UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA DR ALBERTO SÁNCHEZ
Ciudad Universitaria,

PERFIL DE TRABAJO DE TRABAJO DE GRADUACIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE:
LICENCIADA EN MATEMÁTICA.

TITULADA:
OPTIMIZACIÓN CONVEXA NO DIFERENCIABLE CON APLICACIONES EN CONTROL ÓPTIMO ESTOCÁSTICO.

PRESENTADA POR:
BR. LENNY IVONNE ARIAS RIVERA

ASESOR:
MSC. PORFIRIO ARMANDO RODRÍGUEZ

CIUDAD UNIVERSITARIA 16 DE MARZO del 2017

Contents

1 Introducción

La Matemática, como ciencia, vive y crece con el intercambio de ideas entre los cultivadores de diferentes áreas del conocimiento y el desarrollo del quehacer de la vida del ser humano y su entorno; y por su puesto alimentada por investigaciones que se desarrollan día a día en distintos espacios de la actividad científica.
La programación lineal fue planteada como un modelo matemático desarrollado durante la Segunda Guerra Mundial a fin de reducir los costos del ejército y aumentar las pérdidas del enemigo. Los fundadores de la técnica son George Dantzig, quien publicó el método simplex en 1947, John Von Neumann, que desarrolló la teoría de la dualidad en el mismo año y Leonid Kantoróvich, un matemático ruso que utilizó técnicas similares aplicadas a la economía antes de Dantzig y que ganó el premio Nobel en economía en 1975. (Véase [3])
La teoría moderna de optimización comenzó esencialmente con el desarrollo del método simplex.
En el área de investigación operativa la programación lineal estudia la minimización de una función lineal restringida a un conjunto definido por desigualdades lineales. Es una herramienta fundamental gracias a:

- La existencia de algoritmos eficientes de resolución como el simplex (a pesar de su complejidad no polinomial).
- La cantidad de problemas que entran dentro de este paradigma.
- La cantidad de problemas que sin ser lineales pueden ser aproximados o resueltos mediante sucesiones de éstos.

Sin embargo existen numerosos problemas en los que las herramientas lineales no son suficientes (Véase [9])y que dan lugar al estudio de la programación no lineal. Por ejemplo:

- Problemas de mínimos cuadrados originados en la teoría de redes neuronales.
- Problemas de diseño estructural (mecánico, eléctrico, espacial).
- Problemas de control óptimo (lanzamiento cohete-satélite, planeamiento de producción).
- Problemas de control óptimo estocástico (operación de una represa, de una red de poliductos, portafolio de acciones).
- Problemas de ruteo en redes de transporte y telecomunicaciones.

Las funciones no diferenciables en todas partes, marcaron un paradigma en la teoría moderna de optimización. Se encontró que muchos problemas de optimización convexa no eran diferenciables en el punto mínimo. Así, un enfoque completamente diferente se desarrolló, en donde fue concebida la noción de subdiferencial.
De esta forma los enfoques modernos de la teoría de optimización deben sus orígenes al cálculo de variaciones que ha sido estudiado por más de tres siglos y que también fue crucial en el desarrollo del análisis funcional.

Los métodos computacionales para la optimización no diferenciable se desarrollaron en dos direcciones:

a) La investigación dirigida a resolver determinados tipos de problemas de minimización con funciones no diferenciables que tienen una estructura especial y que se define de forma explícita

b) La investigación sobre la elaboración de métodos para resolver clases más generales de problemas, que no suponen de antemano el conocimiento de la estructura específica de la función a minimizar pero requieren la evaluación de la función y sus gradientes (o sus análogos en el caso no diferenciable) en cualquier punto dado.

Para el primer grupo, se cuenta con numerosos trabajos sobre métodos para resolver problemas de minimización. Para el segundo grupo una serie de obras se dedican a la minimización de funciones convexas lineales a trozos. Para la solución de varios problemas de optimización no diferenciable en los que se dan explícitamente los puntos donde la función no es diferenciable (por ejemplo, las funciones con valores absolutos), se han desarrollado técnicas especiales de suavizado.

En lo que se refiere a métodos generales de optimización no diferenciable, se pueden distinguir dos clases básicas en los cuales se requiere el cálculo del subdiferencial.

1. Métodos de gradiente generalizado. Para minimizar funciones diferenciables, se usan con frecuencia múltiples versiones del método de gradiente, que es natural, ya que la dirección negativa del gradiente en un punto dado es una dirección de descenso. La selección de un tamaño de paso en la mayoría de estos métodos tiene por objeto disminuir de manera significativa el valor de la función objetivo en cada iteración.

2. El método de planos de corte para resolver problemas convexos. El de Kelley por ejemplo se basa en aproximaciones lineales a la gráfica de una función convexa por medio de hiperplanos de soporte; en cada iteración el método resuelve un problema de programación lineal con un número creciente de restricciones en cada iteración.

2 Justificación

En nuestra realidad, vivimos con ciertas comodidades y a medida que han transcurrido los años los conocimientos han avanzado tanto que hoy en día poseemos un sistema númerico universal, celulares, computadoras, internet, tarjetas de crédito, tarjetas de débito, identificaciones, cohetes espaciales, etc. Pero todos éstos pusieron un reto a la humanidad y para poder resolverlo no solo observamos que teníamos un problema sino que lo planteamos y buscamos una solución a éste. Claramente esto requiere de tiempo y conocimiento, debido a ello es importante que sepamos y resaltemos que todo tiene una base matemática.

En medicina por ejemplo, cuando se crean curas para enfermedades se deben tomar desiciones sobre que elementos químicos utilizar porque unos son más agresivos que otros.

Podemos decir que la optimización matemática es el estudio sobre cómo hacer la mejor elección cuando estamos limitados por un conjunto de requerimientos.

En cultivos (cualquier tipo de cultivo) se debe saber que tipo de fertilizantes e insecticidas usar y la cantidad correcta, de lo contrario eso resultaría en una mala cosecha o la muerte del cultivo.

Eso es solo por mencionar algunos ejemplos, lo que hacemos está ligado a las matemáticas y en cuanto a toma de desiciones no solo podemos estudiar estadística sino que tambien la optimización matemática se encuentra inmersa.

Ahora bien, sabemos que en las matemáticas trabajamos con espacios y funciones teóricas, por así decirlo, pero: que pasa cuando aplicamos nuestros conocimientos a la realidad? si bien es cierto que nuestro espacio y todas sus maravillosas propiedades se mantienen. Cuando hablamos en concreto de optimización convexa (cuya base teórica es el análisis convexo) nos referimos a minimizar funciones convexas reales definidas para una variable contenida dentro de un subconjunto convexo de un espacio vectorial.

Las funciones no diferenciables en todas partes, marcaron un paradigma en la teoría moderna de optimización [3]. Se encontró que muchos problemas de optimización convexa no eran diferenciables en el punto mínimo y aquí es donde surge la necesidad e importancia de estudiar Optimización Convexa no Diferenciable ya que debido a ella no solo se desarolló nueva teoría matemática sino que también las técnicas abarcadas por este campo de estudio son importantes en aplicaciones de ingeniería (problemas de gestión hidrotérmica, problemas de control óptimo de producción, reconstrucción óptima de imágenes tomográficas, problemas de ruteo en redes de transporte y telecomunicaciones.) porque al plantear un problema de diseño en forma convexa se puede identificar la estructura de la solución óptima, que suele revelar aspectos importantes de dicho diseño.

3 Objetivos

Objetivo General

Analizar la teoría de conjuntos y funciones convexas para obtener las herramientas matemáticas necesarias para afrontar la resolución de problemas de optimización con restricciones (programacion no lineal) y aplicaciones en áreas donde desempeñan un papel central.

Objetivos Específicos

1. Proporcionar conocimientos del análisis convexo para estudiar, formular y resolver problemas de optimización convexa en los cuales éstos son no diferenciables.

2. Brindar un material bibliográfico para el estudio de la optimización convexa en general y no diferenciable en particular que sirva como base de estudio para materias electivas.

3. Plantear aplicaciones y soluciones reales a problemas de programación no lineal que aparecen en el ámbito científico, industrial, telecomunicaciones, transporte, entre otros, con las herramientas que nos proporcionan las distintas areas de la matemática.

4. Vincular la teoría matemática, tanto la planteada en los planes de estudio de la carrera como como aquella que queda fuera y de la cual se ha hecho una recopilación, específicamente del análisis convexo con aplicaciones en distintas áreas mediante la investigación científica.

4 Sección I: Análisis convexo

4.1 Conjuntos convexos [1]

A lo largo de este proyecto trabajaremos en un espacio vectorial normado real $(X*, |\cdot|)$ en dualidad con su espacio dual topológico X^* dotado de la norma usual:

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

donde $\langle \cdot, \cdot \rangle$ denota el producto de dualidad entre X y X^*, es decir para una funcion lineal continua $x \in X^*$ definida sobre el espacio X a valores en $\mathbb{R}$, se tiene $x^* : X \longmapsto \mathbb{R}$ y escribiremos

$$\langle x, x^* \rangle = x^*(x) \quad \forall x \in X \tag{1}$$

Recordemos que la topología más pequeña en X que hace continua a los elementos en X^* es la *topología débil.* Por otro lado, dado $x \in X$ se tienen que el operador *evaluacion* (que notaremos de igual manera) $x : X^* \longmapsto \mathbb{R}$ definido por

$$x(x^*) = x^*(x) = \langle x, x^* \rangle$$

Es lineal continuo (para la topología inducida por $\| \cdot \|_*$ en X), por lo tanto $X \subset X^{**}$. La topología más pequeña en X^* que hace ser contimuos a los operadores evaluación (es decir, a los elementos de X) es la topología $*-$débil.

Se tendrá que un espacio vectorial normado X dotado de la topología débil y su dual X^* dotado de la topología $*-$débil estan en *dualidad* gracias al par de dualidad definido en (1).

Definición 4.1. *: Un conjunto $C \subset X$ se dice convexo si para todo par de puntos $x, y \in C$ se tiene*

$$[x, y] := \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\} \subset C$$

Definición 4.2. *: La suma vetorial [2]*

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_m x_m$$

es llamada **combinación convexa** *de $x_1, x_2, \ldots, x_m$ si los coeficientes λ_i son no negativos y $\sum_{i=1}^{m} \lambda_i = 1$*

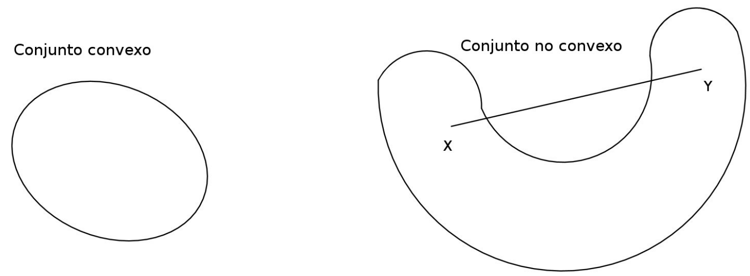


Figure 1: Conjunto convexo y no convexo

Definición 4.3. *Dado un conjunto $S \subset X$ no vacío, diremos que la cápsula convexa de S la cual denotaremos por $Co(S)$, es la intersección de todos los conjuntos convexos que contienen a S, es decir:*

$$Co(S) = \bigcap_{S \subset T,\ T\ convexo} T$$

La cápsula convexa de un conjunto S, $Co(S)$, es convexa. Más aún, la cápsula convexa de un conjunto S es el menor conjunto convexo que contiene a S.

