partial h^* \ \setminus \partial h_d^*$$

La dimensión de ∂h_d^* es $l_d^*(l_d^* \leq l^*)$. Sea los vectores básicos $d_i^*, i = \overline{1, l_d^*}$ de $\partial h_d^* - \lambda^*$, entonces:

$$\langle D^*\big(c(x^*)\big) d_i^*, d\rangle = 0, i = \overline{1, l_d^*}$$

o

$$\langle D^*\big(c(x^*)\big) d_i^*, d\rangle < 0, i = \overline{l_d^*, l^*}$$

Si $l_d^* = n$ entonces $d = 0$ lo que contradice $s \in G^*$. Si $l_d^* < n$ es posible construir una función diferenciable $arc\ x(\theta), \theta \in [0, \bar{\theta})$ para $x(0) = x^*$ y

$x'(0) = d$, entonces:

$$\max_{\lambda \in \partial h^*} \langle (\lambda - \lambda^*), c\big(x(\theta)\big) - c^*\rangle = 0$$

Por hipótesis $\mathrm{h}(\mathrm{c})$ es localmente lineal en c^*, por lo que existe una vecindad c^* tal que:

$$\mathrm{h}\Big(\mathrm{c}\big(x(\theta)\big)\Big) = h(c^*) + \langle c\big(x(\theta)\big) - c^*, \lambda^*\rangle$$

y tomando una sucesión $\{\theta_k\}_k$convergente a cero, da una sucesión direccional que es factible para $d \in \mathcal{G}^*$.

Estas no son las únicas condiciones suficientes para tener $\mathcal{G}^* = G^*$, por ejemplo, el mismo resultado se tiene si c es una función afin o si se puede probar que $G^* = \phi$. (Fletcher, 2013)

Ahora es posible establecer condiciones de segundo orden, para lo que supondremos que f y c son de clase $\mathbb{C}^2$. Como es usual, las condiciones de regularidad solo son utilizadas para establecer condiciones necesarias. (Fletcher, 2013)

TEOREMA 3.2.2.- (Condiciones necesarias de segundo orden) Si x^* minimiza $\phi(x)$ y $\mathcal{G}^* = G^*$ se mantiene, entonces:

$$\langle d, \nabla^2 \mathcal{L}(x^*, \lambda^*) d \rangle \geq 0$$

para todo $d \in G^*$, con el mismo λ^* del teorema (2.8.1)

DEMOSTRACIÓN (Fletcher, 2013)

Sea $d \in \mathcal{G}^*$, con d es una dirección factible normalizada con respecto a X en x^*, existe una sucesión $\{x_n\}_{n\in\mathbb{N}}$ contenida en X tal que:

$$x_n \to x^*; \text{ con } x^* \text{ que minimiza a } \phi$$

denotaremos:

$$\delta_n = \|x_n - x^*\| \text{ y } e_n = x_n - x^*$$

Sea λ^* que satisface las condiciones del teorema 3.2.1.

La expansión de Taylor de $\mathcal{L}(x, \lambda^*)$ alrededor de x^* estará dada por:

$\mathcal{L}(x_n, x^*) = f(x^*) + \langle c(x^*)\ , \lambda^* \rangle + \langle \nabla \mathcal{L}(x^*, \lambda^*)\ , e_n \rangle + \frac{1}{2} \langle e_n, \nabla^2 \mathcal{L}(x^*, \lambda^*) e_n \rangle +$

$o(\|e_n\|^2)$

$$\mathcal{L}(x_n, \lambda^*) = f(x^*) + \langle c(x^*)\ , \lambda^* \rangle + \frac{1}{2} \delta_n^2 \langle d_n \nabla^2 \mathcal{L}(x^*, \lambda^*), d_n \rangle + o({\delta_n}^2)$$

$$\mathcal{L}(x_n, \lambda^*) = f(x^*) + \langle c(x^*)\ , \lambda^* \rangle + \frac{1}{2} \langle d_n, d_n \nabla^2 \mathcal{L}(x^*, \lambda^*) \rangle$$

Por la definición de ϕ, se tiene:

$$\phi(x_n) = \phi(x^*) + \frac{1}{2} \delta_n^2 \langle d_n \nabla^2 \mathcal{L}(x^*, \lambda^*), d_n \rangle \tag{3.2.7}$$

Dado que x^* es un mínimo local, $\phi(x_n) - \phi(x^*) \geq 0$ (para n suficientemente grande), con lo que (3.2.7) se escribe como:

$$0 \leq \frac{1}{2} \delta_n^2 \langle d_n, d_n \nabla^2 \mathcal{L}(x^*, \lambda^*) \rangle$$

Y como $\frac{1}{2} \delta_n^2 > 0$, se tiene:

$$0 \leq \langle d_n, d_n \nabla^2 \mathcal{L}(x^*, \lambda^*) \rangle.$$

CONCLUSIONES

1. Se encontraron valores óptimos para funciones convexas no diferenciables bajo las condiciones necesarias de primer y segundo orden, solo con la condición de continuidad.

2. El subdiferencial es la generalización de la derivada en el caso de funciones convexas diferenciables. El subdiferencial es el conjunto de subgradientes ξ.

3. Se construyó la función $\phi: \mathbb{R}^n \to \mathbb{R}$, definida por:

$$\phi(x) = f(x) + h\big(c(x)\big)$$

   donde:

   $f: \mathbb{R}^n \to \mathbb{R}$ función de clase C^1.

   $c: \mathbb{R}^n \to \mathbb{R}^m$ función de clase C^1.

