UNIVERSIDAD NACIONAL
"JOSÉ FAUSTINO SÁNCHEZ CARRIÓN".

FACULTAD DE CIENCIAS

TESIS
"OPTIMIZACIÓN CONVEXA"

PARA OPTAR EL TITULO PROFESIONAL DE LICENCIADO EN MATEMÁTICA APLICADA

AUTOR:

Bach. CAPUÑAY VILCHEZ, EDGAR JOSÉ

DOCENTE ASESOR: Mo. ISIDRO JAVIER RÍOS PÉREZ

HUACHO – PERÚ
2020

“OPTIMIZACIÓN CONVEXA”

MIEMBROS DEL JURADO

Mg. CARLOS ROBERTO PESANTES ROJAS
PRESIDENTE

Mg. JORGE LUIS ROJAS PAZ
SECRETARIO

Dr. CRISTIAN IVÁN ESCURRA ESTRADA
VOCAL

Mo. ISISDRO JAVIER RÍOS PÉREZ
ASESOR

DEDICATORIA

PARA MIS PADRES:

La presente tesis está dedica a mi familia, principalmente a mi padre Capuñay Agapito que ha sido un pilar fundamental como padre, por brindarme la confianza, consejos, oportunidades y recursos para lograr ser profesional. Con todo cariño, aprecio y mucho respeto a mi madre Raquel Vílchez Zambrano.

Agradezco profundamente a mis padres:

JOSE MERCEDES CAPUÑAY AGAPITO

Y

ANGELITA RAQUEL VILCHEZ ZAMBRANO

Edgar José

AGRADECIMIENTO

La tesis: **OPTIMIZACIÓN CONVEXA,** se ha hecho posible gracias al apoyo de mis padres, mi esposa y demás familiares por la ayuda económica y emocional para el desarrollo de la tesis Agradezco también a todos los Docentes de la Facultad de Ciencias y de otras Facultades por los conocimientos que me impartieron durante 5 años de formación en lo académico.

INDICE

CAPITULO III

RESUMEN

. El objetivo del proyecto de tesis es resolver un programa de optimización convexa.

Materiales y métodos. El diseño metodológico utilizado es del tipo teórico documental. En la investigación se utilizó el diseño Descriptivo y analítico cualitativo, no se usa estadística.

Para la contratación de las Hipótesis se utilizó el método analítico

Conclusiones. Desde el punto de vista educativo se concluye que para la solución de un programa convexo es fundamental el estudio de los conjuntos convexos y las funciones convexas tanto para programas convexos libres y restringidos con igualdad y con desigualdad

PALABRAS CLAVES: Conjunto Convexo, función convexa, optimización, Extremos, programa convexo, punto crítico, solución óptima.

ABSTRACT.

. The objective of the thesis project is to solve a convex optimization program

Materials and methods. The methodological design used is of the theoretical documentary type. In the research the descriptive and qualitative analytical design was used, no statistics are used.

For the contracting of the hypotheses the analytical method was used

Conclusions From the educational point of view it is concluded that for the solution of a convex program it is essential to study convex sets and convex functions for both free and restricted convex programs with equality and inequality

KEY WORDS: Convex set, convex function, optimization, Extremes, convex program, critical point, optimal solution.

INTRODUCCIÓN

Durante el estudio del Cálculo diferencial se estudia optimización usando los criterios de la primera derivada y de la segunda derivada; así como las funciones cóncavas hacia arriba y hacia abajo., sin considerar las formas cuadráticas que son de vital importancia en el estudio de optimización convexa y de optimización convexa en cálculo de variaciones.

La optimización convexa estudia el problema general de maximizar una función convexa sobre un conjunto factible también convexo. Para este estudio se utiliza las formas cuadráticas

En el estudio de optimización convexa se establece resultados tales como:

Condiciones para óptimos locales

1. **Condición necesaria de primer orden**:

   Sea f diferenciable en $\overline{x}$. Si $\overline{x}$ es un mínimo o máximo local de f entonces

   $$\nabla f(\overline{x}) = 0$$

2. **Condiciones necesarias de segundo orden**:

   Sea f dos veces diferenciable en $\overline{x}$. Si $\overline{x}$ es un mínimo local de f entonces

   2.1. $\nabla f(\overline{x}) = 0$

   2.2. $\mathrm{H}f(\overline{x})$ es semi definida positiva

3. **Condiciones suficientes de segundo orden**:

   Sea f dos veces diferenciable en $\overline{x}$. Si

   3.1. $\nabla f(\overline{x}) = 0$

   3.2. $Hf(\overline{x})$ es definida positiva entonces $\overline{x}$ es un mínimo local (estricto) de f

Los resultados del estudio e optimización convexa son de gran beneficio para los estudiantes de Ciencias e Ingeniería y a profesionales de matemáticas e Ingenieros, dado que les permite encontrar aplicaciones prácticas y manejar eficientemente los algoritmos de optimización.

El desarrollo de la tesis consiste preliminarmente en establecer los conocimientos básicos para luego estudiar la optimización de los programas convexos sujetos a restricciones de igualdad y de desigualdad.

CAPÍTULO I

1. PLANTEAMIENTO DEL PROBLEMA

1.1. DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA

La matemática es una herramienta fundamental en el desarrollo de otras ciencias. Específicamente el creciente avance Y técnicas de optimización en ciencias tales como economía, Ingeniería, Administración de negocios, etc. nos motiva desarrollar esta tesis. Nos preguntamos si las formas cuadráticas y las funciones convexas tienen importancia en la optimización. La respuesta a esta pregunta se explicará en el desarrollo de la presente tesis

1.2. FORMULACIÓN DEL PROBLEMA

1.2.1. PROBLEMA GENERAL

¿Es importante la convexidad en la optimización matemática?

1.2.2. PROBLEMAS ESPECÍFICOS

- ¿Es fundamental el uso de funciones convexas en la optimización?
- ¿Cómo estudiar optimización usando concavidad?
- ¿Es posible establecer criterios de optimización convexa?

1.3. OBJETIVOS DE LA INVESTIGACIÓN

1.3.1. OBJETIVO GENERAL.

Resolver un programa de optimización convexa.

1.3.2. OBJETIVOS ESPECIFICOS.

OE1. Resolver un programa de optimización libre.
OE2. Resolver un programa de optimización con restricciones de igualdad.
O.E3. Resolver un programa de optimización con restricciones de desigualdad

CAPITULO II

2. MARCO TEÓRICO

2.1. ANTECEDENTES DE LA INVESTIGACIÓN.

- ✓ Olvi L Mangasarian (1069) en su obra Non Linear Programming, presenta un estudio sobre funciones cóncavas y convexas.
- ✓ En 1972 ante la Operations Research of American, el Institute of Management Sciences y el American Institute of Electrical Enginners, R.I Ackoff afirmó: Creo estamos terminando una era cultural y tecnológica y empezando otra, creo que estamos en las primeras etapas de un profundo cambio en la concepción del mundo en la tecnología que desarrollamos con el objeto de satisfacer nuestros propósitos
- ✓ Bruno Philippi (1982) en su obra Optimización de sistemas escribe algunas nociones básicas de convexidad hablando de conjuntos conexos y funciones cóncavas y convexas.
- ✓ Saúl I Gass (1064) en la parte 4 de la obra Programación no lineal describe algunos resultados sobre convexidad optimización.
- ✓ Leonelli. Héctor y Rumbros. Beatríz (2001) en su obra Métodos dinámicos en Economía. (Otra búsqueda del tiempo pérfido) en el capítulo IV llamado optimización estática presentan un estudio de las funciones convexas,
- ✓ Bonifaz. José L. y Lama. Rudy (1999) en su obra Optimización Dinámica y teoría económica en el capítulo III, llamado Control Óptimo presentan un estudio de funciones cóncavas y convexas

2.2. BASES TEORICAS.

Introducción. En esta sección resumiremos definiciones y propiedades topológicas de conjuntos en R^n, vectores y matrices con algunas propiedades de espacios vectoriales, así como funciones diferenciables. Presentaremos el teorema de la función implícita y formas cuadráticas.

ALGUNOS CONOCIMIENTOS BÁSICOS EN $\boldsymbol{R^n}$.

1. CONJUNTOS ABIERTOS Y CERRADOS EN $\boldsymbol{R^n}$

1.1. BOLA ABIERTA.

Dado un punto x_0 en R^n y un número real $\rho > 0$, el conjunto

$$B(\boldsymbol{x_0}, \rho) = \{ \boldsymbol{x} \in R^n \ / \ \|\boldsymbol{x} - \boldsymbol{x_0}\| < \rho\}$$

se llama bola abierta de centro $\boldsymbol{x_0}$ y radio ρ

1.2. BOLA CERRADA.

Dado un punto $\boldsymbol{x_0}$ en R^n y un número real $\rho > 0$, el conjunto

$$\bar{B}(x_0 \, , \rho) = \left\{ \begin{matrix} x \\ x \in R^n \ / \ \|x - x_0\| \leq \rho \end{matrix} \right\}$$

se llama bola cerrada de centro x_0 y radio ρ

1.2.1. PUNTO INTERIOR.

Un punto $\boldsymbol{x}$ en R^n es un punto interior del conjunto $S \subset R^n$ si existe un número real $\varepsilon > 0$ tal que $B(\boldsymbol{x}, \varepsilon) \subset S$.

1.2.2. PUNTO DE CLAUSURA.

Un punto $\boldsymbol{x}$ en R^n es de clausura del conjunto $S \subset R^n$ si existe un número positivo $\varepsilon > 0$ tal que $B(\boldsymbol{x}, \varepsilon) \cap S \neq \emptyset$.

Es necesario hacer notar que un punto de clausura no necesita pertenecer al conjunto S. Por ejemplo $0 \in R$ es un punto de clausura del conjunto infinito $S = \{1, 1/2, 1/4, \cdots\}$, pero $0 \notin S$. Por otro lado, todo punto del conjunto S es también un punto de clausura de S. En otras palabras, un punto de clausura $\boldsymbol{x}$ de un conjunto S es un punto tal que existen puntos en S que están arbitrariamente cercanos a él.

1.2.3. CONJUNTO ABIERTO.

Un conjunto $S \subset R^n$ tal que todo punto de S es un punto interior se dice que es un conjunto abierto.

1.2.4. CONJUNTO CERRADO.

Un subconjunto S en R^n tal que todo punto de clausura de S está en S se dice que es un conjunto cerrado.

1.2.5. CLAUSURA DE UN CONJUNTO.

La clausura $\bar{S}$ de un subconjunto $S \subset R^n$ es el conjunto de todos los puntos de clausura de S. Obviamente para un conjunto abierto $S \subset \bar{S}$ y para cada conjunto cerrado $S = \bar{S}$

1.2.5.1. INTERIOR DE UN CONJUNTO.

El interior $Int(S)$ de un subconjunto $S \subset R^n$ es el conjunto de puntos interiores de S. Obviamente para un conjunto cerrado $Int(S) \subset S$ y para un conjunto abierto $Int(S) = S$

1.3. TEOREMAS RELATIVOS A CONJUNTOS ABIERTOS Y CERRADOS.

1.3.1. TEOREMA.

La familia de conjuntos abiertos en R^n tiene las siguientes propiedades.

(I) Toda unión arbitraria de conjuntos abiertos es abierto

(II) Toda intersección finita de conjuntos abiertos es abierto

(III) Los conjuntos vacío $\emptyset$ y R^n son abiertos.

DEMOSTRACIÓN.

(I) Sea $(S_i)_{i \in I}$ una familia, finita o infinita, de conjuntos abiertos en R^n.

Si $x \in S = \bigcup_{i \in I} S_i$ entonces $x \in S_i$ para algún $i \in I$, y existe un $\varepsilon > 0$ tal que:

$$B(x, \varepsilon) \subset S_i \subset S$$

Por consiguiente x es un punto interior de S, y por definición S es abierto.

(II) Sea $S_{i \in I}$ una familia finita de conjuntos abiertos en R^n.

Si $x \in S = \bigcap_{i \in I} S_i$ entonces $x \in S_i$ para cada $i \in I$. Puesto que cada S_i es abierto, entonces existe $\varepsilon_i > 0$ tal que $B(x, \varepsilon_i) \subset S_i$ para $\forall i \in I$.

Tomar $\varepsilon = \min_{i \in I} \varepsilon_i > 0$, entonces $B(x, \varepsilon) \subset S$ y por definición de conjunto abierto, seguimos que S es abierto.

(III) Puesto que el conjunto vacío ∅ no contiene puntos, no es necesario hallar una bola abierta alrededor de cualquier punto, y por consiguiente ∅ es abierto.

1.3.2. TEOREMA.

Sea $S \subset R^n$, entonces $R^n \sim \overline{S} = Int(R^n \sim S)$

DEMOSTRACIÓN.

Sea $\boldsymbol{x} \in R^n$. Entonces $\boldsymbol{x} \in \overline{S} \Leftrightarrow \boldsymbol{x} \notin Int(R^n \sim S)$

Por lo tanto $\boldsymbol{x} \notin \overline{S} \Leftrightarrow \boldsymbol{x} \in Int(R^n \sim S)$

1.3.3. COROLARIO.

El complemento (relativo a R^n) de un conjunto abierto en R^n es cerrado, y viceversa.

DEMOSTRACIÓN.

S es cerrado $\Leftrightarrow S = \overline{S} \Leftrightarrow R^n \sim S = R^n \sim \overline{S} = int(R^n \sim S) \Leftrightarrow R^n \sim S$ es abierto.

Usando el corolario 1.2.3, el siguiente teorema es una consecuencia directa del teorema 1.2.1.

1.3.4. TEOREMA.

La familia de conjuntos cerrados en R^n tiene las propiedades siguientes.

(I) Toda intersección de conjuntos cerrados es cerrado

(II) Toda unión finita de conjuntos cerrados es cerrado.

(III) Los conjuntos vacíos ∅ y R^n son cerrados.

DEMOSTRACIÓN.

La demostración de este teorema se deduce directamente del teorema 1.2.1 y usando las leyes de De Morgan siguientes:

$$S \sim \bigcap_{i=1}^{n} A_i = \bigcup_{i=1}^{n} (S \sim A_i)\,; \qquad S \sim \bigcup_{i=1}^{n} A_i = \bigcap_{i=1}^{n} (S \sim A_i)$$

1.4. SUCESIONES Y ACOTAMIENTO.

1.4.1. SUCESIÓN.

Una sucesión en un conjunto X es una función f del conjunto I de todos los enteros positivos en el conjunto X. Si $f(n) = x_n \in X$ para $n \in I$, es costumbre denotar la sucesión f por el símbolo $\{x_n\}$ o por $x_1, x_2, \cdots$ Para cualquier sucesión de enteros positivos $n_1, n_2, \cdots$, tal que $n_1 < n_2 < \cdots$, la sucesión $x_{n_1}, x_{n_2}, \cdots$ se llama una subsucesión de $x_1, x_2, \cdots$ Si $X = R^n$, entonces la sucesión $\{x_n\}$ ó $x_1, x_2, \cdots$ se llama sucesión de puntos en R^n, y si $X = R$, entonces $x_1, x_2, \cdots$ es la conocida sucesión de números reales.

1.4.2. PUNTO LÍMITE.

Sea $\{x_n\}$ ó $x_1, x_2, \cdots$ una sucesión de puntos en R^n. Un punto $\boldsymbol{x} \in R^n$ se dice que es un punto límite de la sucesión si

$\varepsilon > 0, N \in \mathbb{Z}^+ \Longrightarrow \|x_n - \overline{x}\| <\in$ Para algún $n \geq N$

Nota. Frecuentemente, a un punto límite también se llama punto de acumulación o punto clausura.

1.4.3. LÍMITE DE UNA SUCESIÓN.

Sea $\{x_n\}$ ó $x_1, x_2, \cdots$ una sucesión de puntos en R^n. A un punto $\boldsymbol{x_0} \in R^n$ se llama punto límite de la sucesión, o decimos que la sucesión converge a $\boldsymbol{x_0}$ si:

$$\forall\ \varepsilon > 0, \exists\ n_0 \in \mathbb{N}\ /\ \|x_n - x_0\| < \varepsilon, \forall\ n \geq n_0.$$

y escribimos $\lim\limits_{x \to \infty} x_n = \boldsymbol{x_0}$

OBSERVACIÓN. Obviamente el límite de una sucesión es también un punto límite de la sucesión, pero el reciproco no es necesariamente válido. Por ejemplo: la sucesión $1, -1, 1, -1, \cdots$ tiene los puntos límites $1\ y -1$, pero no tiene límite. La sucesión $1, 1/2,\ 1/4, \cdots$ tiene un límite, y por consiguiente un punto límite cero. También note que si $\boldsymbol{x_0}$ es un límite de $x_1, x_2, \cdots$, entonces él también es un límite de cada subsucesión de $x_1, x_2, \cdots$

1.4.4. TEOREMA.

(I) Si $\boldsymbol{\overline{x}}$ es un punto límite de la sucesión $x_1, x_2, \cdots$, entonces existe una subsucesión $x_{n_1}, x_{n_2}, \cdots$ la cual tiene a $\boldsymbol{\overline{x}}$ como su límite, y recíprocamente

(II) Si una sucesión $x_1, x_2, \cdots$, tiende a un límite $\boldsymbol{x_0}$, entonces este no puede ser punto límite (y por consiguiente no es límite) más que $\boldsymbol{x_0}$.

DEMOSTRACIÓN.

(I) Sea $\boldsymbol{\overline{x}}$ un punto límite de $x_1, x_2, \cdots$, Entonces

$$\forall\, \varepsilon > 0, \exists\, n_1 : \left\|x_{n_1} - x_0\right\| < \varepsilon$$

$$\Longrightarrow\ \exists\, n_2 \geq n_1 + 1 : \left\|x_{n_1} - x_0\right\| < \varepsilon/2$$

$$\cdots$$

La sub sucesión $x_{n_1}, x_{n_2}, \cdots$ converge a $\boldsymbol{\overline{x}}$

Recíprocamente, si la sub sucesión $x_{n_1}, x_{n_2}, \cdots$ converge a $\overline{x}$, entonces para $\varepsilon > 0$ y $\overline{n}$ dados, existe un $n_i \geq \overline{n}$ (en realidad para todo $n_i \geq \overline{n}$) tal que:

$$\left\|x_{n_i} - \overline{x}\right\| < \varepsilon,$$

y por consiguiente $\boldsymbol{\overline{x}}$ es un punto límite de $x_1, x_2, \cdots$

(II). Sea la sucesión $x_1, x_2, \cdots$, que tiende al punto límite $\boldsymbol{x_0}$ y $\boldsymbol{\overline{x}} \neq \boldsymbol{x_0}$. Entonces para $0 < \varepsilon < \|\overline{x} - x_0\|$, entonces existe un entero n_0 tal que $n \geq n_0 \ \Longrightarrow \|x_n - x_0\| < \varepsilon$

$\Longrightarrow \varepsilon + \|x_n - \overline{x}\| > \|x_0 - \overline{x}\|$ (Por desigualdad triangular)

$\Longrightarrow \|x_n - \overline{x}\| > \|x_0 - \overline{x}\| - \varepsilon > 0$

y por consiguiente $\boldsymbol{\overline{x}}$ no es punto límite de la sucesión.

OBSERVACIÓN. Nótese que el punto límite (y en consecuencia un límite) de una sucesión $x_1, x_2, \cdots$ en R^n es un punto de clausura de cualquier conjunto S en R^n que contiene $\{x_1, x_2, \cdots\}$. Recíprocamente, si $\boldsymbol{\overline{x}}$ es un punto de clausura del conjunto $S \subset R^n$, entonces existe una sucesión $x_1, x_2, \cdots \subset S$ (y en consecuencia también una sub sucesión $x_{n_1}, x_{n_2}, \cdots$) tal que $\boldsymbol{\overline{x}}$ es un punto límite de $x_1, x_2, \cdots$ (y en consecuencia un límite de la $x_{n_1}, x_{n_2}, \cdots$)

1.4.5. COTAS SUPERIOR E INFERIOR.

Sea S un conjunto no vacío de números reales. Entonces:

a) α es una cota inferior de $S \Leftrightarrow (x \in S \Rightarrow x \geq \alpha)$

b) β es una cota superior de $S \Leftrightarrow (x \in S \Rightarrow x \leq \beta)$

1.4.6. MÁXIMA COTA INFERIOR Y MÍNIMA COTA SUPERIOR.

Sea S un conjunto no vacío de números reales.

a) Una cota inferior $\overline{\alpha}$ es una máxima cota inferior de S (ínfimo de S, o inf S) si ningún número mayor es una cota inferior de S.

Equivalentemente:

$$\overline{\alpha} = \inf S \Leftrightarrow \begin{cases} x \in S \Rightarrow x \geq \overline{\alpha} \\ \forall\, \varepsilon > 0 \Rightarrow \exists x \in S: \overline{\alpha} + \varepsilon > x \end{cases}$$

b) Una cota superior $\overline{\beta}$ es una mínima cota superior de S (supremo de S, o sup S) si ningún número menor es una cota superior de S.

Equivalentemente:

$$\overline{\beta} = \sup S \Leftrightarrow \begin{cases} x \in S \Rightarrow x \leq \overline{\beta} \\ \forall\, \varepsilon > 0 \Rightarrow \exists x \in S: \overline{\beta} - \varepsilon < x \end{cases}$$

Tomamos el siguiente axioma como uno de los axiomas del sistema de los números reales.

1.4.7. AXIOMA.

Cualquier conjunto no vacío S de números reales que tiene una cota inferior (superior) tiene una máxima (mínima) cota inferior (superior)

Si un conjunto no tiene ínfimo (o equivalentemente por el axioma anterior si él no tiene cota inferior), decimos que S es no acotado por abajo y escribimos $ínf\ S = -\infty$. Similarmente, si S no tiene supremo (o equivalentemente por el axioma anterior si él no tiene cota superior), decimos que S es no acotado por arriba y escribimos $sup\ S = +\infty$ Por consiguiente aumentando a la línea Euclidiana R los dos símbolos $+\infty$ y $-\infty$, cualquier conjunto no vacío S tiene un ínfimo que puede ser $-\infty$, y un supremo, que puede ser $+\infty$. Convencionalmente escribiremos $ínf\ S = +\infty$ y $\sup S = -\infty$

Obsérvese que ni $ínf\, S$ ni $\sup S$ necesitan estar en S. Por ejemplo:

$ínf\{1, 1/2, 1/3, \cdots\}$ es 0, pero 0 no está en el conjunto $\{1, 1/2, 1/3, \cdots\}$

1.4.8. TEOREMA.

a) Toda sucesión acotada no decreciente de números reales tiene un límite.
b) Toda sucesión acotada no creciente de números reales tiene un límite

DEMOSTRACIÓN.

a) Sea $\{x_n\}$ ó $x_1, x_2, x_3, \cdots$ una sucesión acotada no decreciente de números reales. Por el axioma 1.3.7, la sucesión tiene una mínima cota superior $\overline{\beta}$.
Por consiguiente por 1.3.6 tenemos

$$\forall\, \varepsilon > 0, \exists\, \overline{n}: \overline{\beta} - x_n < \varepsilon$$

$\Longrightarrow\ \overline{\beta} - x_n < \varepsilon, \forall\, n \geq \overline{n}$ (Puesto que la sucesión es no decreciente

$\Longrightarrow -\varepsilon < 0 \leq \overline{\beta} - x_n < \varepsilon,\ \forall\, n \geq \overline{n}$ (Puesto que $\overline{\beta} \geq x_n\ \forall\, n$)

$\Longrightarrow\ \overline{\beta}$ es un límite de la sucesión $x_1, x_2, x_3, \cdots$ ■

b) La prueba para una sucesión acotada no creciente es similar)

1.4.9. SUCESIÓN DE CAUCHY Y CRITERIO DE CONVERGENCIA DE CAUCHY.

a) SUCESIÓN DE CAUCHY.