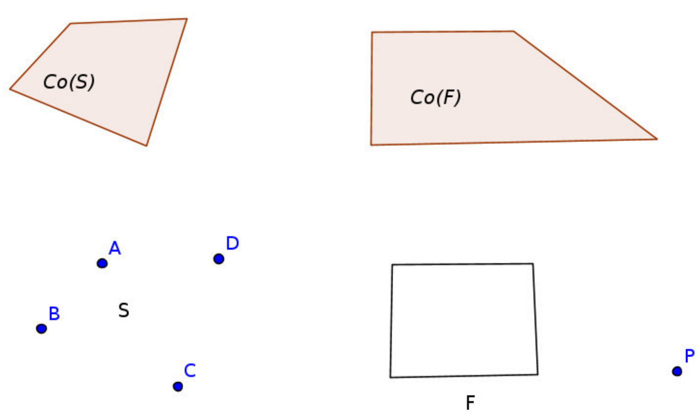


Figure 2: Cápsula convexa

Proposición 4.1. *La cápsula convexa de un conjunto $C \subset X$, no vacío, es el conjunto de todas las combinaciones convexas de elementos de C.*

También podemos definir la *cápsula afín* de C como la colección de todas las combinaciones afines de puntos en C. Esta es la dimensión afín más pequeña de un subespacio que contiene a C. Por ejemplo: la cápsula convexa de dos puntos distintos es una recta unidimensional

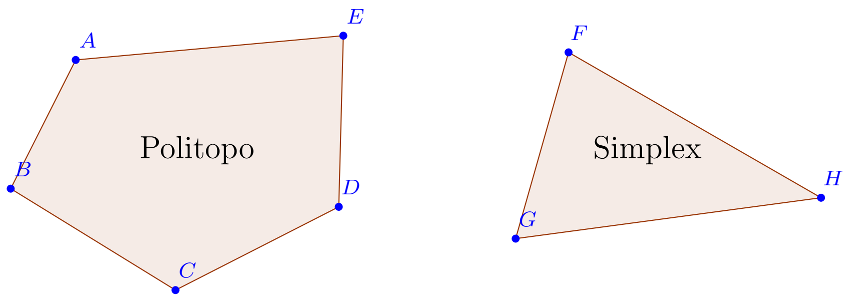


Figure 3: Politopo y simplex

conteniendo estos dos puntos. Similarmente, la *cápsula lineal* de C es la colección de todas las combinaciones lineales de puntos en C [2]

Definición 4.4. *:*

- *La cápsula convexa de un número finito de puntos $x_1, x_2, \ldots, x_{k+1} \in K$ es llamada politopo.*

- *Si $x_1, x_2, \ldots, x_k$ y x_{k+1} son independiente afines, es decir que $x_2 - x_1$, $x_3 - x_1, \ldots,$ $x_{k+1} - x_1$ son linealmente independientes, entonces la $Co(x_1, x_2, \ldots, x_{k+1})$ es llamada un simplex con vértives $x_1, x_2, \ldots, x_{k+1}$*

La figura 3 nos muestra un ejemplo de un politopo y un simplex en $\mathbb{R}^n$. Note que el número máximo de vectores linealmente independientes en $\mathbb{R}^n$ es n y por lo tanto no podría haber ningún simplex en $\mathbb{R}^n$ teniendo más de $n+1$ vértices.

Corolario 4.1. $C_1 \subset C_2 \Longrightarrow co(C_1) \subset co(C_2)$

Teorema de Caratheodory [2]
Por definición, un punto en el conjunto de una cápsula convexa puede ser representado coomo una combinación convexa de un número finito de puntos en el conjunto. El siguiente teorema muestra que cualquier punto x en la cápsula convexa de un conjunto C puede ser representado como una combinación de a lo sumo $n+1$ puntos en C. El teorema es trivialmente cierto para $x \in C$.

Teorema 4.1. *: Sea C un conjunto arbitrario en $\mathbb{R}^n$. Si $x \in Co(C)$, $x \in Co(x_1, \ldots, x_{n+1})$ donde $x_j \in C$, para $j = 1, \ldots, n+1$. En otras palabras x se puede representar como:*

$$x = \sum_{j=1}^{n+1} \lambda_j x_j; \quad x_j \in C \textit{ para } j = 1, \ldots, n+1$$
$$\sum_{j=1}^{n+1} \lambda_j = 1; \quad \lambda_j \geqslant 0 \textit{ para } j = 1, \ldots, n+1$$

4.2 Propiedades topológicas de los conjuntos convexos

A continuación desarrollaremos algunas propiedades topológicas para conjuntos convexos [2]. Como un preliminar a esta parte tendremos que dado un punto $x \in \mathbb{R}^n$ un ε-vecindario alrededor del conjunto es $N_\varepsilon(x) = \{y : \| y - x \| < \varepsilon\}$. Primero revisaremps las definiciones de: clausura, interior y frontera de un conjunto arbitrario en $\mathbb{R}^n$, usando el concepto de $\varepsilon-$vecindario.

Definición 4.5. *: Sea S un conjunto arbitrario en $\mathbb{R}^n$.*

- **Clausura***: Se dice que un punto $x \in \mathbb{R}^n$ está en la clausura de S si* $\forall\, \varepsilon > 0, \quad S \cap N_\varepsilon(x) \neq \emptyset$
- **Interior***:Un punto x se dice que está en el interior de S si* $\exists \varepsilon > 0$ *tal que* $N_\varepsilon(x) \subset S$.
- **Frontera***: Un punto x se dice que está en la frontera si $\forall\, \varepsilon > 0$ se tienen que* $N_\varepsilon(x) \cap S \neq \emptyset$ *y* $N_\varepsilon(x) \cap S^c \neq \emptyset$
- **Acumulación***: Un punto x es de acumulación de S si* $\forall\ \varepsilon > 0$ *se tiene que* $(N_\varepsilon(x) - \{x\}) \cap S \neq \emptyset$

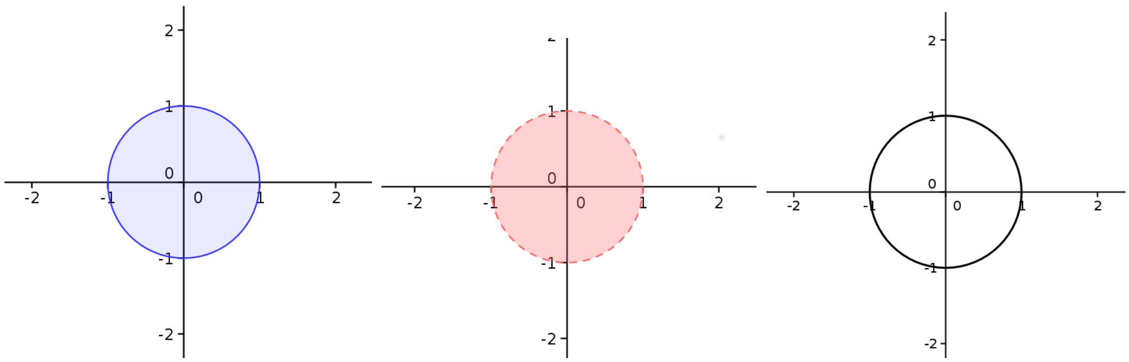
Notación:

- Denotaremos por $\overline{S}$ a la clausura del conjunto S. Además se tiene que S es cerrado si $\overline{S} = S$.
- Denotaremos por $\mathring{S}$ al interior del conjunto S y éste es abierto cuando $\mathring{S} = S$
- Denotaremos por $fr(S)$ a la frontera del conjunto S.
- Denotaremos por S' al conjunto de puntos de acumulación de S

Finalmente, un conjunto es acotado si puede ser contenido en una bola bola de radio suficientemente grande. Un conjunto es *compacto* si es cerrado y acotado. Note que el complememeto de un conjunto abierto en un conjunto cerrado (viceversa) y que los puntos de acumulación de cualquier conjunto y su complemento son el mismo.

una manera un poco más fácil de asimilar éstas definiciones es de forma gráfica considerando $S = \{(x, y) : x^2 + y^2 \leqslant 1\}$, este conjunto representa todos los puntos dentro del circulo con centro $C(0, 0)$ y radio $r = 1$. (Véase 4.2)

Fácilmente se puede verificar que S es cerrado, es decir, $\overline{S} = S$.
Además $\mathring{S}$ consiste en todos los puntos que están estrictamente dentro del circulo, esto es, $\mathring{S} = \{(x, y) : x^2 + y^2 < 1\}$.
Finalmente, S' consiste de los puntos en el circulo, esto es, $S' = \{(x, y) : x^2 + y^2 = 1\}$.



(a) Cerradura $\overline{S}$ (b) Interior $\mathring{S}$ (c) Acumulación S'

Figure 4: Topología del conjunto $S = \{(x, y) : x^2 + y^2 \leqslant 1\}$

Caracterizaciones de conjuntos

- Un conjunto S es cerrado si y solo si contiene todos sus puntos de acumualción, es decir $S' \subset S$.
- $\overline{S} = S \cup S'$ es el cerrado más pequeño que contiene al conjunto S.
- Un conjunto S es abierto si y solo si éste no contiene ninguno de sus puntos frontera, mas precisamente $S' \cap S = \emptyset$.
- $\mathring{S} \subseteq S$, por la tanto, tenemos que $\mathring{S} = S - S'$ mientras que necesariamente $S' \neq S - \mathring{S}$.

Un conjunto puede ser ni abierto ni cerrado, los únicos conjuntos que son abiertos y cerrados a la vez son el conjunto vacío ($\emptyset$) y el mismo $\mathbb{R}^n$. También consideremos que ningún punto $\overline{x} \in S$ puede ser un punto interior o de acumulación de S. Sin embargo $S \neq \mathring{S} \cup S'$, es decir que S no necesita contener sus puntos de acumulación.

Existe otra definición equivalente de conjunto cerrado, el cual es muy importante demostrando desde el punto de vista que es un conjunto cerrado.
Esta definición está basada en las sucesiones de puntos contenidos en S.

Un conjunto es cerrado si y solo si para cualquier sucesión convergente de puntos $\{x_k\}_{k\in\mathbb{N}} \in S$ con punto límite $\overline{x}$, así mismo tenemos que $\overline{x} \in S$.
La equivalencia de ésta y la definición previa de cerradura es fácilmente de ver ya que el punto límite $\overline{x}$ de cualquier sucesión convergente de puntos en S debe encontrarse en el interior o de acumulación de S, en otras palabras debería existir un $\varepsilon > 0$ tal que $\{x :\parallel x - \overline{x} \parallel < 0\} \cap S = \emptyset$ contradiciendo que $\overline{x}$ es el punto límite de una sucesión contenida en S.