   $h \circ c: \mathbb{R}^n \to \mathbb{R}$ función convexa de clase C^0, no diferenciable.

   Para el estudio de funciones convexas no diferenciables. Se determinó el conjunto $\{\nabla f(x^*) + D^* c(x^*)\xi : \xi \in \partial h(x^*)\}$ que es el equivalente al subdiferencial.

4. Se determinó la condición necesaria de primer orden para la existencia de valores óptimos. Para determinar la condición de segundo orden se necesitó una condición de regularidad.

BIBLIOGRAFIA

Bauschke H., C. P. (2011). *Convex Analusis and Monotone Operator Theiry in Hilbert Spaces.* Canada: Springer.

Borda Marcatinco, D. C. (10 de 2013). *repositorio.unac.edu.pe*. Obtenido de http://repositorio.unac.edu.pe/bitstream/handle/UNAC/110/DanteCesar_Tesis_tituloprofesional_2013.pdf?sequence=3&isAllowed=y

Clarke, F. (1990). *Optimization and Nonsmooth Analysis, Classics in Applied Mathematics.* SIAM: Wiley.

Crisóstomo, R. R. (s.f.). *Metodología de la Investigación Científica.* Lima: Universidad Tecnológica del Perú.

Fletcher, R. (2013). *Practical Methods of Optimization.* Wiley.

Gajardo, P. (2006). Introducción al Análisis Convexo. Valparaíso, Chile.

H., B. (1983). *Analyse fonctionnelle.* Paris: Masson.

Jhan, J. (1996). *Introduction to Theory of Nonlinear Optimization.* Belin: Springer.

M. Bazzara, H. S. (2006). *Nonlinear Programming.* Wiley.

Manzo, R. A. (2015). Funciones convexas no Diferenciabes. Mexico D.F.

Moreau, J., Strang, G., & Panagiotopoulos, P. (1988). *Topics in Nonsmooth Mechanics.* Boston: Birkhäuser.

Panagiotopoulos, P. (1985). *Inequality Problems in Mechanics and Applications. Convex and Nonconvex Energy Funcionals.* Boston: Birkhäuser.

Persson, L. (2006). *Convex Funtions and Their Applications.* Springer and Science Media.

R., F. (2013). *Practical Methods of optimization.* Wiley.

Rockafellar, R. (1970). *Convex Analysis.* Princeton University Press.

Rosenberg, B. (1973). Linear regression with randomly dispersed parameters. *Oxford Journals*, 65-72.

S. Boyd, A. M. (2006). Subgradient methods. *Stanford University*.

Xu, S. (2001). *Smoothing Methods for Minimax Problems.*

LINKOGRAFIA

(s.f.). Obtenido de http://tarwi.lamolina.edu.pe/~corihuela/mpeconomistas/capitulo%204-Optimizacion.pdf

(13 de 05 de 19). Obtenido de https://es.slideshare.net/jonciosito/anlisis-estructuraliti

(12 de 9 de 19). Obtenido de https://www.alamy.es/foto-armadura-de-acero-del-ferrocarril-automovil-cable-puente-peatonal-de-hormigon-y-proporcionar-acceso-al-centro-de-dallas-a-traves-del-trinity-river-124261907.html

Nava Manzo, R. A. (13 de febrero de 2015). *mat.izt.uam.mx/mat/*. Obtenido de http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.12-O/Nava_Manzo_Rafael_Alejandro.pdf

Vera Donoso, C. A. (2006). *repositorio.conicyt.cl*. Obtenido de http://repositorio.conicyt.cl/handle/10533/179324

http://programacionlinealunadibague.blogspot.com/2011/05/funciones-convexas-y-concavas.html. (Mayo de 2011). Obtenido de http://programacionlinealunadibague.blogspot.com/2011/05/funciones-convexas-y-concavas.html

ANEXOS

MATRIZ DE CONSISTENCIA

TÍTULO: OPTIMIZACIÓN DE FUNCIONES CONVEXAS NO DIFERENCIABLES

PROBLEMA GENERAL	**OBJETIVO GENERAL**	**HIPOTESIS GENERAL**	
¿Es posible encontrar valores óptimos para funciones convexas no diferenciables, solo con la condición de continuidad?	Determinar la existencia de valores óptimos para funciones convexas no diferenciables, solo con la condición de continuidad.	Existen valores óptimos para funciones convexas no diferenciables, solo con la condición de continuidad.	**VARIABLE INDEPENDIENTE:** Funciones continuas
PROBLEMAS ESPECÍFICOS	**OBJETIVOS ESPECÍFICOS**	**HIPÓTESIS ESPECÍFICOS**	
1. ¿Es posible generalizar el concepto de derivada para funciones convexas diferenciables? 2. ¿Es posible determinar un equivalente de la derivada para funciones convexas no diferenciables? 3. ¿Es posible determinar un valor óptimo para funciones convexas no diferenciables?	1. Obtener la generalización de la derivada para funciones convexas diferenciables. 2. Determinar un equivalente de la derivada para funciones convexas no diferenciables. 3. Determinar las condiciones de existencia de valores óptimos para funciones convexas no diferenciables.	1. Existe una generalización de la derivada para funciones convexas diferenciables. 2. Es posible determinar un equivalente derivada para funciones convexas no diferenciables. 3. Existen condiciones de existencia de valores óptimos para funciones convexas no diferenciables.	**VARIABLES DEPENDIENTES:** Valores óptimos de funciones convexas no diferenciables