Se dice que una sucesión $\{x_n\}$ ó $x_1, x_2, x_3, \cdots$ en R^n es una sucesión de cauchy si para cualquier $\varepsilon > 0$ existe (∃) un número natural n^* tal que para todos los números naturales $n, m \geq n^*$ se tiene $\|x_n - x_m\| < \varepsilon$

b) CRITERIO DE CONVERGENCIA DE CAUCHY.

Una sucesión $\{x_n\}$ ó $x_1, x_2, x_3, \cdots$ en R^n converge a un límite $\boldsymbol{x_0}$ si y sólo si es una sucesión de Cauchy, esto es:

$$\forall\, \varepsilon > 0, \exists\, n^*: \|x_n - x_m\| < \varepsilon, \forall\, n, m \geq n^*.$$

1.5. CONJUNTOS ACOTADOS Y COMPACTOS EN $\boldsymbol{R^n}$.

1.5.1. CONJUNTO ACOTADO.

Un subconjunto S en R^n está acotado si existe un número real M tal que para todo $\boldsymbol{x}$ de S, $\|\boldsymbol{x}\| \leq M$.

1.5.2. CONJUNTOS COMPACTOS.

Un subconjunto S en R^n se dice que es compacto si satisface cualquiera de las siguientes propiedades:

(i) S es cerrado y acotado.

(ii) (BOLZANO-WEIERSTRASS) Toda sucesión de puntos en S tiene un punto límite en S.

(iii) (PROPIEDAD DE INTERSECCIÓN FINITA) Para cualquier familia $(S_i)_{i\in I}$ de conjunto cerrados relativos a S, seguimos que

$$\bigcap_{i\in I} S_i = \emptyset \implies \begin{cases} S_{i_1} \cap S_{I_2} \cap \cdots \cap S_{i_m} = \emptyset \\ para\ alg\acute{u}n\ i_1, i_2, \cdots, i_m \in I \end{cases}$$

ó equivalentemente

$$\bigcap_{i\in I} S_i \neq \emptyset \Longleftarrow \begin{cases} S_{i_1} \cap S_{I_2} \cap \cdots \cap S_{i_m} \neq \emptyset \\ para\ todo\ i_1, i_2, \cdots, i_m \in I \end{cases}$$

(iv) (HEINE- BOREL) De toda familia $(S_i)_{i\in I}$ de conjuntos abiertos cuya unión $\cup S_i$ contiene a S, podemos extraer una subfamilia $i_1, i_2, \cdots, i_m$ cuya unión $i_1 \cup i_2 \cup \cdots \cup i_m$ contiene a S (o equivalentemente: Todo cubrimiento abierto de S tiene un subcubrimiento finito)

1.5.3. COROLARIO.

Sean S y A dos conjuntos en R^n los cuales son compactos y cerrados, respectivamente. Entonces la suma

$$T = S + A = \{\boldsymbol{x} + \boldsymbol{y} \,/\, \boldsymbol{x} \in S, \boldsymbol{y} \in A\}$$

es cerrada.

1.6. VECTORES.

1.6.1. Definición (Vector). Un vector $n - dimensional\, \boldsymbol{x}$, para cualquier entero positivo n es una $n - tupla$ $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ de números reales. Al número real x_i se llama la i –ésima componente del vector $\boldsymbol{x}$

El espacio euclidiano R^n es el conjunto de todos los vectores n-dimensional

$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n),$$

Siendo cada vector $\boldsymbol{x}$ un punto de R^n. Entonces los vectores son puntos de R^n

1.6.2. NOTACIÓN.

En general las letras latinas minúsculas en negritas tales como $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ representan vectores y las letras en negritas tales como $\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}$ representan funciones vectoriales. Por otro lado las letras i, j, k, m, n y algunas otras denotan números enteros.

Las letras griegas minúsculas tales como α, β, μ, ν representan números reales (puntos en R) y las letras $\theta, \emptyset, \varphi$ representan funciones numéricas (reales).

Los números reales x_i representan componentes de un vector y las letras griegas minúsculas con un subíndice tales como $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_n$ indican números reales.

1.7. ADICIÓN DE VECTORES Y MULTIPLICACIÓN POR UN NÚMERO REAL.

Sean x e y dos vectores en R^n y sea $\alpha \in R$.

La suma, denotada por $\boldsymbol{x} + \boldsymbol{y}$, se define como:

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \cdots, x_n + y_n)$$

La sustracción o diferencia, denotada por $\boldsymbol{x} - \boldsymbol{y}$, se define como:

$$\boldsymbol{x} - \boldsymbol{y} = (x_1 - y_1, x_2 - y_2, x_3 - y_3, \cdots, x_n - y_n)$$

La multiplicación por un número real α denotada por $\alpha\, \boldsymbol{x}$, se define como:

$$\alpha\boldsymbol{x} = (\alpha x_1, \alpha x_2, \cdots, \alpha x_n).$$

1.7.1. COMBINACIÓN LINEAL

Un vector $\boldsymbol{x} \in R^n$ es una combinación lineal de los vectores $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_m}$ en R^n si $\boldsymbol{x} = \lambda_1\boldsymbol{x_1} + \lambda_2\boldsymbol{x_2} + \cdots + \lambda_m\boldsymbol{x_m}$ para algunos $\lambda_1, \lambda_2, \cdots, \lambda_m \in R$, y es una combinación lineal no negativa de $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_m} \in R^n$ si además, de la igualdad anterior, se cumple $\lambda_1, \lambda_2, \cdots, \lambda_m \geq 0$.

Los números $\lambda_1, \lambda_2, \cdots, \lambda_m$ se llaman pesos.

1.7.2. INDEPENDENCIA Y DEPENDENCIA LINEAL.

Los vectores $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_m} \in R^n$ son linealmente independientes (L.I), si de la relación de dependencia lineal

$$\lambda_1\boldsymbol{x_1} + \lambda_2\boldsymbol{x_2} + \cdots + \lambda_m\boldsymbol{x_m} = \boldsymbol{0};\ \lambda_1, \lambda_2, \cdots, \lambda_m \in R$$

se sigue que

$$\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0.$$

Los vectores $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_m} \in R^n$ son linealmente dependientes (L.D) si existen escalares $\lambda_1, \lambda_2, \cdots, \lambda_m$ no todos ceros tales que

$$\lambda_1\boldsymbol{x_1} + \lambda_2\boldsymbol{x_2} + \cdots + \lambda_m\boldsymbol{x_m} = 0$$

1.7.3. PRODUCTO ESCALAR

El producto escalar de dos vectores x e y de R^n denotado por $\boldsymbol{x} \cdot \boldsymbol{y}$ se define como

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1y_1, +x_2y_2 + x_3y_3 + \cdots + x_ny_n$$

1.7.4. NORMA DE UN VECTOR.

La norma de un vector $\boldsymbol{x} \in R^n$ denotada por $\|\boldsymbol{x}\|$ se define por

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{(x_1)^2 + (x_2)^2 + \cdots + (x_n)^2}$$

1.7.5. DESIGUIALDAD DE CAUCHY – SCHWARZ.

Sean x e y dos vectores en R^n. Entonces

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

Donde $|\boldsymbol{x} \cdot \boldsymbol{y}|$ es el valor absoluto del número real $\boldsymbol{x} \cdot \boldsymbol{y}$

1.7.6. DISTANCIA ENTRE DOS PUNTOS.

Sean $\boldsymbol{x}, \boldsymbol{y} \in R^n$. La distancia entre los puntos (vectores) $\boldsymbol{x}, \boldsymbol{y} \in R^n$ es el número no negativo

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{(\boldsymbol{x} - \boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})}$$

Obsérvese que

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{(\boldsymbol{x} - \boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})}$$

$$= \sqrt{(x_1 - y_1, x_2 - y_2, \cdots, x_n - y_n) \cdot (x_1 - y_1, x_2 - y_2, \cdots, x_n - y_n}$$

$= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$ $\therefore\ d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$

1.7.7. TEOREMA FUNDAMENTAL DE ESPACIOS VECTORIALES.

Si cada uno de los vectores $\boldsymbol{y_0}, \boldsymbol{y_1}, \cdots, \boldsymbol{y_m} \in R^n$ es una combinación lineal de $\boldsymbol{x_1},\ \boldsymbol{x_2}, \cdots, \boldsymbol{x_m} \in R^n$; entonces $\boldsymbol{y_0}, \boldsymbol{y_1}, \cdots, \boldsymbol{y_m}$ son linealmente dependientes.

DEMOSTRACIÓN.

La prueba se hace por inducción sobre m. Si $m = 1$, entonces $\boldsymbol{y_0} = p_0\boldsymbol{x_1},\ \boldsymbol{y_1} = p_1\boldsymbol{x_1}$. Si $p_0 = p_1 = 0$ entonces $\boldsymbol{y_0} = \boldsymbol{y_1} = \boldsymbol{0}$ y así $\boldsymbol{y_0}$ y $\boldsymbol{y_1}$ son linealmente dependientes. Si no, entonces digamos por ejemplo que $p_0 \neq 0$. Entonces $p_1\boldsymbol{y_0} - p_0\boldsymbol{y_1} = p_1p_0\boldsymbol{x_1} - p_0p_1x_1 = 0$ y así $\boldsymbol{y_0}$ y $\boldsymbol{y_1}$ son linealmente dependientes puesto que $p_0 \neq 0$

Ahora asumamos que el teorema es válido para $m = k - 1$ y mostraremos que también es válido para $m = k$. Por hipótesis tenemos que

$$\boldsymbol{y_j} = \sum_{i=1}^{k} p_i^j \boldsymbol{x}_i\,, \qquad j = 0, 1, 2, \cdots, k$$

Si todo los p_i^j es cero entonces todos los $\boldsymbol{y_j}$ son ceros y por consiguiente linealmente dependientes.

Asuma ahora que al menos un p_i^j es diferente de cero, digamos por ejemplo p_1^0. Definamos

$$\boldsymbol{z_j} = \boldsymbol{y_j} - \frac{p_i^j}{p_0^1}\boldsymbol{y_0} = \sum_{i=2}^{k}\left(p_i^j - \frac{p_1^j}{p_1^0}p_i^0\right)\boldsymbol{x_i} \qquad j = 1, 2, 3, \cdots, k$$

Entonces cada uno de los vectores $\boldsymbol{z_j}$ es una combinación lineal de los $k - 1$ vectores $\boldsymbol{x_2}, \boldsymbol{x_3}, \cdots, \boldsymbol{x_k}$, y por consiguiente por la hipótesis inductiva, los $\boldsymbol{z_j}$ son linealmente dependientes, esto es, existen números $q_1, q_2, \cdots, q_k$ no todos ceros, tal que

$$0 = \sum_{j=1}^{k} q_j\boldsymbol{z_j} = \sum_{j=1}^{k} q_j\boldsymbol{y_j} - \frac{1}{p_1^0}\left(\sum_{j=1}^{k} q_j p_1^j\right)\boldsymbol{y_0}$$

lo cual muestra que los $\boldsymbol{y_j}$ son linealmente dependientes.

1.7.8. COROLARIO.

Cualquier conjunto $S = \{\boldsymbol{y_1}, \boldsymbol{y_2}, \cdots, \boldsymbol{y_n}, \boldsymbol{y_{n+1}}\}$ de $n+1$ vectores en R^n son linealmente dependientes.

DEMOSTRACIÓN.

Sea $\boldsymbol{e_i}$ un vector en R^n cuyos elementos son ceros a excepción del i-ésimo elemento, el cual es 1. Entonces cualquier vector $\boldsymbol{x} \in R^n$ es una combinación lineal del vector $\boldsymbol{e_i}$, esto es

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i e_i$$

Entonces por el teorema fundamental de espacios vectoriales seguimos que el conjunto S el linealmente dependiente.

1.7.9. COROLARIO.

Cualquier conjunto $S = \{\boldsymbol{y_1}, \boldsymbol{y_2}, \cdots, \boldsymbol{y_m}\}$ de m vectores en R^n son linealmente dependientes si $m > n$

DEMOSTRACIÓN.

Por corolario 1.6.8 cualquier conjunto S de $n+1$ vectores de los m vectores es linealmente dependiente, por consiguiente los m vectores son linealmente dependientes.

1.7.10. COROLARIO.

Cada sistema de n ecuaciones lineales homogéneas con m incógnitas, $m > n$ tiene una solución diferente de cero.

DEMOSTRACIÓN.

Considere la ecuación matricial $\boldsymbol{Ax} = \boldsymbol{0}$ donde A es cualquier matriz $n \times m$. Las m columnas $A_{.j}$ de A son vectores en R^n y por consiguiente por corolario 1.4.12 son linealmente dependientes. Por lo tanto existe un $x \neq 0$ tal que $\boldsymbol{Ax} = \boldsymbol{0}$.

1.7.11. BASES DE UN SUBESPACIO DE $\boldsymbol{R^n}$

Definición. Sea S un subespacio de R^n, y sea r el máximo número de vectores linealmente independientes que pueden ser seleccionados de S. Cualquier conjunto de r vectores linealmente independientes de S se llama una base para S.

1.7.12. TEOREMA DE LA BASE.

El conjunto $T = \{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \cdots, \boldsymbol{x_r}\}$ de vectores linealmente independientes es una base para un subespacio S de R^n si y sólo si cada vector $\boldsymbol{y}$ en S $(y \in S)$ es una combinación lineal de los $\boldsymbol{x_i}$, $i = 1, 2, \cdots, r$

DEMOSTRACIÓN.

Supongamos que cada vector $\boldsymbol{y}$ en S es una combinación lineal de los $\boldsymbol{x_i}$, es decir

$\boldsymbol{y} = \sum_{\boldsymbol{i=1}}^{r} p_i \boldsymbol{x_i}$.

Entonces el conjunto S ya no contiene vectores linealmente independientes, para cualquier conjunto de más que r vectores debe ser dependiente ya que ellos son combinación lineal de los $\boldsymbol{x_i}$. Por lo tanto r es el máximo número de vectores linealmente independientes que pueden ser seleccionados de S y por consiguientes el conjunto de los $\boldsymbol{x_i}$ son una base para S.

Recíprocamente. Supongamos que el conjunto T de los $\boldsymbol{x_i}$ es una base para S. Entonces por definición, r es el número de vectores en el conjunto más grande de vectores linealmente independientes que pueden ser hallados en S. Así si $\boldsymbol{y}$ está en S, entonces $T = \{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \cdots, \boldsymbol{x_r}\}$ es linealmente dependiente, esto es

$$\sum_{i=1}^{r} p_i x_i + p_0 y = 0,$$

y $p_0 \neq 0$, de otro modo los x_i serían linealmente dependientes. Por consiguiente

$$y = -\frac{1}{p_0} \sum_{i=1}^{r} p_i x_i$$

la cual es la combinación lineal deseada.

1.8. MATRICES.

1.8.1. Definición. Una matriz A de orden $m \times n$ es una configuración rectangular de elementos encerrados entre paréntesis o corchetes de la siguiente forma:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

El i-ésimo renglón de la matriz A se denotado por $\boldsymbol{A_i}$. será un n-vector y se escribirá como:

$$A_i. = (A_{i1}, A_{i2}, \cdots, A_{in}) \quad i = 1, 2, \cdots, m$$

La j-ésima columna de la matriz A denotada por $A._j$ será un m-vector y se escribirá como

$$A._j = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix} \quad j = 1, 2, \cdots, n$$

1.8.2. MATRIZ TRANSPUESTA.

La transpuesta de la matriz A de orden $m \times n$ se denota por A^T y se determina cambiando las filas por las columnas, entonces la transpuesta queda definida por:

$$A^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{m1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{mn} \end{bmatrix}$$

1.8.3. RANGO FILA Y COLUMNA DE UNA MATRIZ.

El máximo número de renglones o filas linealmente independientes de una matriz se llama rango fila de una matriz, y el máximo número de columnas linealmente independientes de una matriz se llama rango columna de una matriz. El rango de una matriz es el rango fila o columna de la matriz

1.8.4. TEOREMA DEL RANGO.

Para cualquier matriz A, el rango fila y rango columna son iguales.

DEMOSTRACIÓN.

Sea r el rango fila y s el rango columna de la matriz A tal que $r < s$. Seleccionemos un renglón básico para A, el cual podemos asumir que consiste de los renglones $A_{1.}, A_{2.}, \cdots, A_{r.}$ y una columna básica, la cual la asumiremos que consiste de las columnas $A_{\cdot 1}, A_{\cdot 2}, \cdots, A_{\cdot s}$

Sea $\overline{A_\iota} = (A_{i1}, A_{i2}, \cdots, A_{is})$ y note que las ecuaciones

$$\overline{A_\iota} y = 0 \qquad i = 1, 2, \cdots, r$$

son r ecuaciones en s $(s > r)$ desconocidas y por corolario 1.4.13 tiene una solución $\bar{y}$ diferente de cero. También puesto que $A_1, A_2, \cdots, A_r$ es un renglón base, seguimos del teorema de la base 1.5.2 que para todo k, $A_k = \sum_{i=1}^{r} p_i{}^k A_i$ para algunos números $p_i{}^k$.

Por consiguiente

$$\overline{A_k} = \sum_{i=1}^{r} p_i{}^k \, \overline{A_\iota}$$

y

$$\overline{A_k}\bar{y} = \sum_{i=1}^{r} p_i{}^k \, (\overline{A_\iota}\bar{y}) = 0, \forall k$$

las cuales son equivalentes a

$$\sum_{j=1}^{s} \overline{A_{.J}}\,\overline{y_J} = 0$$

Esto muestra que las columnas $A_{\cdot 1}, A_{\cdot 2}, \cdots, A_{\cdot s}$ de A son linealmente dependientes, lo cual contradice la suposición que ellas fueron una base. Esta contradicción muestra que $r \geq s$. El mismo argumento aplicado a la transpuesta de A muestra que $s \geq r$. Por consiguiente $s = r$

1.8.5. COROLARIO.

Sea A una matriz $r \times n$ con rango r $(r \leq n)$. Entonces para cualquier $\boldsymbol{B} \in R^r$, el sistema $\boldsymbol{A}x = \boldsymbol{B}$ tiene una solución $x \in R^{n.}$

DEMOSTRACIÓN.

Por teorema del rango (Teorema 1.6.4), el rango columna de A es r, y así si $A_{\cdot 1}, A_{\cdot 2}, \cdots, A_{\cdot r}$ decimos, es una base columna, tenemos r vectores linealmente independientes en R^r, y el vector $\boldsymbol{B} \in R^r$ se puede expresar como una combinación lineal de ellos, esto es,

$$B = \sum_{j=1}^{r} A_{.j}\, x_j$$

Por definición, $x_j = 0$ para $j = r + 1, \cdots, n$, tenemos que $\boldsymbol{B} = \boldsymbol{A}x$

1.8.6. MATRIZ NO SINGULAR.

Una matriz $n \times n$ se llama no singular si tiene rango n.

1.8.7. MATRIZ SEMIDEFINIDA.

Una matriz A $n \times n$ se llama semidefinida positiva si

$$\boldsymbol{x}A\boldsymbol{x} \geq 0, \qquad \forall \boldsymbol{x} \in R^n$$

y semi definida negativa si

$$\boldsymbol{x}A\boldsymbol{x} \leq 0, \qquad \forall \boldsymbol{x} \in R^n$$

1.8.8. MATRIZ DEFINIDA.

Una matriz A $n \times n$ se llama matriz definida positiva si $\boldsymbol{x} \neq 0 \Longrightarrow \boldsymbol{x}A\boldsymbol{x} > 0$

y definida negativa si

$$\boldsymbol{x} \neq 0 \Longrightarrow \boldsymbol{x}A\boldsymbol{x} < 0$$

Obviamente el negativo de una matriz (definida) semidefinida positiva es una matriz (definida) semidefinida negativa y recíprocamente. También cada matriz definida (negativa) positiva también es semidefinida positiva (negativa)

1.8.9. TEOREMA.

Cada matriz definida positiva o negativa es no singular.

DEMOSTRACIÓN.

Sea A una matriz $n \times n$ definida positiva o negativa. Si A es singular, entonces su rango es menor que n, y existe un $\boldsymbol{x} \in R^n$, $\boldsymbol{x} \neq 0$, tal que $A\boldsymbol{x} = 0$. Por consiguiente $\boldsymbol{x}A\boldsymbol{x} = 0$ para algún $\boldsymbol{x} \neq 0$, lo cual contradice la suposición que A es definida positiva o negativa. Por consiguiente A es no singular.

1.9. FUNCIONES CONTINUA Y SEMICONTINUAS, MAXIMO Y MÍNIMO.

En esta sección se establece algunas definiciones básicas y propiedades de funciones continuas y semicontinuas definidas en un subconjunto de R^n. También establecemos algunos hechos acerca de máximo y mínimo de estas funciones.

1.9.1. FUNCIONES CONTINUAS Y SEMICONTINUAS.

1.9.1.1. FUNCIÓN NUMÉRICA O REAL.

Una función numérica o real θ es una función del conjunto $S \subset R^n$ en R. En otras palabras, una función numérica es una correspondencia que asocia un número real con cada $\boldsymbol{x}$ de S. Se denota por

$\theta: S \subset \mathbb{R}^n \to \mathbb{R}$, tal que $\boldsymbol{y} = \theta(\boldsymbol{x})$ con $\boldsymbol{x} \in \mathbb{R}^n$; $\boldsymbol{y} \in \mathbb{R}$

$\boldsymbol{y} = \theta(\boldsymbol{x})$ se llama regla de correspondencia.

1.9.1.2. FUNCIÓN CONTINUA.

Una función numérica θ definida sobre un conjunto $S \subset R^n$ se llama continua en $x_0 \in S$ (con respecto a S si cualquiera de las dos condiciones siguientes equivalentes son satisfechas.

(I) Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que si $\boldsymbol{x} \in S$, y $\|\boldsymbol{x} - \boldsymbol{x_0}\| < \delta$ entonces

$$\|\theta(x) - \theta(\boldsymbol{x_0})\| < \varepsilon$$

(II) Para cada sucesión $x_1, x_2, \cdots,$ en S que converge a $\boldsymbol{x_0}$,

$$\lim_{n\to\infty} \boldsymbol{\theta}(\boldsymbol{x_n}) = \boldsymbol{\theta}\,(\lim_{n\to\infty} \boldsymbol{x_n}) = \boldsymbol{\theta}(\boldsymbol{x_0})$$

se dice que θ es continua sobre S (con respecto a S) si es continua (con respecto a S) en cada x_0 de S, o equivalentemente si cualquiera de las condiciones siguientes equivalentes son válidas:

(III) Los conjuntos: $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(\boldsymbol{x}) \leq \alpha\}$ y $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(\boldsymbol{x}) \geq \alpha\}$ son cerrados relativo a S para cada número real α

(IV) Los conjuntos $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(\boldsymbol{x}) > \alpha\}$ y $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(\boldsymbol{x}) < \alpha\}$ son abiertos relativo a S para cada número real α.