Segmento de linea entre puntos del interior y cerradura de un conjunto [2]

Dado un conjunto convexo con interior no vacío, el segmento de linea (excluyendo los puntos finales) que une un punto del interior del conjunto con un punto de la clausura de este pertenece al interior del conjunto. Este resultado se muestra acontinuación.

Teorema 4.2. *Sea S un conjunto convexo en $\mathbb{R}^n$ con interior no vacío. Sea $x_1 \in \overline{S}$ y $x_2\mathring{S}$. Entonces $\lambda x_1 + (1-\lambda)x_2 \in \mathring{S} \quad \forall \quad \lambda \in (0,1)$*

Demostración:
Como $x_2 \in \mathring{S}$ existe un $\varepsilon > 0$ tal que $\{z :\parallel z - x_2 \parallel < \varepsilon\} \in S$. Sea y tal que

$$y = \lambda x_1 + (1-\lambda)x_2; \ \lambda \in (0,1) \tag{2}$$

Para probar que $y \in \mathring{S}$ es suficiente construir un vecindario sobre y que también pertenece a S. En particular mostraremos que $\{z :\parallel z - y \parallel < (1-\lambda)\varepsilon\}$. Sea z tal que $\parallel z - y \parallel < (1-\lambda)\varepsilon$ (véase 2), ahora bien, como $x_1 \in \overline{S}$

$$\left\{x :\parallel x - x_1 \parallel < \frac{(1-\lambda)\varepsilon - \parallel z - y \parallel}{\lambda}\right\} \cap S$$

es no vacío, en particular, existe $z_1 \in S$ tal que

$$\parallel z_1 - x_1 \parallel < \frac{(1-\lambda)\varepsilon - \parallel z - y \parallel}{\lambda} \tag{3}$$

Ahora, sea $z_2 = \dfrac{z - \lambda z_1}{1-\lambda}$ de (2), la desigualdad de Schwartz y (3) tenemos:

$$\begin{aligned}
\parallel z_2 - x_2 \parallel = \parallel \frac{z - \lambda z_1}{1-\lambda} - x_2 \parallel &= \parallel \frac{(z - \lambda z_1) - (y - \lambda x_1)}{1-\lambda} \parallel \\
&= \frac{1}{1-\lambda} \|(z - y) + (x_1 - \lambda x_1)\| \\
&\leqslant \frac{1}{1-\lambda} (\parallel z - y \parallel + \lambda \parallel x_1 - z_1 \parallel) \\
&< \varepsilon
\end{aligned}$$

Por lo tanto, $z_2 \in S$. Dada la definición de z_2 notemos que $z = \lambda z_1 + (1-\lambda)z_2$, como z_1 y z_2 pertenecen a S, entonces z también pertenece a S. Hemos mostrado que para cualquier z con $\parallel z - y \parallel < (1-\lambda)\varepsilon$ pertenece a S. Por lo tanto $y \in \mathring{S}$.

$\square$

Corolario 4.2. *Sea S un cojunto convexo. Entonces $\mathring{S}$ es convexo.*

Corolario 4.3. *Sea S un conjunto convexo con interior no vacío. Entonces $\overline{S}$ es convexa.*

Demostración
Asumiendo que $\mathring{S} \neq \emptyset$. Sea $x_1, x_2 \in \overline{S}$, por teorema tenemos que

$$\lambda x_2 + (1-\lambda)z \in \mathring{S} \ \ \forall \ \lambda \in (0,1)$$

Sea $\mu \in (0,1)$ fijo. Por teorema 4.2:

$$\mu x_1 + (1-\mu)[\lambda x_2 + (1-\lambda)z] \in \mathring{S} \subset S \ \ \forall \lambda \in (0,1)$$

Si tomamos el límite cuando λ se apróxima 1 se tiene que: $\mu x_1 + (1-\mu)x_2 \in \overline{S}$.

$\square$

Corolario 4.4. *Sea S un conjunto convexo con interior no vacío. Entonces $\overline{(\mathring{S})} = \overline{S}$.*

Demostración
Claramente $\overline{(\mathring{S})} \subseteq \overline{S}$. Sea $x \in \overline{S}$ escogemos $y \in \mathring{S}$ (asumieno que $\mathring{S} \neq \emptyset$). Entonces:

$$\lambda x + (1-\lambda)y \in \mathring{S} \ \ \forall \lambda \in (0,1)$$

Haciendo $\lambda \to 1^-$ se sigue que $x \in \overline{(\mathring{S})}$.

$\square$

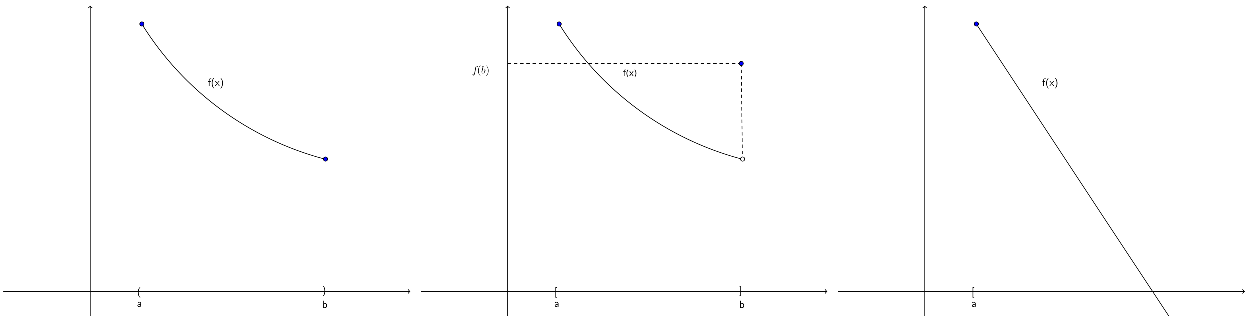
Corolario 4.5. *Sea S un conjunto convexo con interior no vacío. Entonces $(\mathring{\overline{S}}) = \mathring{S}$*

Demostración:
Note que $\mathring{S} \subseteq (\mathring{\overline{S}})$. Sea $x_1 \in (\mathring{\overline{S}})$, necesitamos mostrar que $x \in \mathring{S}$. Existe un $\varepsilon > 0$ tal que $\| y - x_1 \| < \varepsilon$ implica que $y \in \overline{S}$. Ahora, sea $x_2 \neq x_2 \in \mathring{S}$ y sea $y = (1+\delta)x_1 - \delta x_2$, donde $\delta = \dfrac{\varepsilon}{2\,\|\|}$. Como $\|\| = \dfrac{\varepsilon}{2}$, $y \in \overline{S}$. Pero

$$\lambda = \frac{1}{1+\delta} \in (0,1)$$

Como $y \in \overline{S}$ y $x_2 \in \mathring{S}$ entonces por teorema 4.2, $x_1 \in \mathring{S}$

$\square$



(a) S es no cerrado (b) f discontinua (c) f no acotada

Figure 5: Inexistencia de solución de minimización [2]

4.3 Teorema de Weierstrass

Un resultado muy importante y ampliamente utilizado se basa en los conceptos anteriores. Este resultado relaciona la existencia de una solución de minimización para un problema de optimización [2].

- Podemos decir que $\overline{x}$ es una solución de minimización al problema $\min\{f(x); x \in S\}$, siempre que $\overline{x} \in S$ y $f(\overline{x}) \leqslant f(x) \ \ \forall x \in S$. En tal caso, decimos que existe un mínimo.

- Por otra parte, decimos que $\alpha = \inf\{f(x)|x \in S\}$ si α es la mayor de las cotas inferiores de f en S; estos es: $\alpha \leqslant f(x) \ \ \forall \ x \in S$ es decir que no hay $\overline{\alpha} > \alpha$ tal que $\overline{\alpha} \leqslant f(x) \ \ \forall x \in S$.

- Similarmente, $\alpha = \max\{f(x)|x \in S\}$ si existe una solución $\overline{x} \in S$ tal que $\alpha = f(\overline{x}) \geqslant f(x) \ \ \forall x \in S$.

- Por otro lado, $\alpha = \sup\{f(x)|x \in S\}$ si α es la menor de las cotas superiores de f en S; esto es $\alpha \geqslant f(x) \ \ \forall x \in S$, y no hay otro $\overline{\alpha} < \alpha$ tal que $\overline{\alpha} \geqslant f(x) \ \ \forall x \in S$

En la Figura (5) se ilustran tres momentos en los que el mínimo no existe.
En la Figura (5a) el ínfimo de de f sobre (a, b) está dado por $f(b)$ pero como S no es cerrado y en particular $b \notin S$ el ínimo no existe.
En la Figura (5b) tenemos que $\inf\{f(x);\ x \in [a, b]\}$ esá dado por el límite de $f(x)$ cuando x tiende a b por la izquierda. $(\lim_{x \to b^-} f(x))$. Como f es discontinua en b no existe una solución de minimización.
La figura (5c) ilustra cuando f no es acotada sobre un conjunto $S = \{x : \ x \geqslant a\}$
Ahora probaremos formalmente este resultado cuando S es no vació, cerrado y acotado, f continua en S, con éstas condiciones el mínimo en la Figura 5 si existe.

Teorema 4.3. *Sea S un conjunto no vacío, compacto y sea $f \longmapsto \mathbb{R}$ continua en S entonces el problema $\min\{f(x) : x \in S\}$ alcanza su mínimo; esto es, existe una solucion de minimización para este problema.*

Demostración:

Como f es continua en S y S es cerrado y acotado, f también está acotada (recordar que f es continua), como $S \neq \emptyset$, existe una cota inferior (la más grande) $\alpha \equiv \min\{f(x) : x \in S\}$. Ahora, sea $0 < \varepsilon < 1$ y consideremos el conjunto $S_k = \{x \in S : \alpha \leqslant f(x) \leqslant \alpha + \varepsilon^k\}\ \forall\, k = 1, 2, \ldots$. Por la definición de mínimo se tiene que $S_k \neq \emptyset\ \forall\, k$, entonces, también podemos construir una sucesión de puntos $\{x_k\} \subseteq S$ seleccionando un punto $x_k \in S_k\ \forall\, k = 1, 2, \ldots$ Como S es acotado, existe existe una subseción convergente $\{x_k\}_K \to \overline{x}$ indexada por el conjunto K. Por la cerradura de S tenemos $\overline{x} \in S$; y por la continuidad f se tiene que como $\alpha \leqslant f(x) \leqslant \alpha + \varepsilon^k\ \forall\, k$ tenemos que:

$$\alpha = \lim_{k\to\infty;\ k\in K} f(x_k) = f(\overline{x})$$

Por lo tanto, hemos demostrado que existe una solución $\overline{x} \in S$ tal que $f(\overline{x}) = \alpha = \inf\{f(x) : x \in S\}$, entonces $\overline{x}$ es una solución de minimización.

□

4.3.1 Separación y soporte de conjuntos

Las nociones de hiperplano soporte y separación de conjuntos convexos son muy importantes en optimización. Casi todas las condiciones de optimalidad y dualidad se relacionan con algún tipo de separación o soporte de conjuntos convexos [2]. Los resultados que veremos acontinuación están basados en el siguiente hecho geométrico: Dado un conjunto cerrado S y un punto $y \in S$ existe un único punto $\overline{x} \in S$ con con una distancia mínima a y y un hiperplano que separa a y y S.