(V) El epígrafo (gráfica superior) de θ

$$G_\theta = \{(x, \xi) \,/\, x \in S, \xi \in R, \theta(x) \leq \xi\}$$

y el hipográfo (gráfica inferior) de θ

$$H_\theta = \{(x, \xi) \,/\, x \in S, \xi \in R, \theta(x) \geq \xi\}$$

son cerrados respecto a $S \times R$

En cada una de las condiciones anteriores se observan un par de requerimientos simétricos. Anulando un requerimiento de cada par, llegamos al concepto de funciones semicontinuas.

1.9.1.3. FUNCIÓN SEMICONTINUA INFERIOR.

Una función numérica θ definida sobre un conjunto $S \subset R^n$ se dice que es semicontinua inferior en $x_0 \in S$ (con respecto a S) si cualquiera de las dos siguientes condiciones equivalentes se satisfacen:

(I) Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que si $\boldsymbol{x} \in S$, y $\|\boldsymbol{x} - \boldsymbol{x_0}\| < \delta$ entonces

$$-\varepsilon < \theta(\boldsymbol{x}) - \theta(\boldsymbol{x_0})$$

(II) Para cada sucesión $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots,$ en S que converge a $\boldsymbol{x_0}$,

$$\lim_{n\to\infty} \boldsymbol{inf}\, \boldsymbol{\theta}(\boldsymbol{x_n}) \geq \boldsymbol{\theta}\,(\lim_{n\to\infty} \boldsymbol{x_n}) = \boldsymbol{\theta}(\boldsymbol{x_0})$$

donde $\lim_{n\to\infty} inferior\, \theta(x_n)$ denota el ínfimo de los puntos límites de la sucesión de números reales $\theta(\boldsymbol{x_1}), \theta(\boldsymbol{x_2}), \cdots$

Se dice que θ es semicontinua inferior en S (con respecto a S) si ella es semicontinua inferior (con respecto a S) en cada punto

$\boldsymbol{x_0} \in S$, o equivalentemente si cualquiera de las siguientes condiciones equivalentes son válidas.

(III) El conjunto $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(x) \leq \alpha\}$ es cerrado relativo a S para cada número real α

(IV) El conjunto $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(x) > \alpha\}$ es abierto relativo a S para cada número real α

(V) El epígrafo (gráfica superior) de θ

$$G_\theta = \{(x, \xi\ ) \,/\, x \in S, \xi \in R, \theta(x) \leq \xi\}$$

es cerrado relativo a $S \times R$.

1.9.1.4. FUNCIONES SEMICONTINUAS SUPERIOR.

Una función numérica θ definida sobre un conjunto $S \subset R^n$ se dice que es semicontinua superior en $x_0 \in S$ (con respecto a S) si cualquiera de las dos siguientes condiciones equivalentes se satisfacen:

(I) Para cada $\varepsilon > 0$, existe un $\delta > 0$ tal que si $\boldsymbol{x} \in S$, y $\|\boldsymbol{x} - \boldsymbol{x_0}\| < \delta$ entonces $\theta(\boldsymbol{x}) - \theta(\boldsymbol{x}_0) < \varepsilon$

(II) Para cada sucesión $\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots,$ en S que converge a $\boldsymbol{x_0}$,

$$\lim_{n\to\infty} \boldsymbol{sup}\ \boldsymbol{\theta}(\boldsymbol{x_n}) \leq \boldsymbol{\theta}\,(\lim_{n\to\infty} \boldsymbol{x_n}) = \boldsymbol{\theta}(\boldsymbol{x_0})$$

donde $\lim\limits_{n\to\infty} \boldsymbol{sup}\ \theta(x_n)$ denota el supremo de los puntos límites de la sucesión de números reales $\theta(\boldsymbol{x_1}), \theta(\boldsymbol{x_2}), \cdots$

Se dice que θ es semicontinua superior en S (con respecto a S) si ella es semicontinua superior (con respecto a S) en cada punto $\boldsymbol{x_0} \in S$, o equivalentemente si cualquiera de las siguientes condiciones equivalentes son válidas.

(III) El conjunto $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(x) \geq \alpha\}$ es cerrado relativo a S para cada número real α

(IV) El conjunto $\{\boldsymbol{x} \,/\, \boldsymbol{x} \in S, \theta(x) < \alpha\}$ es abierto relativo a S para cada número real α

(V) El epígrafo (gráfica superior) de θ

$$G_\theta = \{(x, \xi\ ) \,/\, x \in S, \xi \in R, \theta(x) \geq \xi\}$$

es cerrado relativo a $S \times R$.

OBSERVACIÓN. θ es semicontinua inferior $\boldsymbol{x_0} \in S$ (con respecto a S) si y solo si $-\theta$ es semicontinua superior en $\boldsymbol{x_0} \in S$ (con respecto a S). θ es continua en $\boldsymbol{x_0} \in S$ (con respecto a S) si y solo si es semicontinua superior e inferior en $\boldsymbol{x_0} \in S$ (con respecto a S)

Ejemplo.

a) $\theta(x) = \begin{cases} x, \ para\ x \neq 1 \\ \frac{1}{2}\ para\ x = 1 \end{cases}$

es semicontinua inferior en R

b) $(x) = \begin{cases} x^2, \ para\ x \neq 0 \\ \frac{1}{2}\ \ para\ x = 0 \end{cases}$ es semicontinua superior en R

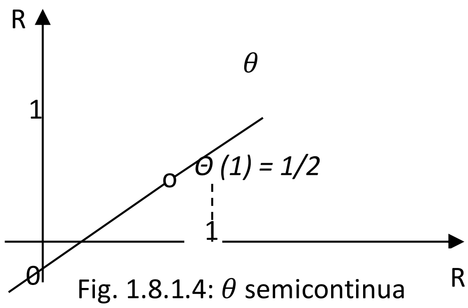


Fig. 1.8.1.4: θ semicontinua inferior sobre R

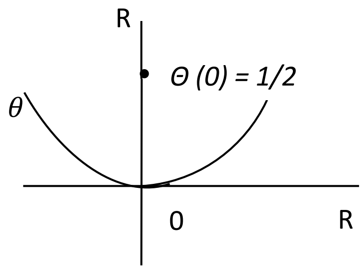


Fig. 1.8.1.4: θ semicontinua superior sobre R

1.9.1.5. TEOREMA.

Sea $(\theta_i)_{i \in I}$ una familia (finita o infinita) de funciones semicontinuas inferior sobre $S \subset R^n$. La mínima cota superior

$$\theta(x) = \sup_{i \in I} \theta_i(x)$$

es semicontinua inferíos sobre S.

Si I es finito, entonces la máxima cota inferior

$$\emptyset(x) = \inf_{i \in I} \theta_i(x)$$

también es semicontinua inferior sobre S.

DEMOSTRACIÓN.

La primera parte se sigue de 1.8.1.3 (iii) anterior y el teorema 1.2.4 (i) si nosotros observamos que para cualquier real λ, el conjunto

$$\{x \,/\, \theta(x) \leq \lambda\} = \left\{x \,/\, \sup_{i\in I} \theta_i(x) \leq \lambda\right\} = \bigcap_{i\in I} \{x \,/\, \theta_i(x) \leq \lambda\}$$

es cerrado (relativo a S). La segunda parte se sigue de 1.8.1.3 (iii) y el teorema 1.2.4 (ii) si observamos que para cualquier real λ el conjunto

$$\{x \,/\, \emptyset(x) \leq \lambda\} = \left\{x \,/\, \inf_{i\in I} \theta_i(x) \leq \lambda\right\} = \bigcup_{i\in I} \{x \,/\, \theta_i(x) \leq \lambda\}$$

es cerrado (relativo a S)

1.9.1.6. COROLARIO.

Sea $(\theta_i)_{i\in I}$ una familia (finita o infinita) de funciones semicontinuas superior sobre $S \subset R^n$. La máxima cota inferior

$$\emptyset(x) = \inf_{i\in I} \theta_i(x)$$

es semicontinua superior sobre S

Si I es finito, entonces la mínima cota superior

$$\theta(x) = \sup_{i\in I} \theta_i(x)$$

también es semicontinua superior sobre S

1.10. FUNCIONES CRECIENTES Y DECRECIENTES EN LOS REALES.

Sea $\theta: S \subset \mathbb{R} \to \mathbb{R}$, tal que $\boldsymbol{y} = \theta(\boldsymbol{x})$ con $\boldsymbol{x} \in \mathbb{R}$; $\boldsymbol{y} \in \mathbb{R}$

- ✓ Se dice que θ es función creciente en S, si $x_1, x_2 \in S$ tal que $x_1 \leq x_2$ entonces $\theta(x_1) \leq \theta(x_2)$.
- ✓ Se dice que θ es función decreciente en S, si $x_1, x_2 \in S$ tal que $x_1 \leq x_2$ entonces $\theta(x_1) \geq \theta(x_2)$.

Por ejemplo

a) $\boldsymbol{\theta(x) = 2x^2 + 4}$ tiene dominio todos los números reales.
Sin embargo, esta función es creciente en $\boldsymbol{S_1} = [\boldsymbol{0}, \infty)$.Esta aseveración se comprueba tomando $\boldsymbol{x_1 = 1, x_2 = 2} \in \boldsymbol{S_1}$ y verificando que $\boldsymbol{\theta(1) = 6 \leq \theta(2) = 12}$ y es decreciente en $\boldsymbol{S_1} = [\![-\infty, \boldsymbol{0}]$.

Obsérvese que $\boldsymbol{S_1} \cup \boldsymbol{S_2} = \mathbb{R}$ =dominio de $\boldsymbol{\theta}$
La gráfica de $\boldsymbol{\theta(x) = 2x^2 + 4}$ es una parábola con vértice en $(\mathbf{0}, \mathbf{4})$ que se abre hacia arriba porque $a = 2 > 0$.
La pendiente de la recta tangente a una función creciente es positiva, mientras que la pendiente a la recta tangente a una función decreciente es negativa.

1.11. ÍNFIMO (SUPREMO) Y MÍNIMO (MÁXIMO) DE UN CONJUNTO DE NÚMEROS REALES.

Recordamos que en la sección 1.3.6, se ha definido el ínfimo y el supremo de un conjunto de números reales de la siguiente manera:

$$\overline{\alpha} = \inf S \Leftrightarrow \begin{cases} x \in S \implies x \geq \overline{\alpha} \\ \forall\, \varepsilon > 0 \implies \exists x \in S: \overline{\alpha} + \varepsilon > x \end{cases}$$

$$\overline{\beta} = \sup S \Leftrightarrow \begin{cases} x \in S \implies x \leq \overline{\beta} \\ \forall\, \varepsilon > 0 \implies \exists x \in S: \overline{\beta} - \varepsilon < x \end{cases}$$

También se observó que ni $\overline{\alpha}$ ni $\overline{\beta}$ necesitan pertenecer a S, pero, cuando $\overline{\alpha}$ y $\overline{\beta}$ pertenecen a S, ellos se llaman min $\boldsymbol{S}$ y máx $\boldsymbol{S}$, respectivamente.

1.11.1. MÍNIMO (MÁXIMO).

Sea S un conjunto de números reales. Si $(\inf S) \in S$, entonces al $\inf S$ se llama el mínimo de S y se denota por $\min S$. Si $(\sup S) \in S$, al $\sup S$ se llama el máximo de S y se denota por $\max S$.

$$\overline{\alpha} = \min S \Leftrightarrow \begin{cases} \overline{\alpha} \in S \\ x \in S \implies x \geq \overline{\alpha} \end{cases}$$

$$\overline{\beta} = \text{máx}\, S \Leftrightarrow \begin{cases} \overline{\beta} \in S \\ x \in S \implies x \leq \overline{\beta} \end{cases}$$

1.11.2. ÍNFIMO (SUPREMO) Y MÍNIMO (MÁXIMO) DE UNA FUNCIÓN NUMÉRICA.

1.11.2.1. FUNCIONES ACOTADAS.

Una función numérica θ definida sobre el conjunto S se dice que está acotada por abajo sobre S si existe un número α tal que:

$$\boldsymbol{x} \in S \implies \theta(\boldsymbol{x}) \geq \alpha.$$

El número α es una cota inferior de θ sobre S.

La función numérica θ se dice que está acotado por arriba sobre S si existe un número β tal que:

$$\boldsymbol{x} \in S \implies \theta(\boldsymbol{x}) \leq \beta$$

El número β es una cota superior de θ sobre S.

Una función numérica θ definida sobre un conjunto S se dice que está acotada si existe un número real $M > 0$ tal que $|\theta(\boldsymbol{x})| \leq M$.

1.11.3. ÍNFIMO DE UNA FUNCIÓN NUMÉRICA.

Sea θ una función numérica definida sobre un conjunto S. Si existe un número $\overline{\alpha}$ tal que:

$x \in S \implies \theta(x) \geq \overline{\alpha}.$ y $\forall \varepsilon > 0 \Longrightarrow \langle \exists\, x \in S : \ \overline{\alpha} + \varepsilon > \theta(x) \rangle$

Entonces $\overline{\alpha}$ se llama el ínfimo de θ sobre S, y se escribe

$$\overline{\alpha} = \inf_{x \in S} \theta(x)$$

1.11.4. SUPREMO DE UNA FUNCIÓN NUMÉRICA.

Sea θ una función numérica definida sobre un conjunto S. Si existe un número $\overline{\beta}$ tal que:

$$x \in S \implies \theta(x) \leq \overline{\beta}.$$

$$y$$

$$\forall \varepsilon > 0 \Longrightarrow \langle \exists\, x \in S : \ \overline{\beta} - \varepsilon < \theta(x) \rangle$$

Entonces $\overline{\beta}$ se llama el supremo de θ sobre S, y se escribe

$$\overline{\beta} = sup_{x \in S}\, \theta(x)$$

Si admitimos los puntos $\pm\infty$, entonces cada función numérica θ tiene un supremo y un ínfimo sobre el conjunto S en el cual está definida.

MÍNIMO DE UNA FUNCIÓN NÚMERICA.

Sea θ una función numérica definida sobre un conjunto S. Si existe un número $\overline{x}$ tal que:

$x \in S \implies \theta(x) \geq \theta(\overline{x})$.

Entonces $\theta(\overline{x})$ se llama el mínimo de θ sobre S, y se escribe

$$\theta(\overline{x}) = \min_{x \in S} \theta(x)$$

Hacemos las siguientes observaciones:

(I) No toda función numérica tiene un mínimo; por ejemplo e^{-x} y x no tienen mínimo sobre R.
(II) El mínimo de una función numérica, si existe, debe ser finito
(III) El máximo de una función numérica, si existe, es un ínfimo alcanzado, esto es:

$$\theta(\overline{x}) \min_{x\varepsilon R} \theta(x) = \inf_{x\varepsilon R} \theta(x)$$

1.11.5. MÁXIMO DE UNA FUNCIÓN NUMÉRICA.

Sea θ una función numérica definida sobre un conjunto S. Si existe un número $\overline{x}$ tal que:

$x \in S \implies \theta(x) \leq \theta(\overline{x})$.

Entonces $\theta(\overline{x})$ se llama el máximo de θ sobre S, y se escribe

$$\theta(\overline{x}) = \max_{x \in S} \theta(x)$$

Observaciones similares a (i), (ii), y (iii) dadas antes, las cuales se aplican al mínimo de una función, también se aplican al máximo de una función.

1.12. EXISTENCIA DE UN MÍNIMO Y UN MÁXIMO DE UNA FUNCIÓN NUMÉRICA.

En seguida presentamos condiciones suficientes para que una función numérica definida sobre un subconjunto de R^n tenga un mínimo y un máximo sobre este subconjunto.

1.12.1. TEOREMA.

Una función semicontinua inferior (superior) definida sobre un conjunto compacto (esto es, cerrado y acotado) S en R^n está acotado por abajo (arriba) y alcanza en S el valor:

$$\overline{\alpha} = \inf_{x \in S} \theta(x) \qquad \left[\overline{\beta} = \sup_{x \in S} \theta(x)\right]$$

En otras palabras existe un $\overline{x} \in S$ tal que:

$$x \in S \implies \theta(x) \geq \theta(\overline{x}) \qquad [x \in S \implies \theta(x) \leq \theta(\overline{x})]$$

DEMOSTRACIÓN.

Probaremos el caso semicontinua inferior. Sea $\gamma > \overline{\alpha},$ entonces el conjunto

$$\Lambda_\gamma = \{x \,/\, x \in S, \theta(x) \leq \gamma\}$$

es no vacío y es cerrado relativo a S por 1.8.1.2 (iii). En consecuencia por teorema de la intersección finita 1.4.9 (iii)

$$\bigcap_{\gamma > \alpha} \Lambda_\gamma \neq \emptyset$$

$$\theta(\overline{x}) \min_{x \varepsilon R} \theta(x) = \inf_{x \varepsilon R} \theta(x)$$

1.12.2. MÁXIMO DE UNA FUNCIÓN NUMÉRICA.

Sea θ una función numérica definida sobre un conjunto S. Si existe un número $\overline{x}$ tal que:

$$x \in S \implies \theta(x) \leq \theta(\overline{x}).$$

Entonces $\theta(\overline{x})$ se llama el máximo de θ sobre S, y se escribe

$$\theta(\overline{x}) = \max_{x \in S} \theta(x)$$

Observaciones similares a (i), (ii), y (iii) dadas antes, las cuales se aplican al mínimo de una función, también se aplican al máximo de una función.

1.13. EXISTENCIA DE UN MÍNIMO Y UN MÁXIMO DE UNA FUNCIÓN NUMÉRICA.

En seguida presentamos condiciones suficientes para que una función numérica definida sobre un subconjunto de R^n tenga un mínimo y un máximo sobre este subconjunto.

1.13.1. TEOREMA.

Una función semicontinua inferior (superior) definida sobre un conjunto compacto (esto es, cerrado y acotado) S en R^n está acotado por abajo (arriba) y alcanza en S el valor:

$$\overline{\alpha} = \inf_{x \in S} \theta(x) \qquad \left[\overline{\beta} = \sup_{x \in S} \theta(x)\right]$$

En otras palabras existe un $\overline{x} \in S$ tal que:

$$x \in S \implies \theta(x) \geq \theta(\overline{x}) \qquad [x \in S \implies \theta(x) \leq \theta(\overline{x})]$$

DEMOSTRACIÓN.

Probaremos el caso semicontinua inferior. Sea $\gamma > \overline{\alpha}$, entonces el conjunto

$$\Lambda_\gamma = \{x \,/\, x \in S, \theta(x) \leq \gamma\}$$

es no vacío y es cerrado relativo a S por 1.8.1.2 (iii). En consecuencia por teorema de la intersección finita 1.4.9 (iii)

$$\bigcap_{\gamma > \alpha} \Lambda_\gamma \neq \emptyset$$

Seleccionando $\overline{x}$ en esta intersección; entonces $\theta(\overline{x}) = \overline{\alpha}$, y por consiguiente $\overline{\alpha}$ es finito.

Remarcaremos aquí que el teorema anterior no puede ser fortalecido agregando la semi continuidad o la suposición de compacidad. En otras palabras, a fin de seguirse que

$$\inf_{x \in S} \theta(x) = \min_{x \in S} \theta(x)$$

no podemos debilitar las condiciones que

(I) θ es semicontinua inferior sobre S.

(II) S es cerrado, y

(III) S es acotado.

En seguida presentamos ejemplos donde el ínfimo no es alcanzado siempre que una cualquiera de las condiciones anteriores son violadas.

i) $\theta(x) = \begin{cases} 1/2 \quad para\, x = 0, x \in R \\ x \;\; para \;\; 0 < x \leq 1, x \in R. \end{cases}$

$\inf\limits_{0 \leq x \leq 1} \theta(x) = 0,$ pero no existe el mínimo sobre el conjunto compacto

$$\{x \,/\, 0 \leq x \leq 1\}$$

ii) $\theta(x) = x$ para $0 < x < 1, x \in R$

$\inf\limits_{0 \leq x \leq 1} \theta(x) = 0,$ pero no existe mínimo de esta función continua sobre el conjunto abierto $\{x \,/\, 0 < x < 1\}$

iii) $\theta(x) = e^{-x},\;\; x \in R$

$\inf\limits_{x \in R} \theta(x),$ pero no existe mínimo de esta función continua sobre el conjunto cerrado no acotado R

1.14. FUNCIONES DIFERENCIABLES, TEOREMAS DEL VALOR MEDIO

En esta sección presentaremos definiciones y propiedades de funciones diferenciables y daremos los teoremas del valor medio y de la función implícita.

1.14.1. FUNCIONES DIFERENCIABLES Y DOS VECES DIFERENCIABLES.

1.14.1.1. FUNCIÓN NUMÉRICA DIFERENCIABLE.

Sea θ una función numérica definida sobre un conjunto abierto S en R^n, y sea $\overline{x}$ en S. Se dice que θ es diferenciable en $\overline{x}$ si para todo $x \in R^n$ tal que $\overline{x} + x \in S$ tenemos que

$$\theta(\overline{x} + x) = \theta(x) + t(\overline{x})x + \propto (\overline{x}, x)\|x\|$$

donde $t(\overline{x})$ es un vector acotado n – dimensional, y $\propto$ es una función numérica tal que $\lim_{x \to 0} \propto (\overline{x}, x) = 0$

Se dice que θ es diferenciable en S si es diferenciable en cada $\overline{\boldsymbol{x}}$ de S [Obviamente, si θ es fiferenciable sobre el conjunto abierto S, es también diferenciable sobre cualquier subconjunto T (abierto o no) de S. Sin embargo cuando decimos que θ es diferenciable sobre algún conjunto T (abierto o no), entendemos que θ es diferenciable sobre algún conjunto abierto que contiene a T.]

1.14.1.2. DERIVADAS PARCIALES Y GRADIENTE DE UNA FUNCIÓN NUMÉRICA.

Sea θ una función numérica definida sobre un conjunto abierto S en R^n, y sea $\overline{\boldsymbol{x}}$ en S. Se dice que θ tiene derivada parcial en $\overline{\boldsymbol{x}}$ con respecto a $x_i, i = 1,2, \cdots, n$; si

$$\frac{\theta(\overline{x_1},\ \overline{x_2}, \cdots, \overline{x_{i-1}}, \overline{x_i} + \delta, \overline{x_{i+1}}, \cdots, \overline{x_n}) - \theta(\overline{x_1}, \cdots, \overline{x_n})}{\boldsymbol{\delta}}$$

tiende a un límite finito cuando δ tiende a cero. A este límite se llama la derivada parcial de θ con respecto a x_i en $\overline{\boldsymbol{x}}$ y se denota por $\partial\theta(\overline{x})/\partial x_i$. El vector n- dimensional de las derivadas parciales de θ con respecto a $x_1, x_2, \cdots, x_n$ en $\overline{\boldsymbol{x}}$ se llama el gradiente de $\boldsymbol{\theta}$ en $\overline{\boldsymbol{x}}$ y se denotas por $\boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}})$, esto es,

$$\boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}}) = (\partial\theta(\overline{x})/\partial x_1, \partial\theta(\overline{x})/\partial x_2, \cdots, \partial\theta(\overline{x})/\partial x_n)$$

1.14.1.3. TEOREMA.

Sea θ una función numérica definida sobre un conjunto abierto S en R^n, y sea $\overline{\boldsymbol{x}}$ en S.