Distancia mínima de un punto a un conjunto convexo [2]

Para establecer el importante resultado anterior seguiremos la *ley del paralelogramo*.
Sean a y b dos vectores en $\mathbb{R}^n$. Entonces:

$$\begin{aligned} \| \vec{a} + \vec{b} \|^2 &= \| \vec{a} \|^2 + \| \vec{b} \|^2 + 2\vec{a^t}\vec{b} \\ \| \vec{a} - \vec{b} \|^2 &= \| \vec{a} \|^2 + \| \vec{b} \|^2 - 2\,\vec{a^t}\vec{b} \end{aligned}$$

Sumando ambas se tiene:

$$\| \vec{a} + \vec{b} \|^2 + \| \vec{a} - \vec{b} \|^2 = 2 \| \vec{a} \|^2 + 2 \| \vec{b} \|^2$$

Este resultado se ilustra en la Figura (6) y puede ser representado de la siguiente manera: La suma de las normas al cuadrado de las diagonales de un paralelogramo es igual al doble de la suma de sus normas al cuadrado.
De la teoría básica de geometría en espacios vectoriales normados se tiene que un hiperplano cerrado H en X puede ser reperesentado por: [1]

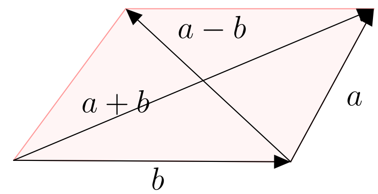


Figure 6: Ley del paralelogramo

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

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

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

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

Teorema 4.4. *Sea $C \subset X$ un conjunto convexo cerrado no vacío del espacio vectorial normado X. Entonces, cada elemento $u \notin C$ puede ser fuertemente separado de C por un hiperplano cerrado, es decir,*

$$\exists\, z^* \in X^*,\ z^* \neq 0,\ \exists \alpha \in \mathbb{R} \textit{ tal que } \langle u, z^* \rangle > \alpha \textit{ y } \langle x, z^* \rangle \leqslant \alpha \ \forall x \in C$$

Para la demostracion del teorema es necesario conocer los resultados del teorema de Hanh-Banach, el cual dice:

Teorema 4.5. *:* **(Primera forma geométrica del Teorema de Hahn-Banach)**
Sean $A \subset C$ y $B \subset C$ subconjuntos convexos no vacíos tales que $A \cap B = \emptyset$. Supongamos que uno de ellos es abierto. Entonces existe un hiperplano cerrado que separa A y B.

Teorema 4.6. *:* **(Segunda forma geométrica del teorema de Hahn-Banach)**
Sean $A \subset C$ y $B \subset C$ subconjuntos convexos no vacíos tales que $A \cap B = \emptyset$. Supongamos que A es cerrado y B es compacto. Entonces existe un hiperplano cerrado que separa estrictamente A y B.

Demostración: (Teorema 4.4)
La demostración está dada cuando el espacio vectorial X es un espacio de Hilbert. Para el caso general, se utiliza la versión analítica del teorema de Hahn-Banach el cual es una consecuencia del lema de Zorn.

Recordemos que si X es un espacio de Hilbert, entonces X^* se identifica con X y se tiene que $\| x \|^2 = \langle x, x \rangle$ (producto interno "=" producto de dualidad).

Sea $\text{Proy}_c(u)$ la proyección de $u \in X$ sobre el conjunto C (ésta existe y es única). $\text{Proy}_c(u)$ está caracterizada por

$$\begin{cases} \langle u - \text{Proy}_c(u), x - \text{Proy}_c(u)\rangle & \leqslant & 0, \ \forall\, x \in C \\ \text{Proy}_c(u) & \in & C \end{cases}$$

Por otro lado

$$\begin{aligned} 0 < \| z^* \|^2 &= \langle z^*,\ z^* \rangle \\ &= \langle u,\ z^* \rangle - \langle \text{Proy}_C(u),\ z^* \rangle && (4) \\ \Longrightarrow & \ \langle u,\ z \rangle > \langle \text{Proy}_C(u),\ z^* \rangle && (5) \end{aligned}$$

Tomemos $\alpha := \langle z,\ \text{Proy}_C(u)\rangle$. Combinando (2) y (3) se obtiene:

$$\sup_{x \in C} \langle x,\ z^* \rangle \leqslant \alpha < \langle u,\ z^* \rangle$$

Lo cual demuestra el resultado.

□

4.4 Funciones convexas

Funciones convexas y cóncavas tienen muchas propiedades importantes y especiales. Por ejemplo: cualquier mínimo local de una función convexa sobre un conjunto convexo es también un mínimo global. Acontinuación introduciremos temas importantes de funciones convexas y desarrollaremos algunas de sus propiedades.

Definición 4.6. *Sea $f : S \longmapsto \mathbb{R}$, donde S es un conjunto convexo no vacío de $\mathbb{R}^n$.*

- *La función S se dice que es* ***convexa*** *en S si:*

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

para cada $x_1, x_2 \in S$ y para cada $\lambda \in (0, 1)$.

- *La función f es* ***estrictamente convexa*** *en S si la desigualdad es estricta:*

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

para cada $x_1, x_2 \in S$ y para cada $\lambda \in (0, 1)$.

- *La función f es llamada* ***cóncava (estrictamente cóncava)*** *en S si $-f$ es convexa (estrictamente convexa) en S.*

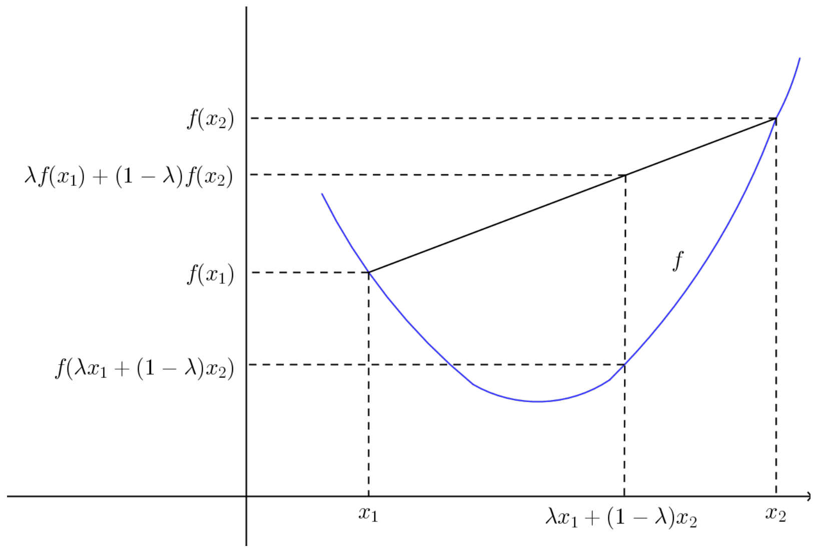


Figure 7: Interpretación geométrica de función convexa en $\mathbb{R}$ [5]

- *La función f diferenciable es* ***pseudo-convexa*** *si:*

$$f(x_1) < f(x_2),\ x_1 \neq x_2 \Longrightarrow \nabla f(x_1)(x_2 - x_1) < 0$$

Sean x_1 y x_2 dos puntos en el dominio de f. Entonces una función f es convexa, si el segmento de recta que une dos puntos $[x_1, f(x_1)]$ y $[x_2, f(x_2)]$ del gráfico de f está por encima de la gráfica de la función de f, como lo muestra la Figura(7)
Los siguientes son algunos ejemplos de funciones convexas.

1. $f(x) = 3x + 4$
2. $f(x) = |x|$
3. $f(x) = x^2 - 2x$
4. $f(x) = (-x)^{\frac{1}{2}}$
5. $f(x_1, x_2) = 2x_1^2 + 2x_2^2 - 2x_1x_2$
6. $f(x_1, x_2, x_3) = x_1^4 + x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3$

Note que en cada uno de los ejemplos las funciones son convexas sobre $\mathbb{R}^n$. Excepto para el ejemplo 4, la función no está definida para $x < 0$. Se pueden construir ejemplos de funciones que son convexas en una región pero no sobre $\mathbb{R}^n$. Por ejemplo $f(x) = x^3$ no es convexa sobre $\mathbb{R}$, pero es convexa sobre $S = \{x : x \geqslant 0\}$. Un ejemplo de ello es $f(x) = x^3$ no es convexa sobre $\mathbb{R}$ pero lo es sobre $S = \{x : x \geqslant 0\}$.

De ahora en adelante nos concentraremos en funciones convexas, pues el resultado para funciones cóncavas puede ser obtenido fácilmente notando que f es cóncava si y sólo si $-f$ es convexa.

Un conjunto asociado a una función convexa f es $S_\alpha = \{x \in S : f(x) \leqslant \alpha\}$, $\alpha \in \mathbb{R}$, usualmente referido como *conjunto de nivel*. A veces este conjunto es llamado un *conjunto subnivel*, para diferenciarlo del *conjunto supernivel* $\{x \in S : f(x) \geqslant \alpha\}$ note que éstos tienen propiedades a las de una función cóncava.

Teorema 4.7. *Sea S un conjunto convexo no vacío en $\mathbb{R}^n$ y sea $f : S \longmapsto \mathbb{R}$ una función convexa. Entonces el conjunto de nivel $S_\alpha = \{x \in S : f(x) \leqslant \alpha\}$; $\alpha \in \mathbb{R}$ donde es un número real, es un conjunto convexo.*

Demostración:
Sean $x_1, x_2 \in S_\alpha$. Entonces $x_1, x_2 \in S$ y $f(x) \leqslant \alpha$ y $f(x) = \alpha$. Ahora, sea $\lambda \in (0,1)$ y $x = \lambda x_1 + (1-\lambda)x_2$. Por la convexidad de S se tiene que $x \in S$, por otro lado, dada la convexidad de f se tiene:

$$f(x) = \lambda f(x_1) + (1-\lambda)f(x_2) \leqslant \lambda\alpha(1-\lambda)\alpha = \alpha$$

Por lo tanto, $x \in S_\alpha$, y por consiguiente, S_α es convexo.

$\square$

Continuidad de funciones convexas

Una propiedad importante de funciones convexas y cóncavas es que son continuas en el interior de su dominio.

Teorema 4.8. *Sea S un conjunto convexo no vacío en $\mathbb{R}^n$ y sea $f : S \longmapsto \mathbb{R}^n$ una función convexa. Entonces f es continua en el interior de S.*

Note que las funciones convexas y cóncavas pueden no ser continuas en todo lugar. Sin embargo, por el Teorema (4.8), los puntos de discontuindad solo son permitidos en la frontera de S, como se ilustra en la siguiente función convexa definida en $S = \{x : -1 \leqslant x \leqslant 1\}$:

$$f(x) = \begin{cases} x^2 & \text{para} \quad |x| < 1 \\ 2 & \text{para} \quad |x| = 1 \end{cases}$$

4.5 Subgradiente

Las funciones convexas poseen algunas propiedades usuales de diferenciabilidad, y una de ellas es la existencia de la derivada direccional lateral (derecha e izquierda) en toda dirección en un punto interior de su dominio, así como en el caso usual de la derivada direccional de una función diferenciable puede ser escrita en términos del vector gradiente, que está asociado con un hiperplano tangente al gráfico de la función, la derivada lateral derecha de una función convexa, no necesariamente diferenciable puede ser escrita en términos de vectores subgradientes que están asociados con hiperplanos soportes al epigrafo de la función. En lo sucesivo nos referiremos a la derivada direccional derecha como derivada direccional [10].