(I) Si θ es diferenciable en $\overline{x}$, entonces θ es continua en $\overline{x}$, y $\boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}})$ existe (pero no recíprocamente), y

$$\theta(\overline{x}+x)=\theta(x)+\nabla\theta(\overline{x})x+\propto(\overline{x},x)\|x\|\ tal\ que$$

$$\lim_{x\to 0}\propto(\overline{x},x)=0, \qquad para\ \overline{x}+x\in S$$

(II) Si θ tiene derivadas parciales continuas en $\overline{x}$ con respecto a $x_1, x_2, \cdots, x_n$, esto es, $\boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}})$ existe y $\boldsymbol{\nabla\theta}$ es continua en $\overline{\boldsymbol{x}}$, entonces θ es diferenciable en $\overline{\boldsymbol{x}}$.

Podemos resumir el resultado anterior como sigue

a) Si θ es diferenciable en $\overline{x}$ entonces θ es continua en $\overline{x}$ y $\nabla\theta(\overline{x})$ existe
b) Si $\nabla\theta(\overline{x})$ existe entonces θ es diferenciable en $\overline{x}$ entonces θ no es diferencuable en $\overline{x}$
c) Si $\nabla\theta$ es continua en $\overline{x}$ y $\nabla\theta(\overline{x})$ eziste entonces Si θ es diferenciable en $\overline{x}$

Equivalentemente:

$$\langle\theta\ diferenciable\ en\ \overline{x}\rangle \underset{\not\Leftarrow}{\Rightarrow} \begin{cases}\theta\ es\ continua\ en\ \overline{x}\\ \boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}})\ existe.\end{cases}$$

$$\langle\theta\ diferenciable\ en\ \overline{x}\rangle \Leftarrow \begin{cases}\boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}})\ existe.\\ \boldsymbol{\nabla\theta}\ es\ continua\ en\ \overline{\boldsymbol{x}}\end{cases}$$

1.14.1.4. FUNCIÓN VECTORIAL DIFERENCIABLE.

Sea f una función vectorial m--dimensional definida sobre un conjunto abierto S en R^n, y sea $\overline{x}$ en S. Se dice que f es diferenciable en $\overline{x}$ (respectivamente sobre S) si cada una de sus componentes $f_1, f_2, \cdots, f_m$ son diferenciables en $\overline{x}$ (respectivamente sobre S)

1.14.1.5. DERIVADAS PARCIALES Y JACOBIANO DE UNA FUNCIÓN VECTORIAL.

Sea f una función vectorial m--dimensional definida sobre un conjunto abierto S en R^n y sea $\overline{x}$ en S. Se dice que f tiene derivadas parciales en $\overline{\boldsymbol{x}}$ con respecto a $x_1, x_2, \cdots, x_n$, si cada una de sus componentes $f_1, f_2, \cdots, f_m$ tienen derivadas parciales en $\overline{\boldsymbol{x}}$ con respecto a $x_1, x_2, \cdots, x_n$.

La matriz mxn definida por

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f_1(\overline{x})}{\partial x_1} & \frac{\partial f_1(\overline{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\overline{x})}{\partial x_n} \\ \frac{\partial f_2(\overline{x})}{\partial x_1} & \frac{\partial f_2(\overline{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\overline{x})}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m(\overline{x})}{\partial x_1} & \frac{\partial f_m(\overline{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\overline{x})}{\partial x_n} \end{bmatrix}$$

se llama Matriz Jacobiana de f en $\overline{x}$ y a su determinante $|\nabla f(x)|$ se llama Jacobiano

1.14.1.6. TEOREMA (REGLA DE LA CADENA).

Sea f una función vectorial m--dimensional definida sobre un conjunto abierto S en R^n y sea ϕ una función numérica definida en R^n. La función numérica θ definida en S por $\boldsymbol{\theta}(\boldsymbol{x}) = \boldsymbol{\phi}[\boldsymbol{f}(\boldsymbol{x})]$ es diferenciable en $\overline{\boldsymbol{x}} \in S$ si f es diferenciable en $\overline{\boldsymbol{x}}$ y si ϕ es diferenciable en

$\overline{\boldsymbol{y}} = \boldsymbol{f}(\overline{\boldsymbol{x}})$,

y

$$\boldsymbol{\nabla\theta}(\overline{\boldsymbol{x}}) = \boldsymbol{\nabla\phi}(\overline{\boldsymbol{y}})\boldsymbol{\nabla f}(\overline{\boldsymbol{x}})$$

1.14.1.7. FUNCIÓN NUMÉRICA DOS VECES DIFERENCIABLE Y SU MATRIZ HESSIANA.

Sea f una función vectorial m--dimensional definida sobre un conjunto abierto S en R^n y sea $\overline{\boldsymbol{x}}$ en S. Se dice que θ es dos veces diferenciable (es decir sus segundas derivadas parciales existen) en $\overline{x}$ si para todo $\boldsymbol{x} \in R^n$ tal que $\overline{\boldsymbol{x}} + \boldsymbol{x} \in S$ tenemos que

$$\theta(\overline{\boldsymbol{x}} + \boldsymbol{x}) = \theta(\overline{\boldsymbol{x}}) + \nabla\theta(\overline{x})\boldsymbol{x} + \frac{1}{2}\nabla^2\theta(\overline{\mathrm{x}})\boldsymbol{x} + \beta(\overline{\boldsymbol{x}}, \boldsymbol{x})(\|x\|)^2$$

donde $\nabla^2\theta(\overline{\boldsymbol{x}})$ es una matriz nxn de elementos acotados, y β es una función numérica de $\boldsymbol{x}$ tal que

$$\lim_{x \to 0} \beta(\overline{\boldsymbol{x}}, \boldsymbol{x}) = 0$$

La matriz $\nabla^2\theta(\bar{\boldsymbol{x}})$ se llama la matriz Hessiana de θ en $\bar{x}$ y sus ij –ésimos elementos se escriben como

$$[\nabla^2\theta(\bar{\boldsymbol{x}})]_{ij} = \frac{\partial^2\theta(\bar{x})}{\partial x_i \partial x_j} \qquad i,j = 1,2,\cdots,n.$$

Obviamente, si θ es dos veces diferenciable en $\bar{\boldsymbol{x}}$, también debe ser diferenciable en $\bar{\boldsymbol{x}}$.

1.14.1.8. TEOREMA.

Sea θ una función numérica definida sobre un conjunto abierto S en R^n, y sea $\bar{\boldsymbol{x}}$ en S. Entonces

(I) Si $\nabla\theta$ es diferenciable en $\bar{\boldsymbol{x}}$, entonces θ es dos diferenciable en $\bar{\boldsymbol{x}}$.

(II) Si $\nabla\theta$ tiene derivadas parciales continuas en $\bar{\boldsymbol{x}}$, entonces θ es dos veces diferenciables en $\bar{\boldsymbol{x}}$.

(III) Si $\nabla^2\theta$ es continua en $\bar{\boldsymbol{x}}$, entonces

$$\frac{\partial^2\theta(\bar{x})}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial\theta(\bar{x})}{\partial x_j}\right)$$

y $\nabla^2\theta(\bar{\boldsymbol{x}})$ es simétrica, esto es, $[\nabla^2\theta(\bar{\boldsymbol{x}})]_{ij} = [\nabla^2\theta(\bar{\boldsymbol{x}})]_{ji}$

OBSERVACIÓN.

1. Los números $\partial\theta(\bar{x})/\partial x_i$, $i = 1,2,\cdots,n$, también son llamados la primera derivada parcial de θ en $\bar{\boldsymbol{x}}$, y $\partial^2\theta(\bar{x})/\partial x_i\,\partial x_j$, $i,j = 1,2,\cdots,n$ también son llamadas las segundas derivadas parciales de θ en $\bar{\boldsymbol{x}}$. De manera análoga podemos definir la $k-$ derivada parcial θ en $\bar{\boldsymbol{x}}$.

2. Sea θ una función numérica definida sobre un conjunto abierto $S \subset R^n \times R^k$ la cual es diferenciable en $(\bar{x},\bar{y}) \in S$.

Entonces definimos los gradientes

$$\nabla_x\theta(\bar{x},\bar{y}) = \left[\frac{\partial\theta(\bar{x},\bar{y})}{\partial x_1} \quad \cdots \quad \frac{\partial\theta(\bar{x},\bar{y})}{\partial x_n}\right]$$

$$\nabla_y\theta(\bar{x},\bar{y}) = \left[\frac{\partial\theta(\bar{x},\bar{y})}{\partial y_1} \quad \cdots \quad \frac{\partial\theta(\bar{x},\bar{y})}{\partial y_n}\right]$$

y, $\nabla\theta(\overline{x},\overline{y}) = \left[\nabla_x\theta(\overline{x},\overline{y}), \nabla_y\theta(\overline{x},\overline{y})\right]$

DEFINICIÓN. Sea f una función vectorial m--dimensional definida sobre un conjunto abierto $S \subset R^n \times R^k$ la cual es diferenciable en $(\overline{x},\overline{y}) \in S$

Entonces definimos las matrices Hessianas siguientes:

$$\nabla_x f(\overline{x},\overline{y}) = \begin{bmatrix} \frac{\partial f_1(\overline{x},\overline{y})}{\partial x_1} & \frac{\partial f_1(\overline{x},\overline{y})}{\partial x_2} & \cdots & \frac{\partial f_1(\overline{x},\overline{y})}{\partial x_n} \\ \frac{\partial f_2(\overline{x},\overline{y})}{\partial x_1} & \frac{\partial f_2(\overline{x},\overline{y})}{\partial x_2} & \cdots & \frac{\partial f_2(\overline{x},\overline{y})}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m(\overline{x},\overline{y})}{\partial x_1} & \frac{\partial f_m(\overline{x},\overline{y})}{\partial x_2} & \cdots & \frac{\partial f_m(\overline{x},\overline{y})}{\partial x_n} \end{bmatrix}$$

$$\nabla_y f(\overline{x},\overline{y}) = \begin{bmatrix} \frac{\partial f_1(\overline{x},\overline{y})}{\partial y_1} & \frac{\partial f_1(\overline{x},\overline{y})}{\partial y_2} & \cdots & \frac{\partial f_1(\overline{x},\overline{y})}{\partial y_k} \\ \frac{\partial f_2(\overline{x},\overline{y})}{\partial y_1} & \frac{\partial f_2(\overline{x},\overline{y})}{\partial y_2} & \cdots & \frac{\partial f_2(\overline{x},\overline{y})}{\partial y_k} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m(\overline{x},\overline{y})}{\partial y_1} & \frac{\partial f_m(\overline{x},\overline{y})}{\partial y_2} & \cdots & \frac{\partial f_m(\overline{x},\overline{y})}{\partial y_k} \end{bmatrix}$$

y el gradiente $\nabla f(\overline{x},\overline{y}) = \left[\nabla_x f(\overline{x},\overline{y}) \;\; \nabla_y f(\overline{x},\overline{y})\right]$

1.15. TEOREMA DEL VALOR MEDIO Y TEOREMA DE TAYLOR.

1.15.1. TEOREMA DEL VALOR MEDIO.

Sea θ una función numérica diferenciable definida sobre un conjunto convexo abierto $S \subset R^n$, y sean $\boldsymbol{x_1}, \boldsymbol{x_2} \in S$. Entonces

$$\theta(\boldsymbol{x_2}) = \theta(\boldsymbol{x_1}) + \nabla\theta[\mathrm{x}_1 + \gamma(\boldsymbol{x_2} - \boldsymbol{x_1})](\boldsymbol{x_2} - \boldsymbol{x_1})$$

para algún número real $\gamma,\ 0 < \gamma < 1$

$$\theta(\boldsymbol{x_2}) = \theta(\boldsymbol{x_1}) + \nabla\theta[\boldsymbol{x_1}](\boldsymbol{x_2} - \boldsymbol{x_1}) + \frac{1}{2}\{(\boldsymbol{x_2} - \boldsymbol{x_1})\nabla^2\theta[\boldsymbol{x_1} + \delta(\boldsymbol{x_2} - \boldsymbol{x_1})](\boldsymbol{x_2} - \boldsymbol{x_1})\}$$

1.15.2. TEOREMA DE TAYLOR (DE SEGUNDO ORDEN)

Sea θ una función numérica dos veces diferenciable definida sobre un conjunto convexo abierto $S \subset R^n$, y sean $\boldsymbol{x_1}, \boldsymbol{x_2} \in S$. Entonces

$$\theta(\boldsymbol{x_2}) = \theta(\boldsymbol{x_1}) + \nabla\theta[\boldsymbol{x_1}](\boldsymbol{x_2} - \boldsymbol{x_1})$$
$$+ \frac{1}{2}\{(\boldsymbol{x_2} - \boldsymbol{x_1})\nabla^2\theta[\boldsymbol{x_1} + \delta(\boldsymbol{x_2} - \boldsymbol{x_1})](\boldsymbol{x_2} - \boldsymbol{x_1})\}$$

para algún número real $\delta,\ 0 < \delta < 1$

2. FORMAS CUADRÁTICAS

En este apartado vamos a estudiar el concepto de forma cuadrática, las formas de expresarlas así como los tipos y propiedades de ellas.

2.1. CONCEPTO DE FORMA CUADRÁTICA.

Para introducir el concepto de forma cuadrática vamos a iniciar el estudio de los llamados polinomios cuadráticos

2.2. DEFINICIÓN (POLINOMIO CUADRÁTICO)

Diremos que un polinomio p en las variables $x_1, x_2, \cdots, x_n$ es cuadrático, si cada uno de sus términos tiene grado dos, es decir

$$\boldsymbol{p}(\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_n}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} \boldsymbol{a_{ij} x_i x_j}$$

donde

$a_{ij} \in R;\ i, j = 1, 2, \cdots, n$ son los coeficientes del polinomio.

Establecido el concepto de polinomio cuadrático, en un conjunto de variables, también podríamos interpretar a este como una aplicación tal que a cada vector n-dimensional $(x_1, x_2, \cdots, x_n)$ de R^n se le asocia el número real $p(\boldsymbol{x})$. En seguida presentamos el concepto de forma cuadrática.

2.3. DEFINICIÓN (FORMA CUADRÁTICA)

Una forma cuadrática Q, es toda aplicación $Q: R^n \to R$ tal que a cada vector columna $n \times 1$, $\boldsymbol{X} = Col(x_1, x_2, \cdots, x_n) \in R^n$ le hace corresponder el valor numérico dado por un polinomio cuadrático

$$Q(\boldsymbol{X}) = q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

donde $A = (a_{ij})$ es una matriz simétrica de orden $n \times n$.:

El vector $\boldsymbol{X^T}$ es un vector fila $\boldsymbol{X^T} = (x_1, x_2, \cdots, x_n) \in R^n$ y $\boldsymbol{X}$ es un vector columna denotado por

$$\boldsymbol{X} = Col(x_1, x_2, \cdots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$$

EXPLICACIÓN. El desarrollo de una forma cuadrática, en términos de las variables $\boldsymbol{x_1, x_2, \cdots, x_n}$ corresponde a un polinomio homogéneo de grado 2, donde los coeficientes de los términos cuadráticos $\boldsymbol{x_i^2}$ son los elementos de la diagonal principal de la matriz simétrica $\boldsymbol{A}$, y cada coeficiente de un término rectangular $\boldsymbol{x_i x_j}$ es el doble del elemento $\boldsymbol{a_{ij}}$ de la misma matriz ($\boldsymbol{i \neq j}$); es decir: coeficiente de un término rectangular $\boldsymbol{x_i x_j = 2a_{ij}}$.

Dada una matriz cuadrada real $M = (m_{ij})$, aunque la matriz no sea simétrica, la función

$$\boldsymbol{Q(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_i x_j}$$

es un polinomio cuadrático. Sea $\boldsymbol{A} = (a_{ij})$ la matriz simétrica asociada a la forma cuadrática, entonces los elementos de A se determinan con

$$\boldsymbol{a_{ij} = \frac{m_{ij} + m_{ji}}{2}}$$

De esta forma: Si $\boldsymbol{M = (m_{ij})}$ es una de las matrices no simétricas asociadas a la forma cuadrática $\boldsymbol{Q}$, entonces su matriz simétrica $\boldsymbol{A}$ asociada a la forma cuadrática se obtiene efectuando la operación

$$\boldsymbol{A = \frac{1}{2}(M + M^T)}$$

De esta manera la forma cuadrática es

$$\boldsymbol{Q}(\boldsymbol{X}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \boldsymbol{a_{ij} x_i x_j}$$

2.4. MATRIZ ASOCIADA A UN FORMA CUADRÁTICA

TEOREMA 01. (Existencia y unicidad de una matriz simétrica asociada a una forma cuadrática)

Dada una forma cuadrática $Q: R^n \rightarrow R$, existe una ÚNICA MATRIZ SIMÉTRICA $\boldsymbol{A}$ de orden $n \times n$ tal que

$$Q(\boldsymbol{X}) = \boldsymbol{X}^T A \boldsymbol{X}$$

y, a esa matriz, se la denomina la MATRIZ ASOCIADA a la forma cuadrática.

Las propiedades de la forma cuadrática $Q(\boldsymbol{X})$ están relacionadas con las propiedades de la matriz simétrica A, y no con todas las posibles matrices M que permiten obtener $Q(X)$ mediante la expresión $Q(\boldsymbol{X}) = \boldsymbol{X}^T M \boldsymbol{X}$

Ejemplos.

1. La matriz simétrica A asociada a la forma cuadrática definida por

$$q(x_1, x_2, x_3) = -x_1^2 + 2x_2^2 + 5x_3^2 + x_1x_2 - 3x_1x_3 + 6x_2x_3$$

es $$A = \begin{bmatrix} -1 & 1/2 & -3/2 \\ 1/2 & 2 & 3 \\ -3/2 & 3 & 5 \end{bmatrix}$$

Explicación.

- Los coeficientes de los términos cuadráticos son $-1, 2, 5$ y forman la diagonal principal de la matriz simétrica asociada a la forma cuadrática.
- Cada elemento a_{ij} es:

$a_{ij} = (Coeficiente\ de\ un\ término\ rectangular\ x_{ij})/2$; esto es $a_{12} = a_{21} = 1/2\,;\ a_{13} = a_{31} = -3/2\,;\ a_{23} = a_{32} = 6/2 = 3$

Si suponemos que $q(x_1, x_2, x_3) = -x_1^2 + 2x_2^2 + 5x_3^2 + x_1x_2 - 3x_1x_3 + 6x_2x_3$ es un polinomio cuadrático y deseamos calcular la matriz simétrica asociada a la forma cuadrática, entonces debemos reescribir el polinomio con todos sus términos rectangulares x_{ij} y x_{ji}, escribir la matriz M y luego calcular $A = (M + M^T)/2$. Esto es:

- $q(x_1, x_2, x_3) = -x_1^2 + 2x_2^2 + 5x_3^2 + x_1x_2 - 3x_1x_3 + 0x_{21} + 6x_2x_3 + 0x_{31} + 0x_{32}$
- La matriz no simétrica M es

$$M = (m_{ij}) = \begin{bmatrix} -1 & 1 & -3 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{bmatrix}$$

El polinomio escrito en forma matricial es $Q(X) = X^T M X$

La matriz simétrica asociada a la forma cuadrática es

$$A = \frac{1}{2}(M + M^T) = \frac{1}{2}\left(\begin{bmatrix} -1 & 1 & -3 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 6 & 5 \end{bmatrix}\right)$$
$$= \begin{bmatrix} -1 & 1/2 & -3/2 \\ 1/2 & 2 & 3 \\ -3/2 & 3 & 5 \end{bmatrix}$$

2. La forma cuadrática asociada a la matriz simétrica

$$A = \begin{bmatrix} 0 & \sqrt{5} & 3 \\ \sqrt{5} & -3 & 0 \\ 3 & 0 & 5 \end{bmatrix}$$

es

$$Q(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}) = \boldsymbol{X}^T A \boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}) \begin{bmatrix} 0 & \sqrt{5} & 3 \\ \sqrt{5} & -3 & 0 \\ 3 & 0 & 5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$Q(x_1, x_2, x_3) = 0x_1^2 - 3x_2^2 + 5x_3^2 + 2\sqrt{5}x_1x_2 + 6x_1x_3 + 0x_2x_3$$

$$Q(x_1, x_2, x_3) = -3x_2^2 + 5x_3^2 + 2\sqrt{5}x_1x_2 + 6x_1x_3$$

3. Sea $M = \begin{bmatrix} 1 & 2 & -1 \\ -1 & -3 & 1 \\ 3 & 0 & 5 \end{bmatrix}$ una matriz cuadrada real no simétrica, la matriz simétrica asociada a la forma cuadrática es

$$A = \frac{1}{2}(A + A^T) = \frac{1}{2}\left(\begin{bmatrix} 1 & 2 & -1 \\ -1 & -3 & 1 \\ 3 & 0 & 5 \end{bmatrix} + \begin{bmatrix} 1 & -1 & 3 \\ 2 & -3 & 0 \\ -1 & 1 & 5 \end{bmatrix}\right) =$$

$$= \frac{1}{2}\begin{bmatrix} 2 & 1 & 2 \\ 1 & -6 & 1 \\ 2 & 1 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 1 \\ 1/2 & -3 & 1/2 \\ 1 & 1/2 & 5 \end{bmatrix}$$

La forma cuadrática es

$$Q(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}) = \boldsymbol{X}^T A \boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}) \begin{bmatrix} 1 & 1/2 & 1 \\ 1/2 & -3 & 1/2 \\ 1 & 1/2 & 5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$Q(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 5x_3^2 + x_1x_2 + x_2x_3 + 2x_1x_3$$

2.5. FORMAS (O MANERAS) DE EXPRESAR UNA FORMA CUADRÁTICA.

Una forma cuadrática se puede escribir en las siguientes formas: Matricial, canónica y polinomial.

2.5.1. Forma matricial de una forma cuadrática.

Definición. Sea forma cuadrática

$$Q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

entonces, la forma matricial es $Q(\boldsymbol{X}) = \boldsymbol{X^T} A\, \boldsymbol{X}$.

donde

$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ y A una matriz cuadrada simétrica de orden $n \times n$ asociada a la forma cuadrática.

1.4.1. Forma canónica de una forma cuadrática.

Definición. Sea

$$Q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

una forma cuadrática, la forma canónica, es decir la forma cuadrática sin términos rectangulares $\boldsymbol{x_i x_j}$ es

$$\boldsymbol{Q(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} a_i\, x_i^2}$$

donde $a_i \in R, i = 1, 2, \cdots . n$.