Definición 4.7. *Sean $f : \mathbb{R}^n \longmapsto \mathbb{R}$, $x \in \mathbb{R}^n$; la derivada direccional de f en x con respecto a y es definida como el siguiente liímite:*

$$f'(x) = \lim_{t \to 0^+} \frac{f(x+ty) - f(x)}{t}$$

si este límite existe.

Hipografo y epigrafo de una función

Una función f en S puede ser completamente descrita por el conjunto $\{[x, f(x)] : x \in S\} \subset \mathbb{R}^n$, cada cual se refiere ***grafo*** de la función. Se puede construir uno de dos conjuntos que están relacionados al grafo de la función f : el *epigrafo* y el *hipografo* de f.

Definición 4.8. *Sea S un conjunto no vacío en $\mathbb{R}^n$, y sea $f : S \longmapsto \mathbb{R}$. El epigrafo de f denotado por epi(f), es un subconjunto de $\mathbb{R}^{n+1}$ definido por:*

$$\{(x, y) : x \in S, y \in \mathbb{R}, y \geqslant f(x)\}$$

El hipografo de f denotado por hyp(f) es un subconjunto de $\mathbb{R}^{n+1}$ definido por:

$$\{(x, y) : x \in S, y \in \mathbb{R}, y \leqslant f(x)\}$$

Teorema 4.9. *Sea S un conjunto convexo no vacío en $\mathbb{R}^n$, y sea $f : S \longmapsto \mathbb{R}$. Entonces f es convexa si y solo si el epi(f) es un conjunto convexo.*

Definición 4.9. *Sean $f : \mathbb{R}^n \longmapsto \mathbb{R},\ x_0 \in \mathbb{R}^n$ Un vector $\vec{s} \in \mathbb{R}^n$ tal que:*

$$f(x) \geqslant f(x_0) + \langle s, x - x_0 \rangle,\ \forall x \in \mathbb{R}^n$$

es llamado un subgradientede f en x_0 [10]

El conjunto de subgradientes de una función f en el punto x_0 es llamado subdiferencial de f en x_0 y se denota por $\partial f(x_0)$.

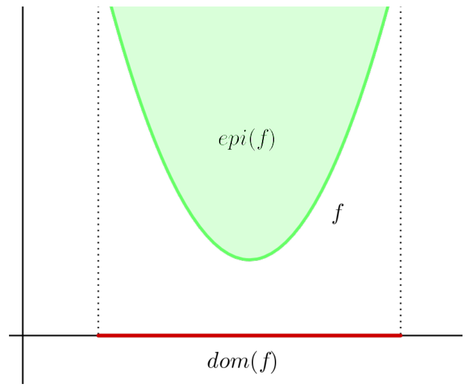


Figure 8: Epigrafo de una función f en $\mathbb{R}$ [8]

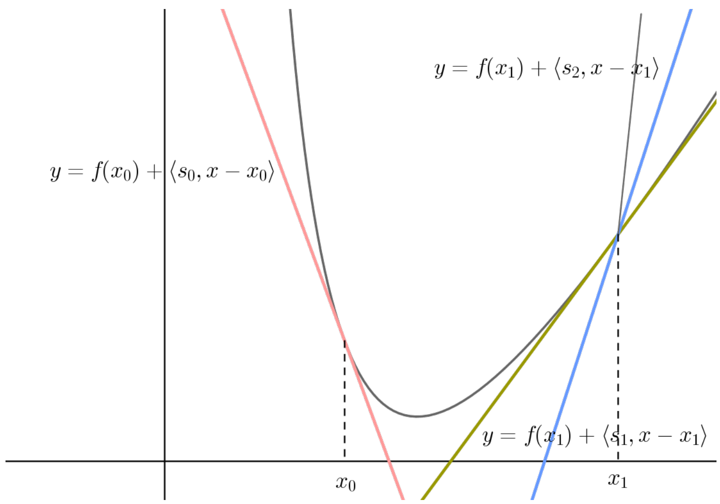


Figure 9: Interpretación geométrica del subgradiente en $\mathbb{R}$ [10]

Definición 4.10. *Si $\partial f(x_0) \neq \emptyset$ diremos que f es diferenciable en x_0.*

Geométricamente, $s \in \mathbb{R}^n$ es un subgradiente de la función convexa f en x_0 si el gráfico de f en $\mathbb{R}^{n+1}$ está por encima del gráfico del hiperplano $y = f(x) + \langle s, x - x_0\rangle$. Como $(x_0, f(x_0))$ está en el hiperplano, este hiperplano constituye un hiperplano soporte a $epi(f)$ en $(x_0, f(x_0))$. Esto es, la existencia de un vector subgradiente establece la existencia de un hiperplano soporte no vertical a $epi(f)$ en $(x_0, f(x_0))$.
El siguiente teorema establece que toda función convexa es subdiferenciable en el interior de su dominio [10].

Teorema 4.10. *Sea f una función convexa definida en un convexo S, entonces en cada punto $x_0 \in \mathring{S}$ se tiene que $\partial f(x_0) \neq \emptyset$*

El recíproco del teorema anterior es falso, pues la condición de ser punto interior es indispensable, por ejemplo, si se considera la función convexa $f : [0, \infty) \longmapsto \mathbb{R}$ dada por $f(x) = -\sqrt{x}$ entonces se tiene que $\partial f(0) = \emptyset$

4.6 Funciones convexas diferenciables

Definición 4.11. *[2] Sea S un conjunto no vacío en $\mathbb{R}^n$, y sea $f : S \longmapsto \mathbb{R}$. Entonces f se dice que es diferenciable en $\overline{x} \in \mathring{S}$ si existe un vector $\nabla f(\overline{x})$, llamado vector gradiente; y una función $\alpha : \mathbb{R}^n \longmapsto \mathbb{R}$ tal que:*

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^t(x - \overline{x}) + \parallel x - \overline{x} \parallel \alpha\langle \overline{x}; x - \overline{x}\rangle \ \ \forall\, x \in S$$

dónde $\lim_{x \to \overline{x}} \alpha\langle \overline{x}, x - \overline{x}\rangle = 0$.

La función f se dice que es diferenciable en $S' \subseteq S$ si ésta es diferenciable en cada punto de S'. Esta representación de f es llamada *expansión (serie de Taylor) de primer orden* de f en el punto $\overline{x}$.
Note que si f es diferenciable en $\overline{x}$, solo podría haber un vector gradiente el cual está dado por:

$$\nabla f(\overline{x}) = \left(\frac{\partial f(\overline{x})}{\partial x_1}, \ldots, \frac{\partial f(\overline{x})}{\partial x_n}\right)^t \equiv (f_1(\overline{x}), \ldots, f_n(\overline{x}))^t$$

dónde $f_i(\overline{x}) = \dfrac{\partial f(\overline{x})}{\partial x_i}$ es la derivada parcial de f respecto a x_i en $\overline{x}$.

Teorema 4.11. *Sea f una función convexa definida en un conjunto convexo S de $\mathbb{R}^n$ y sea x un punto interior de S. f es diferenciable en x si y sólo si posee un único subgradiente en x. En tal caso $\partial f(x) = \{\nabla f(x)\}$.*

Teorema 4.12. *Sea S un conjunto abierto no vacío en $\mathbb{R}^n$ y sea $f : S \longmapsto \mathbb{R}$ diferenciable en $\mathbb{R}$. Entonces f es convexa si y sólo si para cualquier $\overline{x} \in S$ se tiene:*

$$f(x) \geqslant f(\overline{x}) + \nabla f(\overline{x})^t(x - \overline{x}) \ \ \forall x \in S$$

De igual forma, f es estrictamente convexa si y sólo si para cada $x \in S$ se tiene:

$$f(x) > f(\overline{x}) + \nabla f(\overline{x})^t(x - \overline{x}) \ \ \forall x \neq \overline{x} \in S$$

El siguiente teorema da una carcterización suficiente y necesaria para funciones convexas. Para una función f de una variable, la caracterización se reduce a una pendiente creciente.

Teorema 4.13. *Sea S un conjunto abierto no vacío en $\mathbb{R}^n$ y sea $f : S \longmapsto \mathbb{R}$ diferenciable en S. Entonces f es convexa si y sólo si para cada $x_1, x_2 \in S$ se tiene:*

$$[\nabla f(x_2) - \nabla f(x_1)]^t(x_2 - x_1) \geqslant 0.$$

Similarmete, f es esrictamente convexa para cualquier $x_1, x_2 \in S$ distintos, entonces se tiene:

$$[\nabla f(x_2) - \nabla f(x_1)]^t(x_2 - x_1) \geqslant 0.$$

Aunque los teoremas (4.12) y (4.13) dan una caracterización suficiente y necesaria para funciones convexas, comprobar estas condiciones desde el punto de vista computacional es muy difícil. Una caracterización más simple y manejable, al menos para funciones cuadráticas puede obtenerse siempre que la función sea dos veces diferenciable.

Funciones dos veces diferenciables [2]

Una función f que es diferenciable en $\overline{x}$ se dice que es dos veces diferenciable en $\overline{x}$ si la representación de la *expansión de segundo orden (serie de Taylor)* de la siguiente definición existe.

Definición 4.12. *Sea S un conjunto no vacío en $\mathbb{R}^n$ y sea $f : S \longmapsto \mathbb{R}$. Entonces f se dice que es dos veces diferenciable en $\overline{x} \in \mathring{S}$ si existe un vector $\nabla f(\overline{x})$, y una matriz simétrica $H(\overline{x})$ de $n \times n$, llamada **matriz hessiana** y una función $\alpha : \mathbb{R}^n \longmapsto \mathbb{R}$ tal que:*

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

para cada $x \in S$, donde $\lim\limits_{x \longmapsto \overline{x}} \alpha\langle \overline{x}, x - \overline{x}\rangle = 0$.
La función f se dice dos veces diferenciable en el conjunto abierto $S' \subseteq S$ si es dos veces diferenciable en cada punto de S'.

Cabe señalar que para funciones dos veces diferenciables, la matriz hessiana $H(\overline{x})$ está compuesta por las derivadas parciales de orden dos $f_{ij}(\overline{x}) \equiv \dfrac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j}$; $i = 1, \ldots, n$. $j = 1, \ldots, n$ y está dada de la siguiente forma:

$$H(\overline{x}) = \begin{bmatrix} f_{11}(\overline{x}) & f_{12}(\overline{x}) & \ldots & f_{1n}(\overline{x}) \\ f_{21}(\overline{x}) & f_{22}(\overline{x}) & \ldots & f_{2n}(\overline{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}(\overline{x}) & f_{n2}(\overline{x}) & \ldots & f_{nn}(\overline{x}) \end{bmatrix}$$

4.7 Mínimo y máximo de funciones convexas

En esta parte se considerarán problemas de minimizar y maximizar una función convexa sobre un conjunto convexo y se desarrollaran las condiciones necesarias y suficientes para optimalidad [2].