Cualquier forma cuadrática siempre se puede expresar en forma canónica y su matriz asociad es diagonal.

2.5.2. Forma Polinomial

Definición. Una forma cuadrática está en forma Polinomial cuando la sumatoria está desarrollada dando como consecuencia un polinomio homogéneo de segundo grado.

En otras palabras la forma polinomial de una forma cuadrática se obtiene desarrollando la forma matricial $\boldsymbol{X^T} A\, \boldsymbol{X}$, esto es:

$$Q(\boldsymbol{X}) = \boldsymbol{X}^T A\, \boldsymbol{X} = (x_1, x_2, \cdots, x_n)\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Ejemplo 1. La siguiente forma cuadrática está en forma polinomial, escribirlo en forma canónica

$$Q(x, y, z) = x^2 + 2xy + 6y^2 + 4yz + z^2$$

Solución.

Reescribiendo adecuadamente la forma cuadrática

$$\begin{aligned} Q(x, y, z) &= (x^2 + 2xy + y^2) + (5y^2 + 4yz + z^2) \\ &= (x + y)^2 + 5\left(y^2 + \frac{4}{5}yz\right) + z^2 \end{aligned}$$

$$Q(x, y, z) = (x + y)^2 + 5\left(y^2 + \frac{4}{5}yz + \frac{4}{25}z^2 - \frac{4}{25}z^2\right) + z^2$$

$$= (x + y)^2 + 5\left(y^2 + \frac{4}{5}yz + \frac{4}{25}z^2\right) - \frac{4}{5}z^2 + z^2$$

$$Q(x, y, z) = (x + y)^2 + 5\left(y + \frac{2}{5}z\right)^2 + \frac{1}{5}z^2$$

Introduciendo las nuevas variables y_1, y_2, y_3

$$\begin{aligned} &1)\ \ y_1 = x + y \\ &2)\ \ y_2 = y + \frac{2}{5}z \\ &3)\ \ y_3 = z \end{aligned}$$

y remplazando, escribimos la forma cuadrática en su forma canónica

$$Q(y_1, y_2, y_3) = y_1^2 + 5y_2^2 + \frac{1}{5}y_3^2$$

Ejemplo 2. La siguiente forma cuadrática está en forma polinomial, transfórmelo a la forma matricial.

$$Q(x, y, z) = x^2 + 2xy + 6y^2 + 4yz + z^2$$

Solución.

Reescribimos la forma cuadrática como

$$Q(x,y,z) = x^2 + 2xy + 6y^2 + 4yz + z^2 + 0xz$$

Escribiendo la matriz simétrica asociada a la forma cuadrática dada, resulta:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

La forma matricial es

$$Q(x,y,z) = X^T A X = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Ejemplo 3. Dada la forma cuadrática $Q(x,y,z) = x^2 + 2xy + 6y^2 + 4yz + z^2$

a) Obtener dos matrices distintas (no simétricas) M y N tales que $X^T M \boldsymbol{X} = Q(X)$ y $X^T N \boldsymbol{X} = Q(X)$

b) Obtener una matriz simétrica A tales que $X^T A\, \boldsymbol{X} = Q(x,y,z)$

Solución.

a) Para hallar las dos matrices distintas, reescribimos la forma cuadrática como:

i) $Q(x,y,z) = x^2 + 3xy + 0xz - yx + 6y^2 + yz + 0zx + 3zy + z^2$

ii) $Q(x,y,z) = x^2 + 4xy + xz - 2yx + 6y^2 + 2yz - zx + 2zy + z^2$

Las dos matrices distintas y no simétricas son obtenidas de los polinomios anteriores, ellas son:

$$M = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 6 & 1 \\ 0 & 3 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 4 & 1 \\ -2 & 6 & 2 \\ -1 & 2 & 1 \end{bmatrix}$$

Se observa que ambas dan la misma forma cuadrática dada, puesto que:

$$Q(x,y,z) = X^T M X = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ -1 & 6 & 1 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
$$= x^2 + 2xy + 6y^2 + 4yz + z^2$$

$$Q(x,y,z) = X^T N X = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 4 & 1 \\ -2 & 6 & 2 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
$$= x^2 + 2xy + 6y^2 + 4yz + z^2$$

b) La matriz simétrica A asociada a la forma cuadrática dada es

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

Esta matriz también se puede obtener aplicando

$$A = \frac{1}{2}(M + M^T)$$

- Para M, tenemos:

$$A = \frac{1}{2}(M + M^T) = \frac{1}{2}\left(\begin{bmatrix} 1 & 3 & 0 \\ -1 & 6 & 1 \\ 0 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -1 & 0 \\ 3 & 6 & 3 \\ 0 & 1 & 1 \end{bmatrix} \right) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

- Para N, tenemos:

$$A = \frac{1}{2}(N + N^T) = \frac{1}{2}\left(\begin{bmatrix} 1 & 4 & 1 \\ -2 & 6 & 2 \\ -1 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -2 & -1 \\ 4 & 6 & 2 \\ 1 & 2 & 1 \end{bmatrix} \right) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

A continuación presentamos un método basado en el procedimiento algebraico de completar cuadrados para reducir una forma cuadrática dada en forma polinomial que contiene términos rectangulares $x_i x_j$ a una suma de cuadrados, es decir a su forma canónica. Este método, denominado método de LaGrange se basa en dos ideas sencillas: completar cuadrados y, y algunas veces, usar el hecho de que suma por diferencias es igual a diferencia de cuadrados.

Obviamente el carácter de la forma cuadrática no cambia con las operaciones usadas en el método de Lagrange, lo cual permite clasificar la forma cuadrática según su signo cuando la forma cuadrática está escrita en forma canónica.

a, Formulación general del método de Lagrange.

Sea $Q(X) = X^T A X$ una forma cuadrática. El método que consiste en ir completando cuadrados haciendo cambios de variables en los que en cada paso cambia una (o a lo sumo dos) de las variables, suele denominarse Método de Lagrange. Se presentan dos casos:

CASO I. (Método de completar cuadrados) Si para algún índice i se tiene $a_{ii} \neq 0$, podemos completar cuadrados con todos los términos que contengan a x_i para obtener

$$Q(X) = a_{ii}\left(\sum_{j=1}^{n}\frac{a_{ij}}{a_{ii}}x_j\right)^2 + q_1(x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$$

Donde q_1 es una nueva forma cuadrática con $n - 1$ variables a la que se le vuelve a aplicar el proceso. El cambio de variables que se utiliza es

$$y_i = \sum_{j=1}^{n}\frac{a_{ij}}{a_{ii}}x_j\ ,\ \ y_j = x_j \ \ para\ j \neq i$$

Ejemplo 1.

Aplicar el método de Lagrange para reducir a su forma canónica la siguiente forma cuadrática

$$Q(x_1, x_2, x_3) = x_1^2 + 5x_2^2 + 2x_3^2 + 4x_1x_2 + 2x_1x_3 + 6x_2x_3$$

Solución.

Completamos cuadrados en la variable x_1 puesto que aparecen términos en x_1^2, x_1x_2 y en x_1x_3. Sumamos y restamos $4x_2^2$, $4x_2x_3$, y descomponemos el término $2x_3^2$ como $2x_3^2 = x_3^2 + x_3^2$ tenemos

$$Q(x_1, x_2, x_3) = (x_1^2 + 4x_1x_2 + 2x_1x_3 + 4x_2x_3 + 4x_2^2 + x_3^2\ ) - 4x_2^2 + 5x_2^2 + x_3^2 - 4x_2x_3 + 6x_2x_3$$

$$Q(x_1, x_2, x_3) = (x_1 + 2x_2 + x_3)^2 + x_2^2 + x_3^2 + 2x_2x_3$$

En seguida completamos cuadrados en la variable x_2 puesto que aparecen términos en x_2^2 y en x_2x_3

$$Q(x_1, x_2, x_3) = (x_1 + 2x_2 + x_3)^2 + (x_2 + x_3)^2 = y_1^2 + y_2^2$$

Donde se realizó el cambio de variables $y_1 = x_1 + 2x_2 + x_3$ y, $y_2 = x_2 + x_3, y_3 = x_3$

CASO II. (Método de suma por diferencias) Si $a_{ii} = 0$ para todo i elegimos $a_{ij} \neq 0$ (si todos fueran cero tendríamos $Q(X) = 0$ que ya está reducida)

En este caso hacemos el cambio de variables

$x_i = y_i + y_j,\ x_j = y_i - y_j$ y $x_k = y_k$ para $k \neq i, j$

y pasamos de nuevo al caso (I), pues los términos cuadrados $a_{ij}x_ix_j$ se transforman en:

$$a_{ij}x_ix_j = a_{ij}y_i^2 - a_{ij}y_j^2.$$

Ejemplo 1.

Aplicar el método de Lagrange para reducir a su forma canónica la siguiente forma cuadrática

$$Q(x_1, x_2, x_3, x_4) = 3x_1x_2 + 5x_3x_4$$

Solución.

Vemos que no hay ningún término al cuadrado $(a_{ii} = 0)$, recurrimos al método de suma por diferencias.

Hagamos el cambio de variables

$$x_1 = y_1 + y_2, x_2 = y_1 - y_2, x_3 = y_3, x_4 = y_4$$

Remplazando en la forma cuadrática tenemos

$$Q(y_1, y_2, y_3, y_4) = 3(y_1 + y_2)(y_1 - y_2) + 5y_3y_4 = 3y_1^2 - 3y_2^2 + 5y_3y_4$$

En la última forma cuadrática vemos que hay cuadrados en las dos primeras variables y_1 y y_2 mientras que en las otras dos variables y_3 y y_4 no hay ningún término cuadrático. Esto hace ver que es necesario recurrir a la suma por diferencias, para esto hacemos el cambio de variables

$$y_1 = z_1, \qquad y_2 = z_2, \qquad y_3 = z_3 + z_4, \qquad y_4 = z_3 - z_4$$

Remplazando en la última forma cuadrática tenemos

$$\begin{aligned} Q(z_1, z_2, z_3, z_4) &= 3z_1^2 - 3z_2^2 + 5(z_3 + z_4)(z_3 - z_4) \\ &= 3z_1^2 - 3z_2^2 + 5z_3^2 - z_4^2 \end{aligned}$$

Obsérvese que el mismo resultado se obtendría si se hace, a la vez, los cambios de variable

$$x_1 = z_1 + z_2, \qquad x_2 = z_1 - z_2, \qquad x_3 = z_3 + z_4, \qquad x_4 = z_3 - z_4$$

En efecto

$$\begin{aligned} Q(z_1, z_2, z_3, z_4) &= 3(z_1 + z_2)(z_1 - z_2) + 5(z_3 + z_4)(z_3 - z_4) \\ &= 3z_1^2 - 3z_2^2 + 5z_3^2 - 5z_4^2 \end{aligned}$$

2.6. CLASIFICACIÓN DE UNA FORMA CUADRÁTICA SEGÚN SU SIGNO.

El análisis de una forma cuadrática real de n variables se centra, fundamentalmente, en determinar su signo, es decir, si todos los valores reales de $Q(\boldsymbol{X})$, para $\boldsymbol{X} \in R^n$, mantienen el mismo signo o no. Las formas cuadráticas se pueden clasificar como sigue:

Definición.- Sea $Q: R^n \to R$ una forma cuadrática definida por

$$\boldsymbol{Q}(\boldsymbol{X}) = \boldsymbol{X^T A X} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} \boldsymbol{a_{ij} x_i x_j}$$

Se dice que:

1. $Q(\boldsymbol{X})$ es definida positiva (D.P.) si $Q(\boldsymbol{X}) > 0,\ \forall \boldsymbol{X} \in R^n,\ \boldsymbol{X} \neq 0$
2. $Q(\boldsymbol{X})$ es definida negativa (D.N.) si $Q(\boldsymbol{X}) < 0,\ \forall \boldsymbol{X} \in R^n,\ \boldsymbol{X} \neq 0$
3. $Q(\boldsymbol{X})$ es semidefinida positiva (S.D.P.) si $Q(\boldsymbol{X}) \geq 0,\ \forall \boldsymbol{X} \in R^n\ y\ \exists\ \boldsymbol{X} \neq 0 \ / \ Q(\boldsymbol{X}) = 0$
4. $Q(\boldsymbol{X})$ es semidefinida negativa (S.D.N.) si $Q(\boldsymbol{X}) \leq 0,\ \forall \boldsymbol{X} \in R^n\ y\ \exists\ \boldsymbol{X} \neq 0 \ / \ Q(\boldsymbol{X}) = 0$
5. $Q(\boldsymbol{X})$ es indefinida si $\exists\ \boldsymbol{X_1}, \boldsymbol{X_2} \in R^n / \ Q(\boldsymbol{X_1}) < 0\ y\ \ Q(\boldsymbol{X_2}) > 0$

Ejemplo 1.

a) La forma cuadrática $\mathrm{Q}(\mathrm{x}_1, \mathrm{x}_2) = \mathrm{x}_1^2 + \mathrm{x}_2^2$ es D.P, ya que está escrita mediante una suma de cuadrados y, por tanto, toma valores positivos cualquiera que sea el vector que se use, salvo el vector nulo.
b) La forma cuadrática $\mathrm{Q}(\mathrm{x}_1, \mathrm{x}_2) = -\mathrm{x}_1^2 - \mathrm{x}_2^2$ es D.N, ya que puede ser escrita como $\mathrm{q}(\mathrm{x}_1, \mathrm{x}_2) = -(\mathrm{x}_1^2 + \mathrm{x}_2^2)$ y, por tanto, toma valores negativos cualquiera que sea el vector que se use, salvo el vector nulo.
c) La forma cuadrática $\mathrm{q}(\mathrm{x}_1, \mathrm{x}_2) = (\mathrm{x}_1 - \mathrm{x}_2)^2$ es S.D.P. pues, $(\mathrm{x}_1 - \mathrm{x}_2)^2 \geq 0\ \ \forall \mathrm{x} \in \mathrm{R}^2$ y existe vectores $\mathrm{x} = (\mathrm{x}_1, \mathrm{x}_2) \in \mathrm{R}^2 / \ \mathrm{x}_1 = \mathrm{x}_2\ \ \mathrm{y}\ \ \mathrm{q}(\mathrm{x}_1, \mathrm{x}_2) = 0$. Por ejemplo $\exists\ \mathrm{x} = (2, 2) \in \mathrm{R}^2 / \ \mathrm{q}(2,2) = (2 - 2)^2 = 0$
d) La forma cuadrática $\mathrm{q}(\mathrm{x}_1, \mathrm{x}_2) = -(\mathrm{x}_1 - \mathrm{x}_2)^2$ es S.D.N. pues, $-(\mathrm{x}_1 - \mathrm{x}_2)^2 \leq 0\ \ \forall \mathrm{x} \in \mathrm{R}^2$ y existe vectores $\mathrm{x} = (\mathrm{x}_1, \mathrm{x}_2) \in \mathrm{R}^2 / \ \mathrm{x}_1 = \mathrm{x}_2\ \ \mathrm{y}\ \ \mathrm{Q}(\mathrm{x}_1, \mathrm{x}_2) = 0$. Por ejemplo $\exists\ \mathrm{x} = (1, 1) \in \mathrm{R}^2 / \ \mathrm{q}(1,1) = -(1 - 1)^2 = 0$

e) La forma cuadrática $\mathrm{Q}(\mathrm{x}_1, \mathrm{x}_2) = \mathrm{x}_1^2 - \mathrm{x}_2^2$ es INDEFINIDA, pues se pueden encontrar vectores de R^2 en los que tome valores positivos y negativos. Por ejemplo, si $(2,3) \in \mathrm{R}^2 \Longrightarrow \mathrm{Q}(2,3) = 2^2 - 3^2 = -5 < 0$ y si $(3,2) \in \mathrm{R}^2 \Longrightarrow \mathrm{Q}(3,2) = 3^2 - 2^2 = 5 > 0$

f) La forma cuadrática $q(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2$ es S.D.P. Obsérvese que esta forma cuadrática se puede escribir como $q(x_1, x_2) = (\mathrm{x}_1 + \mathrm{x}_2)^2 \geq 0 \ \forall \ (x_1, x_2) = X \in R^2$ y existe el vector $X = (x_1, -x_2) \in R^2, x_1 = x_2$ tal que $q(x_1, -x_2) = 0$

g) La forma cuadrática $q(x_1, x_2) = -x_1^2 + 2x_1x_2 - x_2^2$ es S.D.N. Obsérvese que esta forma cuadrática se puede escribir como $Q(x_1, x_2) = -(\mathrm{x}_1 - \mathrm{x}_2)^2 \leq 0 \ \forall \ (x_1, x_2) = X \in R^2$ y existe el vector $X = (x_1, x_2) \in R^2, x_1 = x_2$ tal que $Q(x_1, x_2) = 0$

h) La forma cuadrática $\mathrm{Q}(\mathrm{x}_1, \mathrm{x}_2) = \mathrm{x}_1\mathrm{x}_2$ es INDEFINIDA, pues se pueden encontrar vectores de $X = (x_1, x_2) \in \mathrm{R}^2 / (\mathrm{x}_1 > 0, \mathrm{x}_2 < 0)$ o $(\mathrm{x}_1 > 0, \mathrm{x}_2 > 0)$ en los que tome valores positivos y negativos. Por ejemplo, si $(2, -3) \in \mathrm{R}^2 \Longrightarrow \mathrm{q}(2, -3) = 2(-3) = -6 < 0$ y si $(3,2) \in \mathrm{R}^2 \Longrightarrow \mathrm{q}(3,2) = (3)2 = 6 > 0$. Nótese que también se puede tomar, por ejemplo si $(-2,3) \in \mathrm{R}^2 \Longrightarrow \mathrm{q}(-2,3) = -2(3) = -6 < 0$ y si $(3,2) \in \mathrm{R}^2 \Longrightarrow \mathrm{q}(3,2) = (3)2 = 6 > 0$

En particular, cuando una forma cuadrática está escrita en su forma canónica

$$\boldsymbol{Q}(\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_n}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{d_i}\, \boldsymbol{x_i^2}, \qquad \boldsymbol{d_i} \in R, i = 1, 2, \cdots. n.$$

Puede ser clasificada según su signo de acuerdo al valor que toman los d_i. Esto se establece en la siguiente definición.

Definición. Sea la forma cuadrática escrita en su forma canónica

$$\boldsymbol{Q}(\boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_n}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{d_i}\, \boldsymbol{x_i^2}, \qquad d_i \in R, i = 1, 2, \cdots. n.$$

Se puede afirmar que.

1. $Q(\boldsymbol{X})$ es definida positiva (D.P.) si y solo si $d_i > 0,\ i = 1,2, \cdots, n$
2. $Q(\boldsymbol{X})$ es definida negativa (D.N.) si y solo si $d_i < 0,\ i = 1,2, \cdots, n$
3. $Q(\boldsymbol{X})$ es semidefinida positiva (S.D.P.) si y solo si $d_i \geq 0,\ i = 1,2, \cdots, n$ y existe al menos i_0 tal que $d_{i_0} = 0$

4. $Q(\boldsymbol{X})$ es semidefinida negativa (S.D.N.) si y solo si $d_i \leq 0,\ i = 1,2,\cdots,n$ y existe al menos i_1 tal que $d_{i_1} = 0$
5. $Q(\boldsymbol{X})$ es indefinida si $\exists\ \boldsymbol{d_1}, \boldsymbol{d_2}\ /\ d_{i_1} < 0\ y\ d_{i_2} > 0$

COMENTARIO. Una expresión diagonal o canónica de una forma cuadrática

$$Q(X) = X^T A X$$

Está dada por

$$Q(x_1, x_2, \cdots, x_n) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$$
$$= [x_1 \quad x_2 \quad \cdots \quad x_n] \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$Q(X) = X^T D X$$

Donde $D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & d_n \end{bmatrix}$ es una matriz diagonal, y $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$,

Es decir, la expresión polinómica solo contiene términos cuadráticos y la matriz asociada a la forma cuadrática es diagonal.

2.7. MÉTODOS PARA CLASIFICAR UNA FORMA CUADRÁTICA.

Vamos a estudiar otras caracterizaciones del signo de una forma cuadrática que vienen dadas, bien a través de los menores principales, o bien a través de lo autovalores de la matriz asociada a la forma cuadrática.

TEOREMA 2.7.1. (Método de Jacobi o método de los menores principales para determinar el signo de una forma cuadrática).

Sea $Q: R^n \longrightarrow R$ una forma cuadrática real de n variables $x_1, x_2, \cdots, x_n$ y sean $\Delta_1, \Delta_2, \cdots, \Delta_n$ los menores principales de la matriz A asociada a $Q(\boldsymbol{X}) = X^T AX$. Se verifica:

1. $Q(\boldsymbol{X})$ es D.P. si y sólo si $\Delta_1 > 0,\ \Delta_2 > 0, \cdots,\ \Delta_n > 0$.
2. $Q(\boldsymbol{X})$ es D.N. si y sólo si $\Delta_1 < 0,\ \Delta_2 > 0, \cdots, \Delta_n \begin{cases} > 0 & si\ n\ es\ par \\ < 0 & si\ n\ es\ impar \end{cases}$

   .
3. Si $\Delta_n \neq 0$ y no se verifica ni 1) ni 2) entonces $Q(\boldsymbol{X})$ es INDEFINIDA.

   Sea $r = rg(A)$: Rango de la matriz A
4. Si $\Delta_i \geq 0, \forall\, i = 1,2, \cdots, r$ entonces $Q(\boldsymbol{X})$ es S.D.P.
5. Si $\Delta_1 \leq 0,\ \Delta_2 \geq 0, \cdots, \Delta_r \begin{cases} \geq 0 & si\ r\ es\ par \\ \leq 0 & si\ r\ es\ impar \end{cases}$ entonces $Q(\boldsymbol{X})$ es S.D.N.
6. Si $\Delta_i \neq 0, \forall\, i = 1,2, \cdots, r$ y no se verifica ni 4) ni 5) entonces $Q(\boldsymbol{X})$ es INDEFINIDA

Procedimiento para calcular el rango de una matriz.

Para calcular el rango de una matriz se sigue los dos pasos siguientes

Paso 1: Utilizando operaciones elementales por renglones se transforma la matriz A en una matriz B en forma escalonada por renglones.

Paso 2. El rango de A es igual al número de renglones no nulos de B.

.Este método también se establece como sigue:

TEOREMA 2.7.2.. Sea $Q(\boldsymbol{X}) = \boldsymbol{X}^{\boldsymbol{T}} A\boldsymbol{X}$ una forma cuadrática en las variables $x_1, x_2, \cdots, x_n$, siendo A la matriz simétrica asociada a la forma cuadrática, y sean $\Delta_1, \Delta_2, \cdots, \Delta_n$ sus menores principales, entonces:

1. $Q(\boldsymbol{X})$ es D.P. $\Leftrightarrow$ $\Delta_i > 0, i = 1,2, \cdots, n$.
2. $Q(\boldsymbol{X})$ es D.N. $\Leftrightarrow$ $(-1)^i \Delta_i > 0, i = 1,2, \cdots, n$
3. $Q(\boldsymbol{X})$ es S.D.P. SI $\Delta_i \geq 0, i = 1,2, \cdots, n-1$ y $\Delta_n = |A| = 0$
4. $Q(\boldsymbol{X})$ es S.D.N. SI $(-1)^i \Delta_i \geq 0, i = 1,2, \cdots, n-1$ y $\Delta_n = |A| = 0$
5. $Q(\boldsymbol{X})$ es INDEFINIDA si no se cumple1), 2), 3) o 4)

COMENTARIO Los numerales 1) y 2) del teorema 2.7.1 y los numerales 1) y 2) del teorema 2.7.2. son equivalentes. Para 1) es evidente. Para el numeral 2), vemos que, si:

$$i = 1 \;\Rightarrow\; (-1)^1\Delta_i \geq 0 \;\Leftrightarrow\; -\Delta_i \leq 0 \;\Leftrightarrow\; \Delta_i \leq 0$$

$$i = 2 \;\Rightarrow\; (-1)^2\Delta_i \geq 0 \;\Leftrightarrow\; \Delta_i \geq 0$$

$$i = 3 \;\Rightarrow\; (-1)^3\Delta_i \geq 0 \;\Leftrightarrow\; -\Delta_i \geq 0 \;\Leftrightarrow\; \Delta_i \leq 0$$

$$i = 4 \;\Rightarrow\; (-1)^4\Delta_i \geq 0 \;\Leftrightarrow\; \Delta_i \geq 0$$

Continuando sucesivamente

Si i es par $\Rightarrow (-1)^i\Delta_i \geq 0$ y si i es impar $\Rightarrow$ $(-1)^i\Delta_i \leq 0$.

Por tanto: $Q(\boldsymbol{X})$ es D.N $\Leftrightarrow$ $(-1)^i\Delta_i \geq 0, i = 1,2,\cdots,n$ es EQUIVALENTE con $Q(\boldsymbol{X})$ es D.N. si y sólo si $\Delta_1 \leq 0$, $\Delta_2 \geq 0, \cdots, \Delta_n \begin{cases} \geq 0 \;\; si\; n\; es\; par \\ \leq 0 \; si\; n\; es\; impar \end{cases}$

Ejemplo 1.

a) Usando el signo de una forma cuadrática, determinar si una empresa tiene utilidad o pérdida al utilizar para sus ventas la función de beneficios dada por la forma cuadrática

$$B(x_1, x_2, x_3) = 3x_1^2 + x_2^2 + 2x_2x_3 - x_3^2$$

Solución.

En este caso la matriz de la forma cuadrática es.

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Los menores principales son:

$$\Delta_1 = 3 > 0;\;\; \Delta_2 = \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} = 3 > 0;\quad \Delta_3 = \begin{vmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{vmatrix} = 3\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}$$
$$= -6 < 0$$

Observamos que $\Delta_1 = 3 > 0;$ $\Delta_2 = 3 > 0$ y $\Delta_3 = -6 < 0$. Como $\Delta_3 = -6 \neq 0$ y no cumple ni 1) ni 2) del teorema 01, entonces la forma cuadrática es INDEFINIDA es decir, a veces es positiva y a veces es negativa, la empresa puede tener pérdidas.

b) Clasifique la forma cuadrática siguiente, usando el método de Jacobi.

$$Q(\boldsymbol{X}) = 3x_1^2 - 4x_1x_2 + 6x_2^2 + x_3^2$$

Solución

La matriz asociada a la forma cuadrática es

$$A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Los menores principales son:

$$\Delta_1 = 3 > 0; \;\; \Delta_2 = \begin{vmatrix} 3 & -2 \\ -2 & 6 \end{vmatrix} = 14 > 0; \quad \Delta_3 = \begin{vmatrix} 3 & -2 & 0 \\ -2 & 6 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= 1\begin{vmatrix} 3 & -2 \\ -2 & 6 \end{vmatrix} = 14 > 0$$

Observamos que $\Delta_1 = 3 > 0;\ \Delta_2 = 14 > 0\ \ y\, \Delta_3 = 14 > 0$, entonces la forma cuadrática es D.P

c) Clasificar la forma cuadrática de acuerdo a su signo.

$$Q(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + x_2^2 - 4x_3^2$$

Solución.

La matriz asociada a la forma cuadrática es

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

Los menores principales son:

$$\Delta_1 = 1 > 0; \;\; \Delta_2 = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0; \quad \Delta_3 = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -4 \end{vmatrix} = -4\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0$$

Por el teorema 1.6.2. parte 5), la forma cuadrática es INDEFINIDA porque no se cumple 1), 2), 3) o 4).

Podemos resolver el problema, usando el teorema 1.6.1., calculando el rango de la matriz A.

$$A = \begin{matrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -4 \end{matrix} \xrightarrow[-R_1 + R_2]{} \begin{matrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -4 \end{matrix} \xrightarrow[R_{23}]{} \begin{matrix} 1 & 1 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{matrix} \xrightarrow[-\frac{1}{4}R_2]{} \begin{matrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix} = B$$

Como hay dos renglones diferentes de cero en la matriz B, entonces el rango es $r = 2$. Entonces $\Delta_1 = 1 > 0$; $\Delta_2 = 0$

Obsérvese que no es posible aplicar el teorema 1.6.1 porque no se cumple ninguno de los 6 puntos o sea no es D.P, D.N, S.D.P, S.D.N o INDEFINIDA, entonces no es posible clasificar la forma cuadrática según su signo.

d) Sea $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ la matriz asociada a la forma cuadrática $Q(\boldsymbol{X}) = \boldsymbol{X}^T A\boldsymbol{X}$. Clasificar la forma cuadrática según su signo usando el método de Jacobi

Solución.

Los menores principales son:

$$\Delta_1 = 1 > 0;\ \ \Delta_2 = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0;\quad \Delta_3 = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 4 \end{vmatrix} = 0$$

Obsérvese que, según el método de Jacobi, la forma cuadrática no es: D.P, D.N. S.D.P, S.D.N ni indefinida.

COROLARIO 2.7.3. . (Expresión diagonal o canónica de Jacobi)

Sea una forma cuadrática $Q(\boldsymbol{X}) = \boldsymbol{X}^T A\boldsymbol{X}$ con su matriz asociada A, sean $\Delta_1, \Delta_2, \cdots, \Delta_n$ los menores principales de A y sea $\boldsymbol{rg}(\boldsymbol{A}) = \boldsymbol{r} < n$. La expresión diagonal de Jacobi de la forma cuadrática q viene dada por:

$$Q(x_1, x_2, \cdots, x_n) = \Delta_1 x_1^2 + \frac{\Delta_2}{\Delta_1} x_2^2 + \frac{\Delta_3}{\Delta_2} x_3^2 + \cdots + \frac{\Delta_r}{\Delta_{r-1}} x_r^2,$$

siempre que $\Delta_1 \neq 0,\ \Delta_2 \neq 0,\ \cdots, \Delta_r \neq 0$

Ejemplo1.

Sea la forma cuadrática $Q(x_1, x_2, x_3) = 3x_1^2 + 3x_2^2 + 5x_3^2 - 4x_1x_2$ expresarlo en la forma diagonal de Jacobi.

Solución.

La matriz asociada a la forma cuadrática es

$$A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Los menores principales son $\Delta_1 = 3,\ \Delta_2 = 5,\ \Delta_3 = 25$

Veamos el rango de la matriz. Aplicamos transformaciones elementales sobre filas en la matriz A de la cual se ha tomado solo los elementos por facilidad

$$A = \begin{matrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{matrix} \xrightarrow{\frac{1}{3}R_1} \begin{matrix} 1 & -2/3 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{matrix} \xrightarrow{2R_1 + R_2} \begin{matrix} 1 & -2/3 & 0 \\ 0 & 5/3 & 0 \\ 0 & 0 & 5 \end{matrix}$$

$$\xrightarrow{\frac{3}{5}R_2} \begin{matrix} 1 & -2/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{matrix} \xrightarrow{\frac{2}{3}R_2 + R_1} \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{matrix} \xrightarrow{\frac{1}{5}R_3} \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} = B$$

Observamos que la matriz B equivalente a la matriz A tiene tres renglones (vectores) diferentes de cero, entonces el rango de la matriz A es $r = 3$

Como el rango de la matriz A es $r = 3$ y los tres menores principales son diferentes de cero, la expresión diagonal de Jacobi es

$$Q(x_1, x_2, x_n) = 3x_1^2 + \frac{5}{3}x_2^2 + \frac{25}{5}x_3^2 = 3x_1^2 + \frac{5}{3}x_2^2 + 5x_3^2$$

Los ejercicios c) y d) del ejemplo 1 que siguen al teorema 2.7.2, hacen ver que es necesario establecer un nuevo método para poder clasificar las formas cuadráticas que no se pueden clasificar por el método de Jacobi.
,n

TEOREMA 1.6.2. (Método de los autovalores para determinar el signo de una forma cuadrática) Sea $Q: R^n \longrightarrow R$ una forma cuadrática real de n variables y sean $\lambda_1, \lambda_2 . \cdots, \lambda_n$ los autovalores de la matriz simétrica A asociada a $Q(\boldsymbol{X}) = \boldsymbol{X^T A X}$. Se verifica:

1. $Q(\boldsymbol{X})$ es D.P. si y sólo si los autovalores de A son todos positivos; esto es

   $Q(\boldsymbol{X})$ es D.P $\Leftrightarrow \lambda_i > 0 \ \forall i = 1, 2, 3, \cdots, n$
2. $Q(\boldsymbol{X})$ es D.N. si y sólo si los autovalores de A son todos negativos: esto es

   $Q(\boldsymbol{X})$ es D.N $\Leftrightarrow \lambda_i < 0 \ \forall i = 1, 2, 3, \cdots, n$
3. $Q(\boldsymbol{X})$ es S.D.P. si y sólo si los autovalores de A son positivos y nulos.
4. $Q(\boldsymbol{X})$ es S.D.N. si y sólo si los autovalores de A son todos negativos y nulos.

5. $Q(\boldsymbol{X})$ es INDEFINIDA si y sólo si los autovalores de A son positivos y negativos.

COMENTARIO . Para caracterizar una forma cuadrática de acuerdo a su signo, usando el teorema 2.7.3., es necesario calcular la matriz simétrica asociada a la forma cuadrática y luego calcular los autovalores de esta matriz. La forma cuadrática en su forma matricial es $\boldsymbol{X}^T A\,\boldsymbol{X}$. Los autovalores se obtienen resolviendo $|A-\lambda I|=0$

Ejemplo 2.

a) Usando el método de los autovalores, clasificar de acuerdo a su signo la forma cuadrática $Q(x_1,x_2,x_3)=x_1^2+2x_1x_2+x_2^2-4x_3^2$

Solución.

Escribiendo la forma cuadrática en su forma equivalente, es decir en forma matricial

$$Q(x_1,x_2,x_3)=[x_1,x_2,x_3]\begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ 0 & 0 & -4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

Ahora se calculan los autovalores de la matriz simétrica $\boldsymbol{A}$ asociada a la forma cuadrática, usando $|\mathrm{A}-\lambda\mathrm{I}|=0$

$$|A-\lambda I|=\begin{vmatrix}1-\lambda & 1 & 0\\ 1 & 1-\lambda & 0\\ 0 & 0 & -4-\lambda\end{vmatrix}=-(4+\lambda)[(1-\lambda)^2-1]$$
$$=-(4+\lambda)\lambda(\lambda-2)=0$$

De donde $\lambda_1=-4;\ \lambda_2=0;\ \lambda_3=2$

Puesto que los autovalores son positivos y negativos, seguimos que la forma cuadrática es INDEFINIDA.

b) Usando el signo de una forma cuadrática, determinar si una empresa tiene utilidad o pérdida al utilizar para sus ventas la función de beneficios dada por la forma cuadrática. Resuelva por el método de los autovalores.

$$B(x_1,x_2,x_3)=3x_1^2+x_2^2+2x_2x_3-x_3^2$$

Solución.

Escribiendo la forma cuadrática en su forma equivalente, es decir en forma matricial

$$B(x_1, x_2, x_3) = [x_1\ x_2\ x_3]\begin{bmatrix}3 & 0 & 0\\ 0 & 1 & 1\\ 0 & 1 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

Ahora se calculan los autovalores de la matriz simétrica $\boldsymbol{A}$ asociada a la forma cuadrática, usando $|\mathrm{A} - \lambda\mathrm{I}| = 0$

$$|A - \lambda I| = \begin{vmatrix}3-\lambda & 0 & 0\\ 0 & 1-\lambda & 1\\ 0 & 1 & -1-\lambda\end{vmatrix} = (3-\lambda)[-(1-\lambda)(1+\lambda) - 1]$$
$$= -(3-\lambda)(2-\lambda^2) = 0$$

De donde $\lambda_1 = -\sqrt{2}$; ; $\lambda_3 = \sqrt{2}$; $\lambda_2 = 3$

Luego, la forma cuadrática es INDEFINIDA, es decir, a veces es positiva y a veces negativa, la empresa puede tener pérdidas.

COROLARIO 1.6.2.. (Expresión diagonal o canónica por autovalores)

Dada la forma cuadrática $Q(\boldsymbol{X}) = \boldsymbol{X}^T A \boldsymbol{X}$ con su matriz asociada A cuyos autovalores son $\lambda_1, \lambda_2, \cdots, \lambda_n$, existe una expresión diagonal o canónica para $q(X)$ dada por

$$Q(x_1, x_2, \cdots, x_n) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + \cdots + \lambda_n x_n^2,$$

Ejemplo1.

Sea la forma cuadrática $Q(x_1, x_2, x_3) = 3x_1^2 + 3x_2^2 + 5x_3^2 - 4x_1x_2$ expresarlo en la forma diagonal por autovalores.

Solución.

La matriz asociada a la forma cuadrática es

$$A = \begin{bmatrix}3 & -2 & 0\\ -2 & 3 & 0\\ 0 & 0 & 5\end{bmatrix}$$

Los autovalores se determinan resolviendo la ecuación $|A - \lambda I| = 0$, esto es:

$$\begin{vmatrix}3-\lambda & -2 & 0\\ -2 & 3-\lambda & 0\\ 0 & 0 & 5-\lambda\end{vmatrix} = 0$$

$$0 = (5-\lambda)\begin{vmatrix} 3-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} = (5-\lambda)(9-6\lambda+\lambda^2-4)$$
$$= (5-\lambda)(\lambda^2-6\lambda+5)$$

$$(5-\lambda)(\lambda-5)(\lambda-1) = 0$$

$$\lambda_1 = 1, \qquad \lambda_2 = \lambda_3 = 5$$

Por tanto, la expresión diagonal o canónica por autovalores es

$$Q(x_1, x_2, x_3) = x_1^2 + 5x_2^2 + 5x_3^2$$

Escrita en forma matricial es

$$Q(x_1, x_2, x_3) = Q(X) = X^T AX = (x_1, x_2, x_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

3. CONJUNTOS CONVEXOS.

INTRODUCCIÓN. Esta sección tiene por objetivo introducir el concepto de conjuntos convexos, dar propiedades de tales conjuntos y derivar los teoremas básicos de separación para conjuntos convexos. Estos teoremas de separación forman la base fundamental en la que se apoyan las condiciones de optimización de programación no lineal.

3.1. CONJUNTOS CONVEXOS EN $\boldsymbol{R^n}$

3.1.1. CONJUNTOS CONVEXOS Y SUS PROPIEDADES.

Con la finalidad de introducir el concepto de conjunto convexo, empezaremos definiendo la recta y el segmento de recta entre dos puntos de R^n.

3.1.2. RECTA.

Sean $\boldsymbol{x_1}, \boldsymbol{x_2} \in R^n$, la recta que pasa a través de los puntos $\boldsymbol{x_1}\ y\ \boldsymbol{x_2}$ se define como el conjunto

$$\{\boldsymbol{x} \in R^n : \boldsymbol{x} = (1-\lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}, \lambda \in R\}$$

o equivalentemente

$$\{\boldsymbol{x} \in R^n : x = p_1\boldsymbol{x_1} + p_2\boldsymbol{x_2}\ para\ \ p_1 + p_2 = 1, p_1, p_2 \in R\}$$

Reescribiendo la primera definición en la forma equivalente

$$\{\boldsymbol{x} \in R^n : \boldsymbol{x} = \boldsymbol{x_1} + \lambda(\boldsymbol{x_2} - \boldsymbol{x_1}), \lambda \in R\}$$

y considerando el caso cuando $\boldsymbol{x} \in R$, es obvio que la ecuación vectorial

$$\{\boldsymbol{x} \in R^n : \boldsymbol{x} = (1 - \lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}, \lambda \in R\}$$

es la ecuación paramétrica de la recta que pasa a través de los puntos $\boldsymbol{x_1}$ y $\boldsymbol{x_2}$ (Ver figura 1.1.2)

3.1.3. SEGMENTOS DE RECTA.

Sean $\boldsymbol{x_1}\ y\ \boldsymbol{x_2} \in R^n$. Definimos los siguientes segmentos de recta que une los puntos $\boldsymbol{x_1}\ y\ \boldsymbol{x_2}$.

a) Segmento de recta cerrado.

$$[\boldsymbol{x_1}, \boldsymbol{x_2}] = \{\boldsymbol{x} \,/\, \boldsymbol{x} = (1 - \lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}, \qquad 0 \le \lambda \le 1\}$$

El segmento incluye los puntos $x_1\ y\ x_2$

b) Segmento de recta abierto.

$$(\boldsymbol{x_1}, \boldsymbol{x_2}) = \{\boldsymbol{x} \,/\, \boldsymbol{x} = (1 - \lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}, \qquad 0 < \lambda < 1\}$$

El segmento no incluye los puntos $x_1\ y\ x_2$

c) Segmento de recta cerrado-abierto.

$$[\boldsymbol{x_1}, \boldsymbol{x_2}) = \{\boldsymbol{x} \,/\, \boldsymbol{x} = (1 - \lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}, \qquad 0 \le \lambda < 1\}$$

El segmento incluye el punto x_1 y excluye el punto x_2.

d) Segmento de recta abierto-cerrado.

$$(\boldsymbol{x_1}, \boldsymbol{x_2}] = \{\boldsymbol{x} \,/\, \boldsymbol{x} = (1 - \lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}, \qquad 0 < \lambda \le 1\}$$

El segmento no excluye el punto x_1 pero incluye el punto x_2.

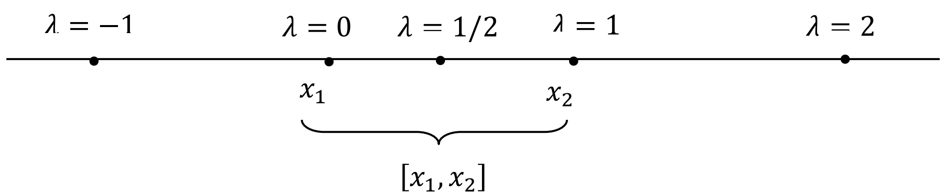


Fig. 3.1.1: Recta y segmento de recta a través de $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$

3.1.4. DEFINICIÓN DE CONJUNTO CONVEXO.

Los contenidos que se presentan a continuación han sido extraídos del libro NON LINEAR PROGRAMMING del autor MNGASARIAN. Olvi L

Definición (conjunto convexo). Un subconjunto de S de R^n es un conjunto convexo si, para dos puntos cualesquiera $\boldsymbol{x_1}\ y\ \boldsymbol{x_2} \in S$ y cualquier λ que satisface $0 \leq \lambda \leq 1, \lambda \in R^n$, entonces el punto $\boldsymbol{x} = \big((1-\lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}\big) \in S$

Simbólicamente:

S es convexo si

$$\boldsymbol{x_1}\ y\ \boldsymbol{x_2} \in S\ \wedge\ 0 \leq \lambda \leq 1, \lambda \in R\ \Rightarrow \boldsymbol{x} = \big((1-\lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}\big) \in S$$

.

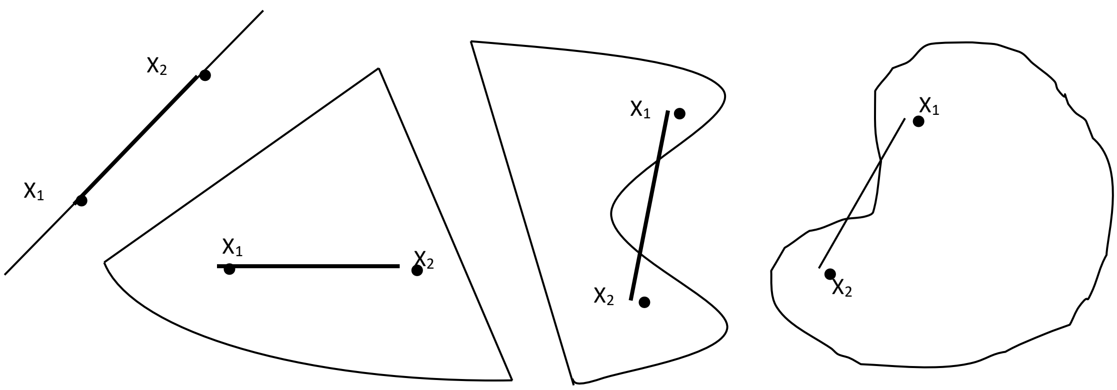


Fig.3.1.4: a) Conjuntos convexos en R^2. b) Conjuntos no convexos en R^2

Intuitivamente el conjunto S es convexo cuando el segmento

$$\boldsymbol{x} = \big((1-\lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}\big)$$

está contenido completamente en el conjunto.

De la definición de conjunto convexo seguimos que los siguientes conjuntos también son convexos.

a) El espacio R^n
b) El conjunto vacío

3.1.5. SEMIESPACIO.

Sea $\boldsymbol{c} \in R^n, \boldsymbol{c} \neq 0$ y $\alpha \in R$. Entonces el conjunto

$$[\, \boldsymbol{x} \in \boldsymbol{R^n} \,/\;\; \boldsymbol{cx} < \alpha]$$

es un semi espacio abierto en R^n y el conjunto

$$[\, \boldsymbol{x} \in \boldsymbol{R^n} \,/\;\; \boldsymbol{cx} \leq \alpha]$$

es un semi espacio cerrado en R^n.

Ambos semi espacio son conjuntos convexos.

3.1.6. PLANO.

Sea $\boldsymbol{c} \in R^n, c \neq 0$ y $\alpha \in R$. Entonces el conjunto

$$[\, \boldsymbol{x} \in \boldsymbol{R^n} \,/\;\; \boldsymbol{cx} = \alpha]$$

se llama plano en R^n.

3.1.7. SUBESPACIO.

Un conjunto $S \subset R^n$ es un subespacio si

$$x_1, x_2 \in S;\; p_1, p_2 \in R \Longrightarrow (p_1 x_1 + p_2 x_2) \in S.$$

El subespacio de R^n que contiene el origen es un conjunto convexo.

Los subespacio de R^3 son el $\emptyset, R^2$, el origen y todas las rectas y planos que pasan por el origen.