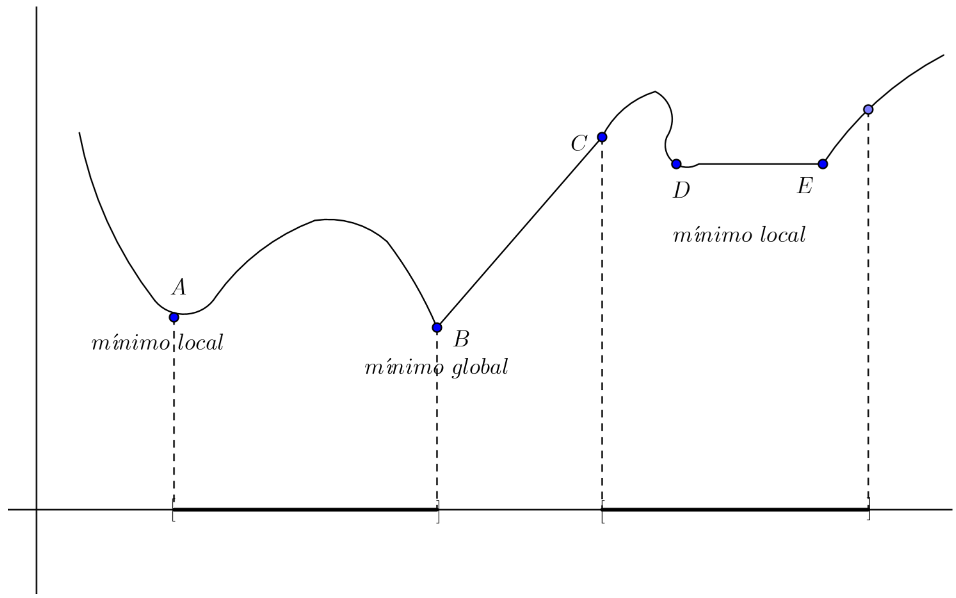


Figure 10: Mínimo local y global [2]

Definición 4.13. *Sea $f : \mathbb{R}^n \longmapsto \mathbb{R}$ y consideremos el problema a minimizar $f(x)$ sujeto a $x \in S$. Un punto $x \in S$ es llamado* ***solución factible*** *para el problema. Si $\overline{x} \in S$ y $f(x) \geqslant f(\overline{x})$ para cada $x \in S$, $\overline{x}$ es llamado una* ***solución optima, solución optima global*** *o simplemente una* ***solución*** *para el problema.*

La colección de soluciones óptimas es llamada ***solución óptima alternativa.***

- Si $\overline{x} \in S$ y si existe un $\varepsilon-$vecindario $N_\varepsilon(\overline{x})$ alrededor de $\overline{x}$ tal que $f(x) \geqslant f(\overline{x})$ para cada $x \in S \cap N_\varepsilon(\overline{x})$, $\overline{x}$ es llamado ***solución óptima local.***
- Si $\overline{x} \in S$ si $f(x) > f(\overline{x})$ para todo $x \in S \cap N_\varepsilon(\overline{x})$, $x \neq \overline{x}$ para algún $\varepsilon > 0$, $\overline{x}$ es llamada ***solución óptima local estricta***.
- Si $\overline{x} \in S$ es el único mínimo local en $S \cap N_\varepsilon(\overline{x})$, para algún vecindario $N_\varepsilon(\overline{x})$ alrededor de $\overline{x}$, x es llamada ***solución óptima local fuerte o aislada***

Todos estos tipos de óptimo local o mínimo local a veces también se denominan como ***mínimo relativo.*** La Figura(10) ilustra ejemplos de mínimos locales y globales para el problema de minimizar $f(x)$ sujeto a $x \in S$

Teorema 4.14. *Sea S un conjunto convexo no vacío en $\mathbb{R}^n$ y sea $f : S \longmapsto \mathbb{R}$ convexa en S. Considere minimizar el problema $f(x)$ sujeto a $x \in S$. Suponga que $\overline{x} \in S$ es una solución óptima local para el problema, entonces:*

1. *$\overline{x}$ es una solución óptima global.*
2. *Si $\overline{x}$ es una solución mínima estricta o f es estrictamente convexa, $\overline{x}$ es la única solución óptima global y también es un mínimo local.*

Acontinuación, se presenta una condición necesaria y suficiente para la existencia de una solución global. Si tal solución óptima no existe, entonces $\inf\{f(x) : x \in S\}$ es finito, pero no se alcanza en ningún punto de S ó es igual a $-\infty$.

Teorema 4.15. *Sea f una función convexa en $\mathbb{R}^n$ y S un conjunto convexo en $\mathbb{R}^n$ Consideremos el problema de optimización*

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

Entonces $\overline{x}$ es un mínimo local de f sobre S si y sólo si existe $s \in \partial f(\overline{x})$ tal que

$$\langle s, x - \overline{x} \rangle \geqslant 0 \quad \textit{para toda } x \in S$$

Maximizar una función convexa

Se desarrollará una condición necesaria para un máximo de una función convexa sobre un conjunto convexo. Desafortunadamene, esta condición no es suficiente. Por lo tanto, es improbable, que varios máximos que satisfagan la condición del Teorema 4.16 existan. En el caso de mimizar, no existe información local de tales soluciones que podrían llevar a mejores puntos. Por lo tanto, maximizar una función convexa es una tarea mucho más dura que minimizar una función convexa.

Teorema 4.16. *Sea $f : \mathbb{R}^n \longmapsto \mathbb{R}$ una función convexa y sea S un conjunto convexo no vacío en $\mathbb{R}^n$. Considere el problema:*

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

Si $\overline{x} \in S$ es una solución óptima local

$$\langle s, x - \overline{x} \rangle \leqslant 0$$

para cada $x \in S$, donde $s \in \partial f(\overline{x})$

5 Sección II: Condiciones de optimalidad y dualidad

5.1 Condicion de optimalidad de Fritz John

Teorema 5.1. ***Condiciones necesarias de Fritz John [2]***
Sea X un conjunto abierto no vacío en $\mathbb{R}^n$ y sea $f : \mathbb{R}^n \longmapsto \mathbb{R}$ y $g_i : \mathbb{R}^n \longmapsto \mathbb{R}$ para $i = 1, \ldots, m$. Considere el Problema P para minimizar $f(x)$ sujeto a $x \in X$ y $g_i(x) \leqslant 0$ para $i = 1, \ldots, m$. Sea $\overline{x}$ una solución factible y denotada por $I = \{i : g_(\overline{x}) = 0\}$. Por lo tanto, suponga que f y g_i para $i \in I$son diferenciables en $\overline{x}$ y que g_i para $i \notin I$ son continuas en $\overline{x}$. Si $\overline{x}$ resuelve el problema P localmente, existen escalares u_0 y u_i para $\imath \in I$ tales que:

$$\begin{aligned} u_0 \nabla f(\overline{x}) + \sum_{i \in I} u_i \nabla g_i(\overline{x}) &= 0 \\ u_0, u_i &\geqslant 0 \quad \textit{para } i \in I \\ (u_0, u_I) &\neq (0, 0) \end{aligned}$$

donde u_I es el vector cuyas componentes sin u_i para $i \in I$. Por lo tanto, si g_i para $i \notin I$ casi son diferenciables en $\overline{x}$, las codiciones anteriores pueden escribirse de forma equivalente:

$$\begin{aligned} u_0 \nabla f(\overline{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\overline{x}) &= 0 \\ u_i g_i(\overline{x}) &= 0 \quad \textit{para } i = 1, \ldots, m \\ u_o, u_i &\geqslant 0 \quad \textit{para } i = 1, \ldots, m \\ (u_0, u) &\neq (0, 0) \end{aligned}$$

donde u es el vector cuyas componentes son u_i para $i = 1, \ldots, m$.

Teorema 5.2. ***Condiciones suficientes de Fritz John [2]***
Sea X un conjunto abierto no vacío en $\mathbb{R}^n$ y sea $f : \mathbb{R}^n \longmapsto \mathbb{R}$ y $g_i : \mathbb{R}^n \longmapsto \mathbb{R}$ para $i = 1, \ldots, m$. Considere el Problema P para minimizar $f(x)$ sujeto a $x \in X$ y $g_i(x) \leqslant 0$ para $i = 1, \ldots, m$. Sea $\overline{x}$ una solución FJ denotada por $I = \{i : g_i(\overline{x}) = 0\}$. Se define S como la región relajada y factible par el problema P en el que se eliminan las restricciones no vinculantes.

a. Si existe un $\varepsilon-$vecindario $N_\varepsilon(\overline{x})$, $\varepsilon > 0$, tal que f es seudoconvexa sobre $N_\varepsilon(\overline{x}) \cap S$, $\overline{x}$ es un mínimo local para el problema P.

b. Si f es seudoconvexa en $\overline{x}$ y si g_i, $i \in I$ ambas son estrictamente pseudoconvexas y cuasiconvexas en $\overline{x}$, entonces $\overline{x}$ es una solucíon óptima global para el problema P. En particular, si éstas suposiciones de convexidad generalizada sólo son válidas restringiendo el dominio de f para $N_\varepsilon(\overline{x})$ para algún $\varepsilon > 0$, $\overline{x}$ es un mínimo local para el problema P.