3.1.8. VÉRTICE.

Sea S un conjunto convexo en R^n. Cada $\boldsymbol{x} \in S$ para el cual no existen dos puntos distintos $\mathbf{x_1},\; \boldsymbol{x_2} \in S$ diferentes de $\boldsymbol{x}$ tales que $\boldsymbol{x} \in [\boldsymbol{x_1}, \boldsymbol{x_2}]$, se llama un vértice de S (o un punto extremo de S)

OBSERVACIÓN:

Un conjunto convexo $S \subset R^n$ puede:

a) No tener vértices.

Por ejemplo: el plano $H = [\, \boldsymbol{x} \in \boldsymbol{R^n} \,/\;\; \boldsymbol{cx} = \alpha]$ y la bola abierta $B(\boldsymbol{x_0}, \rho) = \{\, \boldsymbol{x} \in R^n \;/\; \|\boldsymbol{x} - \boldsymbol{x_0}\| < \rho\}$ no tienen vértices.

b) Un número finito de vértices.

Por ejemplo el conjunto $T = \{ \boldsymbol{x} \in R^n \ / \ \boldsymbol{x} \geq 0, \ \boldsymbol{ex} = 1\}$, donde $\boldsymbol{e} \in R^n$ es un vector de unos, tiene los n vértices $\boldsymbol{e^i}, i = 1, 2, \cdots, n$, donde $\boldsymbol{e^i}$ es un vector n dimensional con $\boldsymbol{e_i^i} = 1$ y $\boldsymbol{e_j^i} = 0$, para $i \neq j$

c) Un número infinito de vértices.

Por ejemplo la bola cerrada $\bar{B}(\boldsymbol{x_0}, \rho) = \{ \boldsymbol{x} \in R^n \ / \ \|\boldsymbol{x} - \boldsymbol{x_0}\| \leq \rho\}$ tiene un número infinito de vértices dados por $\{\boldsymbol{x} \in R^n \ / \ \|\boldsymbol{x} - \overline{\boldsymbol{x}}\| = \rho\}$

3.2. TEOREMA.

Si $(S_i)_{i \in I}$ es una familia (finita o infinita) de conjuntos convexos en R^n, entonces su intersección $\bigcap_{i \in I} S_i$ es un conjunto convexo.

DEMOSTRACIÓN.

Sean $x_1, x_2 \in \bigcap_{i \in I} S_i$ y sea $0 \leq \lambda \leq 1$. Entonces $\forall \ i \in I, x_1, x_2 \in S_i$ y como S_i es convexo, se tiene que $\boldsymbol{x} = (1 - \lambda)\boldsymbol{x_1} + \lambda \boldsymbol{x_2} \in S_i$, es decir $\bigcap_{i \in I} S_i$ es convexo.

3.3. POLÍTOPO Y POLIEDRO.

Un conjunto en R^n el cuál es la intersección de un número finito de semi espacio cerrados en R^n es llamado un polítopo. Si un polítopo es acotado (esto es, para cada $\boldsymbol{x}$ en el politopo $\|\boldsymbol{x}\| \leq \alpha$ para algún $\alpha \in R$ fijo) es llamado un poliedro.

De la convexidad de los semi espacio y el teorema , seguimos que los politopos y poliedros son conjuntos convexos.

3.4. COMBINACIÓN CONVEXA.

Un punto $\boldsymbol{b} \in R^n$ se dice que es una combinación convexa de los vectores $\boldsymbol{a_1}, \boldsymbol{a_2}, \cdots, \boldsymbol{a_m} \in R^n$ si existen m números reales $p_1, p_2, \cdots, p_m$ tal que

$$\boldsymbol{b} = p_1\boldsymbol{a_1} + p_2\boldsymbol{a_2} + \cdots + p_m\boldsymbol{a_m}; \ \ p_1, p_2, \cdots, p_m \geq 0 \colon p_1 + p_2 + \cdots + p_m = 1$$

Equivalentemente, si definimos una matriz $A \ m \times n$ cuyo $i-$ésimo renglón es $A_i = a_i$, y si hacemos $\boldsymbol{p} = (p_1, p_2, \cdots, p_m) \in R^n$ y $\boldsymbol{e}$ un $m-$ vector de unos, entonces tenemos que $\boldsymbol{b}$ es una combinación convexa de las filas de A si el sistema

$$\boldsymbol{b} = A'p, \qquad \boldsymbol{p} \geq 0$$

$$\boldsymbol{ep} = 1$$

tiene una solución $\boldsymbol{p} \in R^m$.

3.5. SIMPLEX.

Sean los $m+1$ puntos distintos $\boldsymbol{x_0}, \boldsymbol{x_1}, \cdots, \boldsymbol{x_m}$ en R^n, con $m \leq n$. Si los vectores $\boldsymbol{x_1} - \boldsymbol{x_0}, \boldsymbol{x_2} - \boldsymbol{x_0}, \cdots, \boldsymbol{x_m} - \boldsymbol{x_0}$ son linealmente independientes, entonces el conjunto de todas las combinaciones convexas de $\boldsymbol{x_0}, \boldsymbol{x_1}, \cdots, \boldsymbol{x_m}$

$$T = \left\{ z / \, z = \sum_{i=0}^{m} p_i x_i, \qquad p_i \in R, \quad p_i \geq 0, \qquad i = 0, 1, \cdots, m, \right.$$
$$\left. \sum_{i=0}^{m} p_i = 1 \right\}$$

es llamado un $\boldsymbol{m} -$ simplex en R^n con vértices $\boldsymbol{x_0}, \boldsymbol{x_1}, \cdots, \boldsymbol{x_m}$

OBSERVACIÓN.

a) Un $\boldsymbol{0} -$ simplex es un punto.

b) Un $\boldsymbol{1} -$ simplex es un segmento de recta cerrado.

c) Un $\boldsymbol{2} -$ simplex es un triangulo.

d) Un $\boldsymbol{3} -$ simplex es un tetraedro.

3.6. TEOREMA.

Un conjunto $S \subset R^n$ es convexo si y solamente si para cada entero $m \geq 1$, toda combinación convexa de m puntos cualesquiera de S están en S.

Equivalentemente, una condición necesaria y suficiente para que el conjunto S sea convexo es que para entero $m \geq 1$

$$\left. \begin{array}{r} \boldsymbol{x_1}, \cdots, \boldsymbol{x_m} \in S \\ p_1, p_2, \cdots, p_m \geq 0 \\ p_1 + p_2 + \cdots + p_m = 1 \end{array} \right\} \Longrightarrow p_1 \boldsymbol{x_1} + p_2 \boldsymbol{x_2} + \cdots + p_m \boldsymbol{x_m} \in S \quad \ldots\ldots\ldots\ldots.. (1)$$

DEMOSTRACIÓN.

La condición suficiente de (1) es trivial; tomar $m = 2$, entonces por definición de conjunto convexo, se sigue que S es convexo.

La condición necesaria de (1) será demostrada por inducción. Para $m = 1$, se sigue que (1) vale trivialmente. Para $m = 2$, se sigue que (1) es válido como consecuencia de la definición de conjunto convexo. Asumiendo ahora que (1) es válido para m, mostraremos que también es válido para $m + 1$.

Sea

$$\begin{aligned} \boldsymbol{x_1}, \boldsymbol{x_2}, \cdots, \boldsymbol{x_m}, \boldsymbol{x_{m+1}} &\in S \\ p_1, p_2, \cdots, p_m, p_{m+1} &\geq 0 \\ p_1 + p_2 + \cdots + p_m + p_{m+1} &= 1 \end{aligned}$$

Si $p_{m+1} = 0$, entonces $p_1\boldsymbol{x_1} + p_2\boldsymbol{x_2} + \cdots + p_m\boldsymbol{x_m} \in S$, porque (1) vale para todo m.

Si $p_{m+1} = 1$, entonces $p_1\boldsymbol{x_1} + p_2\boldsymbol{x_2} + \cdots + p_m\boldsymbol{x_m} + p_{m+1}\boldsymbol{x_{m+1}} = \boldsymbol{x_{m+1}} \in S$

Si $0 < p_{m+1} < 1$, podemos escribir

$$\sum_{i=1}^{m+1} p_i x_i = \underbrace{\left[\sum_{i=1}^{m} p_i\right] \underbrace{\left[\frac{p_1}{\sum_{i=1}^{m} p_i} x_1 + \cdots + \frac{p_m}{\sum_{i=1}^{m} p_i} x_m\right]}_{\substack{Un\ punto\ en\ S, porque \\ (1) vale\ para\ m}} + p_m x_{m+1}}_{Punto\ en\ S,\ porque\ (1) vale\ para\ m=2} \in S \ \blacksquare$$

3.7. TEOREMA DE CARATHÉODORY.

Sea $S \subset R^n$. Si x es una combinación convexa de puntos de S, entonces x es una combinación convexa de $n + 1$ o algunos puntos de S.

DEMOSTRACIÓN.

Sea

$$\boldsymbol{x} = \sum_{i=1}^{m} p_i \boldsymbol{x_i}, \qquad \boldsymbol{x_i} \in S, \qquad p_i \in R, \quad p_i \geq 0,$$

$$p_1 + p_2 + \cdots + p_m = 1$$

Vamos a demostrar que si $m > n + 1$, entonces $\boldsymbol{x}$ se puede escribir como una combinación convexa de $m - 1$ puntos en S. (Esto establecerá el teorema, entonces podremos aplicar el teorema repetidamente hasta que $\boldsymbol{x}$ sea una combinación convexa de $m + 1$ puntos en S.) Si cualquier p_i de la expresión previa es cero, entonces $\boldsymbol{x}$ es una combinación convexa de $m - 1$ o algunos puntos de S. Así que cada $p_i > 0$. Dado que $m > n + 1$, entonces existen $r_1, r_2, \cdots, r_{m-1} \in R$, no todos ceros, tales que

$$r_1(\boldsymbol{x_1} - x_m) + r_2(\boldsymbol{x_2} - x_m) + \cdots + r_{m-1}(x_{m-1} - \boldsymbol{x_m}) = 0$$

Definiendo:

$$r_m = -(r_1 + r_2 + \cdots + r_{m-1})$$

Entonces

$$\sum_{i=1}^{m} r_i = 0 \qquad\qquad \sum_{i=1}^{m} r_i x_i = 0$$

Definiendo

$$q_i = p_i - \alpha r_i \quad para \;\; i = 1, 2, \cdots, m$$

donde α es algún número positivo seleccionado de tal manera que $q_i \geq 0$ para todo i, y al menos algún q_i, digamos q_k es igual a cero. En particular seleccionamos α tal que

$$\frac{1}{\alpha} = \max_i \left\{\frac{r_i}{p_i}\right\} = \frac{r_k}{p_k}$$

Entonces

$$q_i \geq 0, \;\; i = 1, 2, \cdots, m, q_k = 0$$

$$\sum_{\substack{i=1 \\ i\neq k}}^{m} q_i = \sum_{i=1}^{m} q_i = \sum_{i=1}^{m} p_i - \alpha \sum_{i=1}^{m} r_i = \sum_{i=1}^{m} p_i = 1$$

y

$$\boldsymbol{x} = \sum_{i=1}^{m} p_i\, x_i = \sum_{i=1}^{m} q_i x_i + \alpha \sum_{i=1}^{m} r_i\, x_i = \sum_{\substack{i=1 \\ i\neq k}}^{m} q_i\, x_i$$

En consecuencia $\boldsymbol{x}$ es una combinación convexa de $m - 1$ puntos de S

3.8. CASCO CONVEXO

Sea $S \subset R^n$. El casco convexo de S, denotado por $[S]$, es la intersección de todos los conjuntos convexos de R^n que contiene a S. El casco convexo también se llama envoltura convexa o envolvente convexa.

OBSERVACIÓN.

a) Por el teorema 1.1.8, el casco convexo de cualquier conjunto $S \subset R^n$ es convexo.

b) Obviamente si S es convexo, entonces $S = [S]$

3.9. TEOREMA.

El casco convexo $[S]$ de un conjunto $S \subset R^n$ es igual al conjunto de todas las combinaciones convexas de puntos de S.

DEMOSTRACIÓN.

Sea el conjunto

$$A = \left\{ \boldsymbol{x}/\boldsymbol{x} = \sum_{i=1}^{k} p_i a_i, \quad p_i \in R, \ \ a_i \in S \quad p_i \geq 0, \ \ \sum_{i=1}^{k} p_i = 1, \ k \geq 1 \right\}$$

Si $\boldsymbol{x_1}, \boldsymbol{x_2} \in A$, entonces

$$\boldsymbol{x_1} = \sum_{i=1}^{k} p_i a_i, \quad p_i \in R, \ \ a_i \in S, \quad p_i \geq 0, \quad \sum_{i=1}^{k} p_i = 1$$

$$\boldsymbol{x_2} = \sum_{i=1}^{m} q_i b_i, \quad q_i \in R, \ \ b_i \in S, \quad q_i \geq 0, \quad \sum_{i=1}^{m} q_i = 1$$

Por consiguiente para $0 \leq \lambda \leq 1$

$$\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2} = \sum_{i=1}^{k} \lambda\, p_i a_i + \sum_{i=1}^{m} (1-\lambda)\, q_i b_i$$

y

$$\lambda p_i \geq 0, (1-\lambda) q_i \geq 0, \qquad \sum_{i=1}^{k} \lambda p_i + \sum_{i=1}^{m} (1-\lambda) q_i = 1$$

Así $\lambda \boldsymbol{x_1} + (1-\lambda)\boldsymbol{x_2} \in A$, y A es convexo. También es claro que $S \subset A$.

Puesto que A es convexo, entonces $[S] \subset A$. Por el teorema 1.1.12 tenemos que el conjunto convexo $[S]$ que contiene a S también debe contener todas las combinaciones convexas de puntos de S. En consecuencia $A \subset [S]$, y $A = [S]$.

3.10. SUMA DE DOS CONJUNTOS.

Sean $S, A \subset R^n$. La suma de se define por

$$S + A = \{\boldsymbol{z}/\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}, \quad \boldsymbol{x} \in S, \quad \boldsymbol{y} \in A\}$$

3.11. PRODUCTO DE UN NÚMERO REAL POR UN CONJUNTO.

Sea $S \subset R^n$, y sea $\lambda \in R$. El producto λS se define por

$$\lambda S = \{\boldsymbol{z}/\boldsymbol{z} = \lambda \boldsymbol{x}, \boldsymbol{x} \in S\}$$

Nótese que si $\lambda = -1$ y $S \subset A \subset R^n$, entonces $A + \lambda S = A - S$.

Hágase notar que esto no es el complemento de S relativo a A

3.12. TEOREMA.

La suma $S + A$ de dos conjuntos convexos S y A en R^n es un conjunto convexo.

DEMOSTRACIÓN.

Sean $z_1, z_2 \in S + A$, entonces $z_1 = x_1 + y_1$ y $z_2 = x_2 + y_2$, donde $x_1, x_2 \in S$ y $y_1, y_2 \in A$.

Para $0 \leq \lambda \leq 1$:

$$(1-\lambda)z_1 + \lambda z_2 = \underbrace{(1-\lambda)x_1 + \lambda x_2}_{\substack{Un\ punto\ es\ S,\ por \\ convexidad\ de\ S}} + \underbrace{(1-\lambda)y_1 + \lambda y_2}_{\substack{Un\ punto\ es\ A,\ por \\ convexidad\ de\ S}} \in S + A$$

Por lo tanto $S + A$ es convexo.

3.13. TEOREMA. El producto λS de un conjunto convexo S en R^n y el número real λ es un conjunto convexo.

DEMOSTRACIÓN.

Sean $z_1, z_2 \in \lambda S$, entonces $z_1 = \lambda x_1$ y $z_2 = \lambda x_2$, donde $x_1, x_2 \in S$.

Para $0 \leq \lambda \leq 1$:

$$(1-\lambda)z_1 + \lambda z_2 = \lambda[\underbrace{(1-\lambda)x_1 + \lambda x_2}_{\substack{Un\ punto\ es\ S,\ por \\ convexidad\ de\ S}}] \in \lambda S$$

Por lo tanto λS es convexo.

3.13.1. COROLARIO.

Si S y A son dos conjuntos convexos en R^n, entonces $S - A$ es un conjunto convexo.

3.14. TEOREMAS DE SEPARACIÓN PARA CONJUNTOS CONVEXOS.

Es intuitivamente posible que si tenemos dos conjuntos convexos disjuntos en R^n, entonces podemos construir un plano de tal manera que un conjunto estaría a un lado y el otro conjunto al otro lado. A pesar de su simplicidad, este es un resultado un poco oscuro y no es fácil probarlo. Una versión de este resultado, el teorema de HAHN - BANACH, se puede establecer usando solamente las propiedades de un espacio vectorial en R^n sin las propiedades topológicas inducidas por la norma $\|\boldsymbol{x}\|$. Sin embargo, usaremos estas propiedades topológicas de R^n para derivar los teoremas de separación para conjuntos convexos. En particular nuestro método de prueba se hará usando el teorema de la alternativa de GORDAN y el teorema de intersección de conjuntos compactos.

3.14.1. PLANO DE SEPARACIÓN.

El plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \,/\, \boldsymbol{cx} = \alpha, \boldsymbol{c} \neq \boldsymbol{0}]$, se dice que separa (separa estrictamente) dos conjuntos S y A en R^n si

$$\boldsymbol{x} \in S \Rightarrow \boldsymbol{cx} \leq \alpha \qquad (\boldsymbol{cx} < \alpha)$$

$$\boldsymbol{x} \in A \Rightarrow \boldsymbol{cx} \geq \alpha \qquad (cx > \alpha)$$

Si tal plano existe, se dice que los conjuntos S y A son separables (estrictamente separables)

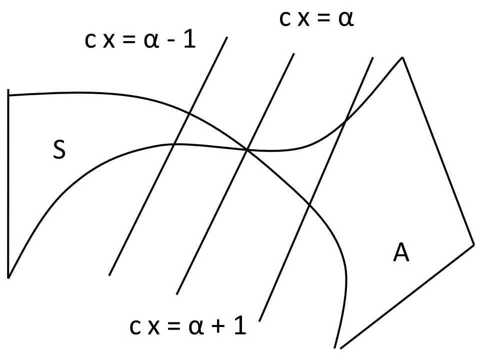


Fig. 3.14.1: a) Conjuntos separables pero no disjuntos

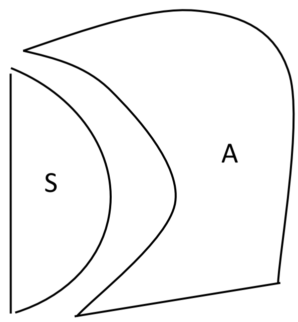


Fig. 3.14.1: b) Conjuntos disjuntos pero no separables

3.14.2. LEMA.

Sea $\mathbf{\Omega}$ un conjunto convexo no vacío en R^n que no contiene el origen 0. Entonces existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \ / \ \boldsymbol{cx} = \alpha, \boldsymbol{c} \neq \mathbf{0}]$ que separa $\mathbf{\Omega}$ y 0, tal que

$$\boldsymbol{x} \in \Omega \implies \mathbf{cx} \geq 0$$

DEMOSTRACIÓN.

Con cada $\boldsymbol{x} \in \mathbf{\Omega}$ asociamos el conjunto cerrado no vacío

$$A_x = \{\boldsymbol{y} \in R^n \ / \ \boldsymbol{yy} = 1, \boldsymbol{xy} \geq 0\}$$

Sea $x_1, x_2, \cdots, x_m$ cualquier conjunto finito de puntos en $\mathbf{\Omega}$. Por definición de convexidad de $\mathbf{\Omega}$, del teorema 3.12 y el hecho de que $\mathbf{0} \notin \mathbf{\Omega}$, seguimos que

$\sum_{i=1}^{m} x_i p_i = 0, \sum_{i=1}^{m} p_i = 1, p_i \geq 0, i = 1, 2, \cdots, m,$

no tiene solución $p \in R^m$. ó equivalentemente

$\sum_{i=1}^{m} x_i p_i = 0, \ p \geq 0$, no tiene solución $\boldsymbol{p} \in R^m$

Por consiguiente por el teorema de Gordan

$x_i y > 0, i = 1, 2, \cdots, m$, tiene una solución $\boldsymbol{y} \in R^n$

Obviamente $\boldsymbol{y} \neq 0$, y podemos tomar y tal que satisfaga $\boldsymbol{yy} = 1$. Entonces

$$y \in \bigcap_{i}^{m} \{\boldsymbol{y} \in R^n \ / \ \boldsymbol{yy} = 1, x_i \boldsymbol{y} \geq 0\} = \bigcap_{i=1}^{m} A_{x_i}$$

$$\therefore \ \bigcap_{i=1}^{m} A_{x_i} \neq \Phi$$

Los conjuntos $(A_x)_{x \in \Omega}$ son cerrados con respecto al conjunto compacto $\{\boldsymbol{y} \in R^n \ / \ \boldsymbol{yy} = 1\}$, por consiguiente por el teorema de la intersección finita tenemos que $\bigcap_{x \in \Omega} A_x \neq \Phi$. Sea $\boldsymbol{c}$ cualquier punto en la intersección. Entonces $\boldsymbol{cc} = 1$ y $\boldsymbol{cx} \geq 0$ para todo $\boldsymbol{x} \in \mathbf{\Omega}$.

Por consiguiente $\{\boldsymbol{x} \in R^n \ / \ \boldsymbol{cx} = 0\}$ es el plano de separación requerido.

Se observará que en el lema anterior no se impuso ninguna otra condición salvo la convexidad.

El siguiente ejemplo muestra que el lema anterior no puede ser más fuerte para $\boldsymbol{x} \in \boldsymbol{\Omega} \Rightarrow \mathrm{cx} > 0$ sin alguna suposición extra. El conjunto

$$\Omega = \{\boldsymbol{x} \in R^2 \ / \ \ \boldsymbol{x} \geq 0\} \cup \{\boldsymbol{x} \in R^2 \ / \ \ \boldsymbol{x_1} > 0, \boldsymbol{x_2} < 0\}$$

es convexo y no contiene el origen, pero no existe ningún plano

$\{\boldsymbol{x} \in R^n \ / \ \ cx = 0\}$ tal que $\boldsymbol{x} \in \boldsymbol{\Omega} \Rightarrow \mathrm{cx} > 0$ (ver figura 1.2.2)

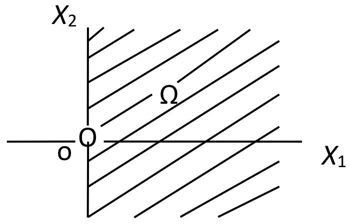


Fig. 3.14.2: Explicación gráfica del ejemplo

Por otro lado asumimos que $\boldsymbol{\Omega}$ es cerrado (o aún cuando al menos lo asumimos, es decir que el origen no es un punto de clausura), entonces podemos establecer un fuerte resultado, esto es, existe un plano que separa estrictamente el origen de $\boldsymbol{\Omega}$.

3.14.3. TEOREMA FUNDAMENTAL DE SEPARACIÓN.

San S y A dos conjuntos convexos disjuntos no vacíos en R^n, entonces existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \ / \ \boldsymbol{cx} = \alpha, \boldsymbol{c} \neq \boldsymbol{0}]$ que los separa, esto es,

$$\boldsymbol{x} \in S \Rightarrow \boldsymbol{cx} \leq \alpha$$

$$\boldsymbol{x} \in A \Rightarrow \boldsymbol{cx} \geq \alpha$$

DEMOSTRACIÓN.