Condiciones de Karush-Kuhn-Thucker [2]

Teorema 5.3. ***(Condiciones necesarias de Karush-Kuhn-Thucker KKT)***
Sea X un conjunto abierto no vacío en $\mathbb{R}^n$ y sea $f : \mathbb{R}^n \longmapsto \mathbb{R}$ y $g_i : \mathbb{R}^n \longmapsto \mathbb{R}$ para $i = 1, \ldots, m$. Considere el problema P para minimizar $f(x)$ sujeto a $x \in X$ y $g_i(x) \leqslant 0$, para $i = 1, \ldots, m$. Sea $\overline{x}$ una solución factible y denotada por $I = \{i : g_i(\overline{x}) = 0\}$. Suponga que f y g_i para $i \in I$ son diferenciables en $\overline{x}$ y que g_i para $i \notin I$ son continuas en $\overline{x}$. Por lanto, suponga que $\nabla g_i(\overline{x})$ para $i \in I$ son linealmente independientes. Si $\overline{x}$ resuelve localmente el problema P, existen escalares u_i para $i \in I$ tal que

$$\begin{array}{rcl} \nabla f(\overline{x}) + \sum\limits_{i \in I} \nabla g_i(\overline{x}) & = & 0 \\ u_i & \geqslant & 0 \quad \textit{para } i \in I \end{array}$$

Además de las suposiciones anteriores, si para cada g_i con $i \notin I$ es casi diferenciable en $\overline{x}$, las condiciones anteriores pueden escribirse de forma equivalente como:

$$\begin{array}{rcl} \nabla f(\overline{x}) + \sum\limits_{i=1}^{m} \nabla g_i(\overline{x}) & = & 0 \\ u_i g_i(\overline{x}) & = & 0 \quad \textit{para } i \in I \\ u_i & \geqslant & \quad \textit{para } i \in I \end{array}$$

5.2 Problemas con restricciones de igualdad y desigualdades

Acontinuación se generalizará las condiciones de optimalidad vistos anteriormente, para ello se sonsidera el siguiente problema de programación no lineal P :

$$\begin{array}{rcl} \min f(x) & & \\ \text{sujeto a } \; g_i(x) & \leqslant & 0 \quad \text{para } \; i = 1, \ldots, m \\ h_i(x) & = & 0 \quad \text{para } \; i = 1, \ldots, l \\ x & \in & X \end{array}$$

5.3 Condición suficiente y necesaria para problemas con restricciones de segundo orden

Basados en varias suposiciones de convexidad (generalizada), se tienen condiciones suficientes para garantizar que dada una solución que satisface las condiciones de óptimalidad de primer orden ésta es local o glabamente óptima. Análogo al caso sin restricciones, ahora se sabe que la derivada de segundo orden es una condición necesaria y suficiente para el problema restringido.[2]
Considere el problema:

$$P : \min\{f(x) : x \in S\} \tag{6a}$$

donde

$$S = \{x : g_i(x) \leqslant 0 \text{ para } i = 1, \ldots, m, \; h_i(x) = 0 \text{ para } i = 1, \ldots, l; \; x \in X\} \tag{6b}$$

Asuma que f, g_i para $i = 1, \ldots, m$ y h_i para $i = 1, \ldots, l$ están definidas en $\mathbb{R}^n \longmapsto \mathbb{R}$, ambas son diferenciables y X es un conjunto abierto no vacío en $\mathbb{R}^n$. La *función lagrangiana* para este problema se define como:

$$\phi(x, u, v) = f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{i=1}^{l} v_i h_i(x) \tag{7}$$

Esta función permite formular una teoría de dualidad para problemas de programación no lineal.

5.4 Dualidad lagrangiana

Dado un problema de programacion no lineal hay otro problema de programación no lineal asociado con el, el primero es llamado *problema primal,* y el otro es llamdo *problema dual lagrangiano.* Bajo ciertas condicines de convexidad y restricciones adecuadas, los problemas dual y primal poseen valores objetivos óptimos iguales, por lo tanto, es posible resolver el problema primal de manera indirecta resolviendo el problema dual [2].

Considere el siguiente problema de programación no lineal, el cual es llamado *problema primal*

Problema primal P:

$$\begin{array}{rlll} \text{minimizar} & f(x) & & \\ \text{sujeto a} & g_i(x) \leqslant 0 & \text{para} & i = 1, \ldots, m \\ & h_i(x) = 0 & \text{para} & i = 1, \ldots, l \\ & x \in X & & \end{array}$$

El *problema dual lagrangiano* se indica acontinuación.

Problema dual lagrangiano

$$\begin{aligned} \text{maximizar} \quad & \theta(u,v) \\ \text{sujeto a} \quad & u \geqslant 0 \\ \text{donde} \quad & \theta(u,v) = \ \inf f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{i=1} l v_i h_i(x) : \ x \in X \end{aligned}$$

Note que la función dual lagrangiana θ puede tomar el valor de $-\infty$ para algunos vectores (u,v). Como el problema dual consiste en maximizar el ínfimo (la mayor de las cotas inferiores) de la función $f(x)+\sum_{i=1}^{m} u_i g_i(x) + \sum_{i=1}^{l} v_i h_i(x)$, por lo cual a veces se refieren a éste como *dual viable.*

El siguiente teorema muestra que el valor objetivo de cualquier solución factible para el problema dual produce una cota inferior en el valor objetivo de cualquier solución factible para el problema primal.

Teorema 5.4. ***Dualidad débil***
Sea x una solución factible para el problema P; esto es $x \in X$, $g(x) \leqslant 0$ y $h(x) = 0$. Sea (u,v) una solución factible para el problema D; esto es $u \geqslant 0$. Entonces $f(x) \geqslant \theta(u,v)$.

5.5 Métodos para optimización no diferenciable

En esta sección se describirán los métodos estudiados para la minimización de funciones que son convexas pero no diferenciables. De igual forma se darán las pruebas de la convergencia de cada método [3].

5.6 Método del subgradiente

Este es el método más sencillo y es muy parecido al método de gradiente para funciones diferenciables pero cuenta con algunas modificaciones [3].

1. El método de subgradiente se aplica directamente a la función no diferenciable.
2. Los tamaños de paso no son elegidos por búsqueda lineal. En la mayoría de los casos, el tamaño de paso es fijo.
3. A diferencia del método de gradiente, el método de subgradiente no es un método de descenso; el valor de la función puede incrementar.

En este caso solo se tratará el problema de minimizar una función sin restricciones, es decir, minimizar $f : \mathbb{R}^n \longmapsto (0, +\infty]$ donde f es convexa.
El método de subgradiente usa la iteración:

$$x^{(k+1)} = x^k - \alpha_k g^k$$

x^k : es el punto resultante de la $k-$ésima iteración.
g^k : es cualquier subgradiente de f en x^k.
$\alpha_k > 0$: es el $k-$ésimo tamaño de paso.

Así, en cada iteración del método de subgradiente, se hace el paso en la dirección del subgradiente negativo.

Cuando la función es diferenciable, la única elección posible de g^k es $\nabla f(x^k)$, y el método del subgradiente se reduce al método de gradiente (excepto por la elección del tamño de paso). Un problema en el método de subgradiente es que se toma cualquier subgradiente en cada iteración, por lo tanto, no se toma en cuenta cual es la elección de subgradiente que hace decresec el valor de la función en el nuevo punto.

5.7 Métodos de plano de corte

El método del plano cortante de Kelley y Chener Goldstein se utiliza para problemas en optimización convexa, no necesariamente diferenciable. Este método se basa en el siguiente resultado: Sea $f : \mathbb{R}^n \longmapsto \mathbb{R}$ una función convexa y $x \in \mathbb{R}$, entonces:

$$f(x) = \max_{y \in \mathbb{R}^n} \{f(y) + \langle s, x - y \rangle : \, s \in \partial f(y)\}$$

Véase la Definición 4.9 de subdiferencial.

Consideremos el problema:

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

Asumamos que M es un conjunto cerrado y que para cada $x \in M$ un subgradiente de f en x puede ser calculado, el método de planos de corte consiste en resolver en la $k-$ésima iteración el problema

$$\min_{x \in M} f_k(x)$$

Esto es, ̃la función objetivo f es reemplazada por una aproximación poliedral f_k, que es construida usando los puntos x_i y sus respectivos subgradientes s_i para $i = 0, 1, 2, \ldots, k-1$ es decir:

$$f(x) = \max\{f(x_0) + \langle x - x_0 \rangle, \, f(x_1) + \langle x - x_1, s_1 \rangle, \ldots, \, f(x_{k-1}) + \langle x - x_{k-1}, s_{k-1} \rangle\}$$

y x minimiza f_k sobre M es decir

$$f_k(x_k) = \min_{x \in M} f_k(x)$$

Este problema es equivalente al problema lineal:

$$\begin{gathered} \min z \\ f(x_i) + s_i(x - x_i) \leqslant z \quad i = 0, 1, 2, \ldots, k-1 \end{gathered}$$

El método de planos de corte esta basado en progresivos refinamientos de aproximaciones poliedrales del epigrafo de f.
Se asume que el mínimo del problema de arriba puede obtenerse para todo k. Para esos valores de k para los cuales esto no ocurre, el conjunto M puede ser reemplazado por un conjunto compacto que contenga al conjunto optimal, estas restriciones son llamadas restricciones de caja.

Cuanto más planos de corte son incrementados, más precisa es la aproximación. Para el caso particular de tener una función lineal por tramos, el método del planos cortante converge en una cantidad finita de pasos. Un problema encontrado en éstos métods es que las funciones afines en cada iteración se acrecientan dificultando la resolución del problema.

5.8 Método del radiente próximo [3]

Considerar el siguiente problema

$$\min h(x) = f(x) + g(x)$$

Donde $f : \mathbb{R}^n \longmapsto \mathbb{R}$ y $g : \mathbb{R}^n \longmapsto (-\infty, +\infty]$ son funciones propias, cerradas, convexas y f es diferenciable. En esta forma se puede partir el problema en dos, en donde una parte es diferenciable. La forma en la cual se parte el problema no es única y puede llevar a diferentes implementaciones del método del gradiente próximo.

El método del gradiente próximo es:

$$x_{k+1} = prox_{\lambda_{kg}}(x_k - \lambda_k \nabla f(x_k))$$

donde $\lambda_k > 0$ es el tamaño de paso.

6 Aplicaciones

6.1 El problema del portafolio [3]

El problema clásico de optimización de portafolio, en el que un agente escoge sus estrategias a modo de maximizar su utilidad esperada ha sido estudiado en profundidad por mucho tiempo. Sin embargo, solo hace poco ha surgido el interés por plasmar el hecho de que la selección de un modelo particular al momento de hacer la toma de decisiones es en sí riesgosa.

El problema del portafolio se define de la siguiente manera:

$$\begin{array}{ll} \min & \frac{1}{2}x^t \sum x + r^t x \\ \text{sujeto a} & \sum x_i = 1 \qquad x_i \geqslant 0 \end{array}$$

donde $\sigma \in \mathbb{R}^{n\times n}$ es la matriz de varianza covarianza y $r \in \mathbb{R}$. El problema del portafolio puede ser convertido al siguiente problema:

$$\min \quad \frac{1}{2}x^t \sum x + r^t x + \mathbb{I}_C$$

donde $\mathbb{I}_C$ es la función indicadora del conjunto $C = \{x \in \mathbb{R}^n | \sum x_i = 1.x_i \geqslant 0\}$. Se puede aplicar el método del gradiente próximo si se calcula el operador próximo de la función indicadora $\mathbb{I}_C$.

Definición 6.1. ***(Operador próximo)*** *Sea $f : \mathbb{R}^n \longmapsto [-\infty, +\infty)$ una función propia, convexa y cerrada. Se define el operador próximo $prox_f : \mathbb{R}^n \longmapsto \mathbb{R}^n$ de f como*

$$proy_f(v) = \arg\min_x \{f(x) + \frac{1}{2}||x - v||^2\} \tag{8}$$

donde $||\cdot||$ es la norma euclidiana.

Dada la definición(**??**), $prox_f(v)$ es un punto que se encuentra entre $x^* = \min\{f\}$ y cerca de v. Cuando f es la función indicadora de $C \subseteq \mathbb{R}^n$ es decir

$$\mathbb{I}_C(x) = \begin{cases} 0 & x \in C \\ \\ +\infty & x \notin C \end{cases}$$

entonces

$$prox_f(v) = \arg\min_{x\in C}\{||x - v||\} \tag{9}$$

Así el operador próximo se puede ver como una proyección generalizada.
Usando (9) se puede obtener éste operador conociendo como se proyecta sobre el conjunto C. La proyección sobre el conjunto C C. La proyección sobre el conjunto C está dada por:

$$\prod_C(v) = \max v_i - \mu^*, 0$$

Donde μ es la raíz de la función

$$f = \sum_{i=1}^{n} \max v_i - \mu, 0 - 1$$

Utilizando (lo que falta), el operador próximo de $\mathbb{I}_C$ está dada por

$$\text{proy}_{\mathbb{I}_C}(v) = \max v_i - \mu^*, 0$$

Ejemplo [3]

Sea $P^1, \ldots, p^5$ los precios de cinco empresas de rama de construcción: Se tomaron los datos del $2009 - 2014$ de los precios diarios de las acciones. Los rendimientos diraios de la empresa i están dados por $r_i^j = \ln, \left(\frac{P_{j+1}^i}{P_j} \right)$. σ es la matriz de varianza covarianza muestral

$$\sigma = \begin{pmatrix} 0.0000778 & 0.00000796 & 0.000000645 & 0.0000541 & 0.00000346 \\ 0.00000796 & 0.000512 & 0.0000432 & 0.0000551 & 0.00000273 \\ 0.000000645 & 0.0000432 & 0.000315 & 0.000305 & 0.0000149 \\ 0.0000541 & 0.0000551 & 0.000305 & 0.0043 & 0.000116 \\ 0.00000346 & 0.00000273 & 0.0000149 & 0.000116 & 0.000208 \end{pmatrix}$$

y el vector $\vec{r}$ es la media muestral de los rendimientos de cada empresa

$$\vec{r} = (0.0004142, 0.0004127, 0.0018, 0.00411, 0.0008422)$$

aplicando el método del gradiente próximo se llegó al punto

$$x^* = (0, 0, 0.9999, 0, 0)$$

en diez iteraciones, en un tiempo de 0.2140 segundos. Lo que significa que todod el dinero se debe invertir en la acción de la empresa 3.

6.2 Procesamiento de imágenes: Transformada distancia [8].

En procesamiento de imágenes, la tranformada distancia, tiene diversas aplicaciones, por ejemplo, en medicina, visualización de estructuras o imágenes geométricas [8].

Para una imagen binaria B definida como una aplicación de $\{1, \ldots, n\} \times \{1, \ldots, m\}$ a $\{0, 1\}$, la transformada distancia es la función que asocia B con un *array* D de la forma $\{1, \ldots, n\} \times \{1, \ldots, m\}$ a $\{0, 1\}$, definida como: Asumimos que B tiene al menos un pixel p con $B[p] = 0$. Para cada pixel p en B, $D[p]$ contiene la distancia euclídea al pixel más cercano en B que contiene el valor 0. En la práctica, para restringir la computación aritmética entera, la transformada distancia D^2 es calculada

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{es} \quad D^2 = \begin{bmatrix} 2 & 1 & 2 & 5 & 8 \\ 1 & 0 & 1 & 4 & 5 \\ 2 & 1 & 2 & 1 & 2 \\ 5 & 4 & 1 & 0 & 1 \\ 8 & 5 & 2 & 1 & 2 \end{bmatrix}$$

Varios algoritmos han sido introducidos para calcular la transformada distancia (Euclidean Distance Transform) ó EDT. Recientes investigaciones se centran en simplificar los algoritmos aun alcanzando complejidad de tiempo lineal. Por ejemplo, el hecho que la computación de la EDT es equivalente a calcular la menor envoltura de funciones cuadráticas y con ello se presentó un algoritmo de complejidad lineal. Otros algoritmos basados en la monotonicidad o propiedades de vecindad son manejados para alcanzar también complejidad lineal.

La relación entre la envoltura de Moreau y la conjugada de Fenchel-Legendre para reducir el cálculo de la transformada distancia usando el algoritmo LLT es esencia, como sigue:

El cuadrado de la transformada distancia es la aplicación D^2 desde $\{1, \ldots, n\} \times \{1, \ldots, m\}$ al conjunto de enteros no negativos definidos por

$$D^2(p) = \min_{q \in O} \| p - q \|^2$$

Dónde $O = \{q : B(q) = 0\}$ es el conjunto de pixeles con valor 0 en B. Usando la función indicador $I(p) = 0$ si $B(p) = 0$ y ∞ en otros casos, podemos encontrar el cuadrado de la transformada distancia de Moreau de I

$$D^2(p) = \min_{q \in O} \| p - q \|^2 + I(q)$$

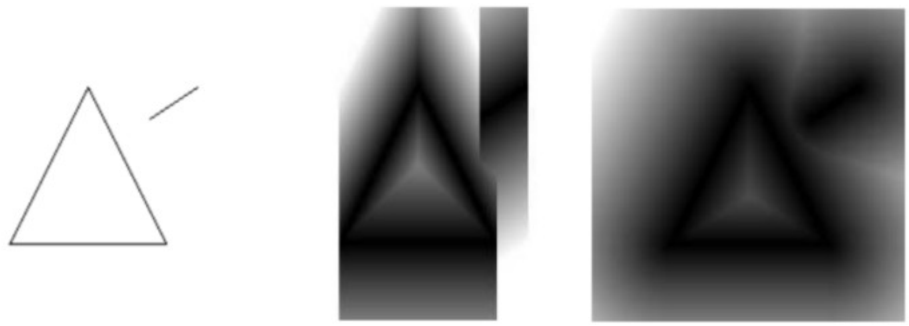
Así los algoritmos para la transformada distancia son casos particulares de los algoritmos para calcular la envoltura de Moreau.

Ejemplos de aplicaciones basadas en transformadas de distancia incluyen la morfometría de los nervios transversales, el registro de imágenes de resonancia magnética, cámara de ruta de planificación en endoscopia virtual y clasificación de tejidos en las imágenes de resonancia magnética.

La siguiente figura basada en el **algoritmo PE** muestra la transformada distancia de una imagen binario con la distancia cuadrada Euclıdea.

- 11(a) muestra la imagen de entrada.
- La transformación de la entrada a lo largo de cada columna se muestra en 11(b).
- La transformada distancia final, obtenida mediante la transformación del resultado intermedio a lo largo de cada fila, se muestra en 11(c).

Los píxeles oscuros corresponden a valores bajos mientras que los píxeles brillantes corresponden a valores altos.



(a) (b) (c)

Figure 11: Transformada distancia de una imagen binario con la distancia cuadrada Euclıdea

7 Metodología

A continuación se describen los aspectos importantes de la metodología de este trabajo.

1. **Tipo de investigación.** Este proyecto es de caracter bibliográfico ilustrativo.

   1.1 **Bibliográfico:**
   Se hará una recopilación de libros impresos, libros obtenidos en internet, tesis de grado y artículos en linea para contar con el material suficiente que cubra las necesidades del estudio y las que puedan surgir más adelante. El objetivo es compilar de forma coherente la información mas útil y destacada de los resultados que se aplicarán en el tema en cuestión.

   1.2 **Ilustrativa:**
   Ya que se pretende estudiar diversas aplicaciones con la la teoría que se presentará en este texto y en base a ello hacer uso de las herramientas que nos proporcione ésta para finalmente *hacer uso de sofware* (MATLAB Ó libre) y aplicar ésto a los problemas que se pretendan estudiar y de esta manera tener una mejor visión de lo importante que es la optimización convexa y los beneficios que nos genera.

2. **Forma de trabajo:**
   Se tendrán reuniones periódicas (una ó dos veces por semana) con el asesor para tratar los aspectos de la investigación tales como analizar y estudiar la teoría, tratar los diferentes aspectos del trabajo escrito y las presentaciones de avances.

3. **Exposiciones.**
   Se tendrán dos exposiciones:

   - **Primera exposición (pública):** Presentación del Perfil del Proyecto de Investigación.
   - **Segunda exposición (pública):** Presentación final del Trabajo de Investigación: resumen de resultados y aplicaciones.

8 Cronograma

Días	Enero				Febrero				Marzo				Abril				Mayo				Junio			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Revisión de trabajo escrito				x	x	x				x														
Estudio y discusión de teoria							x	x		x	x													
Defensa									x		x	xp												xp
Recopilación de material de estudio									x		x	xp												xp
Análisis de material de estudio									x		x	xp												xp
Aprobación de junta directiva									x		x	xp												xp
Entrega de documento final									x		x	xp												xp
Aprobación de tibunal calificador									x		x	xp												xp

xp = Defensa pública

9 Bibliografía

References

[1] R. Tyrrell Rockafellar. Convex Analysis. Princeton University Press, 1996.

[2] Bazaraa M., Sherali H., Shetty C.: Nonlinear programming. Theory and algorithms. Wyley-Interscience, 2006.

Fuentes de Internet

References

[1] Introduccion al Analisis Convexo. Pedro Gajardo. Centro de Modelamiento Matemático. Universidad de Chile
http://www.dim.uchile.cl/~pgajardo/documentos/A_Convexo_pgajardo.pdf

[2] OPTIMIZACION NO DIFERENCIABLE. Hector Manuel Mora Escobar. Dpto. Matemáaticas, U. Nacional, Bogotá
http://www.unicauca.edu.co/matematicas/eventos/optimizacion/Cursillos/Hector-Mora.pdf

[3] Funciones convexas no diferenciables. Tesis de Rafael Alejandro Nava Manzo. Universidad Autónoma Metropolitana. Unidad Iztapalapa
http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.12-O/Nava_Manzo_Rafael_Alejandro.pdf

[4] HERRAMIENTAS DE OPTIMIZACIÓN CONVEXA Y APLICACIONES EN TELECOMUNICACIONES. Emilio Mejía Fernpandez de Velasco. UNIVERSIDAD CARLOS III DE MADRID. ESCUELA POLITÉCNICA SUPERIOR
http://e-archivo.uc3m.es/bitstream/handle/10016/12990/PFC_Emilio_Mejia_Fernandez_Velasco.pdf?sequence=3

[5] Desarrollo del concepto de función convexa. Nelson J. Merentes D. Universidad Central de Venezuela
http://evm.ivic.gob.ve/LibroMerentesRivas.pdf

[6] Análisis convexo y dualidad. Felipe Álvarez con la colaboración de Juan Escobar y Juan Peypouquet. Universidad de Chile
opteq.dim.uchile.cl/public/?/publicaciones/getUsersDocs/pdf/11f75d0f22278872192b00ab3c5ab428/1

[7] Optimizacion. Jorge Amaya A. Departamento de Ingenieria Matematica y Centro de Modelamiento Matematico. Universidad de Chile.
http://escuela.mate304.org/NotasOptimizaci%C3%B3n_Convexa.pdf

[8] Núcleo Promedio de Funciones convexas y Algoritmos en Computación Convexa. César Jesús Lara Ávila. Universidad Nacional de Ingeniería. Lima-Perú. http://cybertesis.uni.edu.pe/bitstream/uni/1165/1/lara_ac.pdf

[9] OPTIMIZACIÓN C ONVEXA NO DIFERENCIABLE. Dr. Pablo A. Lotito. www.google.com

[10] Algunas aplicaciones y extensión del método del subgradiente. Frank Navarro Rojas. UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS. Lima – Perú. http://cybertesis.unmsm.edu.pe/bitstream/cybertesis/4580/1/Navarro_rf.pdf