El conjunto $A - S = \{\boldsymbol{x}/\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{z}, \boldsymbol{y} \in A, \boldsymbol{z} \in S\}$ es convexo por el corolario 1.1.20 y no contiene el origen porque $S \cap A = \Phi$. Según el lema 3.14.2 existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \ / \ \boldsymbol{cx} = \alpha, \boldsymbol{c} \neq \boldsymbol{0}]$ tal que

$$\boldsymbol{x} \in A - S \ \Rightarrow \ c\boldsymbol{x} \geq 0 \qquad \text{ó} \qquad \boldsymbol{y} \in A, \boldsymbol{z} \in S \Rightarrow c(\boldsymbol{y} - \boldsymbol{z}) \geq 0$$

Por lo tanto

$$\beta = \underset{y \in A}{\text{infimo}}\, c\boldsymbol{y} \geq \underset{z \in S}{\text{supremun}}\, c\boldsymbol{z} = \gamma$$

Definiendo

$$\alpha = \frac{\beta + \gamma}{2}$$

Entonces

$$\boldsymbol{z} \in S \implies c\boldsymbol{z} \leq \alpha$$

$$\boldsymbol{y} \in A \implies c\boldsymbol{y} \geq \alpha$$

Ahora, del teorema fundamental de separación derivamos un corolario, y de este corolario deducimos un lema, el cual será usado para establecer el teorema de estricta separación.

3.14.4. COROLARIO.

Sea $\boldsymbol{\Omega}$ un conjunto convexo no vacío en R^n. Si el origen $\boldsymbol{0}$ no es un punto de clausura de $\boldsymbol{\Omega}$ (o equivalentemente si el origen no está en la clausura $\overline{\boldsymbol{\Omega}}$ de $\boldsymbol{\Omega}$), entonces existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \;/\; \boldsymbol{cx} = \alpha,\; \alpha > 0, \boldsymbol{c} \neq \boldsymbol{0}]$ que separa estrictamente $\boldsymbol{\Omega}$ con $\boldsymbol{0}$, y recíprocamente. En otras palabras

$$0 \notin \overline{\Omega} \iff \exists\, c \neq 0, \alpha > 0 : \boldsymbol{x} \in \Omega \implies c\boldsymbol{x} > \alpha$$

DEMOSTRACIÓN.

($\Leftarrow$) Haremos la demostración por el absurdo.

Supongamos que existe $c \neq 0, \alpha > 0$ tal que $cx > \alpha$ para todo $\boldsymbol{x} \in \boldsymbol{\Omega}$. Si $0 \in \overline{\boldsymbol{\Omega}}$, entonces existe $x \in \boldsymbol{\Omega}$ tal que $\|x\| < \alpha/2\|c\|$, y por consiguiente

$$\frac{\alpha}{2} = \|c\| \frac{\alpha}{2\|c\|} > \|c\|\|x\| \geq |cx| > \alpha$$

lo cual es una contradicción. $\therefore\; 0 \notin \overline{\boldsymbol{\Omega}}$.

($\Rightarrow$) Dado que $\boldsymbol{0}$ no es un punto de clausura de $\boldsymbol{\Omega}$, existe una bola abierta
$B(\boldsymbol{0}, \varepsilon) = \{\, \boldsymbol{x} \in R^n \;/\; \|\boldsymbol{x}\| < \varepsilon\}$ centrada en 0 y de radio ε tal que

$$\boldsymbol{B}(\boldsymbol{0}, \varepsilon) \cap \boldsymbol{\Omega} = \boldsymbol{\Phi}.$$

Puesto que la bola $\boldsymbol{B}(\boldsymbol{0}, \varepsilon)$ es convexa, del teorema 3.14.3 (teorema fundamental de separación), seguimos que existe un plano

$H = [\boldsymbol{x} \in \boldsymbol{R^n} \ / \ \ \boldsymbol{cx} = \gamma, \ \boldsymbol{c} \neq \boldsymbol{0}]$, tal que

$$\boldsymbol{x} \in B(\boldsymbol{0}, \varepsilon) \ \Longrightarrow \ c\boldsymbol{x} \leq \gamma$$

$$\boldsymbol{x} \in \Omega \ \Longrightarrow \ \mathrm{c}\mathbf{x} \geq \gamma$$

Puesto que $\boldsymbol{B}(\boldsymbol{0}, \varepsilon)$ es una bola abierta, debe contener el vector $\boldsymbol{\delta c} \neq 0$ para algún $\delta > 0$.

Por consiguiente $\gamma \geq \boldsymbol{\delta c}c > 0$

Haciendo $\alpha = \delta cc/2 > 0$

entonces

$$\boldsymbol{x} \in \Omega \ \Longrightarrow \ \mathrm{c}\mathbf{x} \geq \gamma > \alpha > 0$$

3.14.5. LEMA.

Sea $\boldsymbol{\Omega}$ un conjunto convexo cerrado no vacío en R^n. Si $\boldsymbol{\Omega}$ no contiene el origen, entonces existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \ / \ \ \boldsymbol{cx} = \alpha,$ $\alpha > 0, \boldsymbol{c} \neq \boldsymbol{0}]$ que separa estrictamente $\boldsymbol{\Omega}$ con $\boldsymbol{0}$, y recíprocamente. En otras palabras

$$\boldsymbol{0} \notin \overline{\boldsymbol{\Omega}} \ \Leftrightarrow \ \exists \ c \neq 0, \alpha > 0 \colon \boldsymbol{x} \in \boldsymbol{\Omega} \ \Longrightarrow c\boldsymbol{x} > \alpha$$

DEMOSTRACIÓN.

Este lema se sigue del corolario 3.14.4 observando que los requisitos de que Ω sea cerrado y no contenga el origen $\boldsymbol{0}$ implica que $\boldsymbol{0}$ no es un punto de clausura de $\boldsymbol{\Omega}$, esto es $\boldsymbol{0} \notin \overline{\boldsymbol{\Omega}}$.

TEOREMA FUNDAMENTAL DE ESTRICTA SEPARACIÓN.

San S y A dos conjuntos convexos no vacíos en R^n, con S compacto y A cerrado. Si S y A son disjuntos, entonces existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} \ / \ \ \boldsymbol{cx} = \alpha, \boldsymbol{c} \neq \boldsymbol{0}]$ que los separa estrictamente, y recíprocamente. En otras palabras

$$S \cap A = \Phi \ \Longrightarrow \left\langle \begin{array}{c} \exists \ \boldsymbol{c} \neq \boldsymbol{0} \ y \ \alpha \\ \boldsymbol{x} \in S \ \Longrightarrow \boldsymbol{cx} < \alpha \\ \boldsymbol{x} \in A \ \Longrightarrow \boldsymbol{xc} > \alpha \end{array} \right\rangle$$

DEMOSTRACIÓN.

($\Longleftarrow$) Haremos la demostración por el absurdo.

Si $\boldsymbol{x} \in S \cap A$, entonces $cx < \alpha < cx$ es una contradicción.

($\Longrightarrow$) El conjunto $A - S = \{\boldsymbol{x}/\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{z}, \boldsymbol{y} \in A, \boldsymbol{z} \in S\}$ es convexo por el corolario 2.1.20 y cerrado por el corolario 1.4.10. Por consiguiente por el lema 3.14.5 existe un plano $H = [\boldsymbol{x} \in \boldsymbol{R^n} / \ \boldsymbol{cx} = \mu,\ \mu > 0, \boldsymbol{c} \neq \boldsymbol{0}]$ tal que

$$\boldsymbol{x} \in A - S \ \Longrightarrow\ c\boldsymbol{x} > \mu > 0 \qquad ó \qquad \boldsymbol{y} \in A, \boldsymbol{z} \in S \ \Rightarrow c(\boldsymbol{y} - \boldsymbol{z}) > \mu > 0$$

Por lo tanto

$$\beta = \inf_{\boldsymbol{y} \in A} c\boldsymbol{y} \geq \sup_{\boldsymbol{z} \in S} c\boldsymbol{z} + \mu = \gamma$$

Definiendo

$$\alpha = \frac{\beta + \gamma}{2}$$

entonces

$$\boldsymbol{z} \in S \ \Longrightarrow\ c\boldsymbol{z} < \alpha$$
$$\boldsymbol{y} \in A \ \Longrightarrow\ c\boldsymbol{y} > \alpha$$

Los teoremas de separación serán usados para derivar algunos teoremas fundamentales para funciones convexas, los cuales a su vez serán usados para derivar el criterio fundamental de optimización de Kuhn-Tucker de programación no lineal convexa y también la condición necesaria de optimización llamado principio del mínimo.

Remarcamos aquí que un teorema de la alternativa, el teorema de Gordan, fue fundamental para derivar los teoremas de separación. Podemos invertir el proceso y usar los teoremas de separación para derivar los teoremas de la alternativa. Así para determinar el teorema de Gordan, es decir ó $\boldsymbol{A^T y} = 0, \boldsymbol{y} \geq 0$ tiene solución $\boldsymbol{y} \in R^n$ ó $\boldsymbol{Ax} > 0$ tiene solución $\boldsymbol{x} \in R^n$, observamos que si $\boldsymbol{e} \in R^m$ es un vector de unos, entonces:

$\boldsymbol{A^T y} = 0, \boldsymbol{y} \geq 0$ no tiene solución $\Leftrightarrow 0 \notin \Omega = \{\boldsymbol{z}/\boldsymbol{z} = A^T\boldsymbol{y},\ \boldsymbol{y} \geq 0, \boldsymbol{e}\mathbf{y} = 1\}$

$$\Leftrightarrow\ \exists\ c \neq 0, \alpha > 0: \boldsymbol{z} \in \Omega\ \Longrightarrow \boldsymbol{cz} > \alpha$$
$$\Leftrightarrow \boldsymbol{y} \geq 0, \boldsymbol{ey} = 1\ \Longrightarrow cA^T\boldsymbol{y} > \alpha > 0\ \ \Leftrightarrow\ \ A\boldsymbol{c} > 0$$

Las últimas implicaciones se siguen tomando $\boldsymbol{y} = \boldsymbol{e^i} \in R^m,\ i = 1, 2, \cdots, m,$ donde $\boldsymbol{e^i}$ tiene ceros para todos los elementos excepto 1 para el i-ésimo elemento.

Usando la estructura del lema 1.2.5 podemos dar una interpretación geométrica de los teoremas de Gordan como sigue: o el origen $\boldsymbol{0} \in R^n$ está en el casco convexo de los vectores renglón

$A_1, A_2, \cdots, A_n$ de la matriz A ($A^T\boldsymbol{y} = 0, \boldsymbol{y} \geq 0$ tiene una solución), o no está (en tal caso por el lema 3.14.2 $A\boldsymbol{x} > 0$ tiene una solución $\boldsymbol{x} = \boldsymbol{c}$). Más generalmente, si $\boldsymbol{\Omega}$ es cualquier conjunto convexo cerrado no vacío en R^n, entonces o contiene el origen o no lo contiene (en tal caso por el lema 3.14.5 existe un vector $\boldsymbol{c} \in R^n$ el cual hace un ángulo agudo estricto con cada $\boldsymbol{x} \in \boldsymbol{\Omega}$.

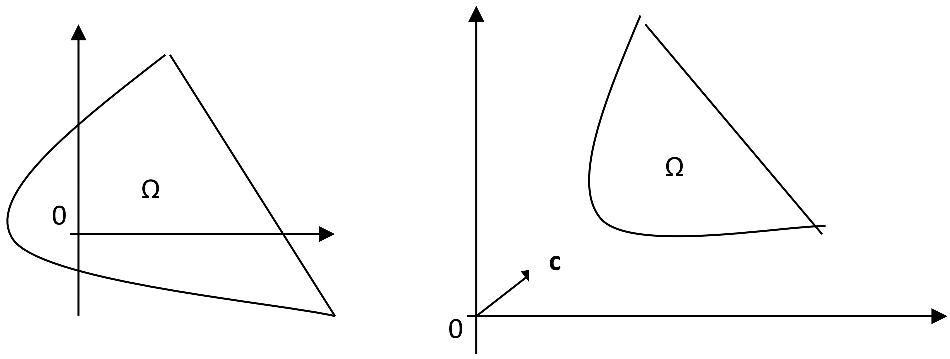
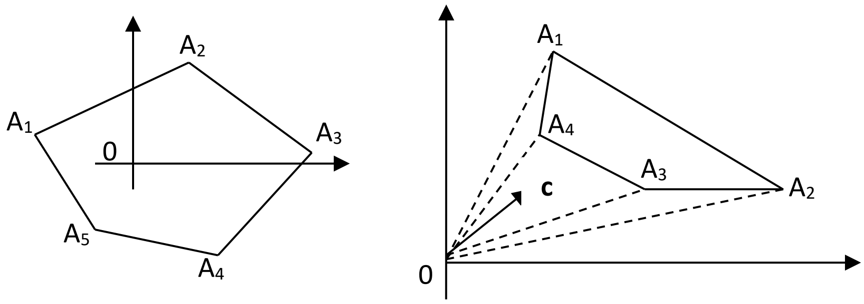


Fig. 3.14.2.: Interpretación geométrica del lema 3.14.5.

a) El conjunto **Ω** contiene el origen
b) El conjunto **Ω** no contiene el origen



Gráfica 3.14.5: Interpretación geométrica del teorema de Gordan usando el lema 3.14.5

a) $A^T y = 0, y \geq 0$ tiene solución; $Ax > 0$ no tiene solución.
b) $Ax > 0$ tiene solución. $A^T y = 0, y \geq 0$ no tiene solución.

4. FUNCIONES CÓNCAVAS Y CONVEXAS.

INTRODUCCIÓN. En esta sección analizaremos las funciones cóncavas, convexas, estrictamente cóncavas y estrictamente convexas definidas sobre subconjuntos de R^n. Las funciones cóncavas y convexas son sumamente importantes en programación no lineal dado que ellas están entre las pocas funciones para las cuales podemos dar criterios suficientes en optimización, y ellas son las únicas funciones, para las que podemos establecer condiciones necesarias de optimización sin linealización; como por ejemplo las condiciones de punto silla de KUHN – TUCKER.

En este capítulo también daremos algunas de las propiedades básicas de las funciones cóncavas y convexas y obtendremos algunos teoremas que involucran estas funciones. Estos teoremas derivados usando los teoremas de separación para conjuntos convexos, son muy parecidos a los teoremas de la alternativa para sistemas lineales. En este sentido las funciones cóncavas y convexas heredan algunas de las propiedades de las funciones lineales.

4.1. FUNCIONES CÓNCAVAS Y CONVEXAS: DEFINICIONES Y PROPIEDADES.

4.1.1. FUNCIÓN CONVEXA.

Definición. Una función numérica θ definida sobre un conjunto $S \subset R^n$ se dice que es convexa en $\bar{\boldsymbol{x}} \in S$ (con respecto a S) si

$$\left.\begin{array}{c} \boldsymbol{x} \in \boldsymbol{S} \\ 0 \leq \lambda \leq 1 \\ (1-\lambda)\bar{\boldsymbol{x}} + \lambda \boldsymbol{x} \in S \end{array}\right\} \Longrightarrow (1-\lambda)\theta(\bar{\boldsymbol{x}}) + \lambda\theta(\boldsymbol{x}) \geq \theta[(1-\lambda)\bar{\boldsymbol{x}} + \lambda\boldsymbol{x}]$$

se dice que θ es convexa sobre S si ella es convexa en cada $\boldsymbol{x} \in S$.

De la definición anterior inmediatamente seguimos que una función numérica θ definida sobre un conjunto convexo S es convexa sobre S si y sólo si

$$\left.\begin{array}{c} x_1,\ x_2 \in S \\ 0 \leq \lambda \leq 1 \end{array}\right\} \Longrightarrow (1-\lambda)\theta(x_1) + \lambda\theta(\boldsymbol{x_2}) \geq \theta[(1-\lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}]$$

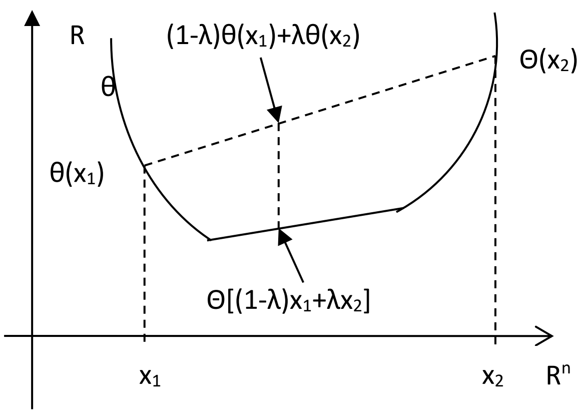


Fig. 4,1,1 (a): Función convexa sobre R

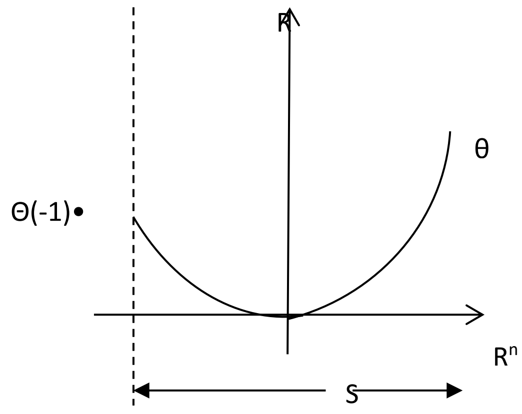


Fig. 4.1.1 (b): Función convexa sobre $S=[-1,\infty)$

4.1.2. FUNCIÓN CÓNCAVA.

Definición. Una función numérica θ definida sobre un conjunto $S \subset R^n$ se dice que es cóncava en $\bar{\boldsymbol{x}} \in S$ (con respecto a S) si

$$\left.\begin{array}{c} \boldsymbol{x} \in \boldsymbol{S} \\ 0 \leq \lambda \leq 1 \\ (1-\lambda)\bar{\boldsymbol{x}} + \lambda \boldsymbol{x} \in S \end{array}\right\} \Longrightarrow (1-\lambda)\theta(\bar{\boldsymbol{x}}) + \lambda\theta(\boldsymbol{x}) \leq \theta[(1-\lambda)\bar{\boldsymbol{x}} + \lambda\boldsymbol{x}]$$

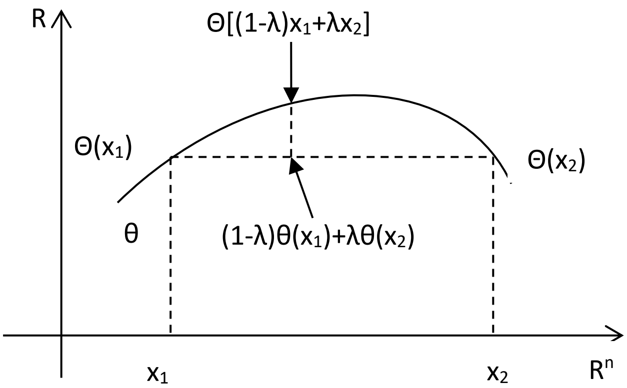
Se dice que θ es cóncava sobre S si ella es cóncava en cada $\boldsymbol{x} \in S$.

Obviamente θ es cóncava en $\bar{\boldsymbol{x}} \in S$ (cóncava sobre S) si y sólo si $-\theta$ es convexa en $\bar{\boldsymbol{x}}$ (convexa sobre S). Los resultados obtenidos para funciones convexas pueden ser cambiados en los resultados para funciones cóncavas por la apropiada multiplicación por -1, y viceversa.

De la definición anterior seguimos inmediatamente que una función numérica θ definida sobre un conjunto convexo S es cóncava sobre S si y sólo si

$$\left.\begin{array}{c} x_1,\ x_2 \in S \\ 0 \leq \lambda \leq 1 \end{array}\right\} \Longrightarrow (1-\lambda)\theta(\boldsymbol{x_1}) + \lambda\theta(\boldsymbol{x_2}) \leq \theta[(1-\lambda)\boldsymbol{x_1} + \lambda\boldsymbol{x_2}]$$

La gráfica 4.1.2 muestra dos funciones cóncavas sobre conjuntos convexos de $R^n = R$



Fig, 4.1.2 (a): Función cóncava θ sobre R

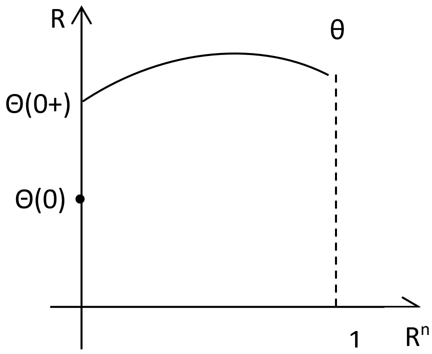


Fig. 24.1.2 (b): Función cóncava θ sobre $S = [0, 1]$

4.1.3. FUNCIÓN ESTRICTAMENTE CONVEXA.

Definición. Una función numérica θ definida sobre un conjunto $S \subset R^n$ se dice que es estrictamente convexa en $\bar{\boldsymbol{x}} \in S$ (con respecto a S) si

$$\left.\begin{array}{c}\boldsymbol{x} \in \boldsymbol{S} \\ x \neq \overline{x} \\ 0 < \lambda < 1 \\ (1-\lambda)\overline{\boldsymbol{x}} + \lambda\boldsymbol{x} \in S\end{array}\right\} \Rightarrow (1-\lambda)\theta(\overline{\boldsymbol{x}}) + \lambda\theta(\boldsymbol{x}) > \theta[(1-\lambda)\overline{\boldsymbol{x}} + \lambda\boldsymbol{x}]$$

Se dice que θ es estrictamente convexa sobre S si ella es estrictamente convexa en cada $\boldsymbol{x} \in S$.

4.1.4. FUNCIÓN ESTRICTAMENTE CÓNCAVA.

Definición. Una función numérica θ definida sobre un conjunto $S \subset R^n$ se dice que es estrictamente cóncava en $\overline{\boldsymbol{x}} \in S$ (con respecto a S) si

$$\left.\begin{array}{c}\boldsymbol{x} \in \boldsymbol{S} \\ x \neq \overline{x} \\ 0 < \lambda < 1 \\ (1-\lambda)\overline{\boldsymbol{x}} + \lambda\boldsymbol{x} \in S\end{array}\right\} \Rightarrow (1-\lambda)\theta(\overline{\boldsymbol{x}}) + \lambda\theta(\boldsymbol{x}) < \theta[(1-\lambda)\overline{\boldsymbol{x}} + \lambda\boldsymbol{x}]$$

Se dice que θ es estrictamente cóncava sobre S si ella es estrictamente cóncava en cada $\boldsymbol{x} \in S$.

Obviamente una función estrictamente convexa (estrictamente cóncava) sobre un conjunto $S \subset R^n$ es convexa (cóncava) sobre S, pero no recíprocamente. Por ejemplo una función constante sobre R^n es cóncava y convexa sobre R^n. En realidad, puede ser fácilmente probado que todas las funciones lineales $\theta(x) = c\boldsymbol{x} + \alpha$ sobre R^n no son ni estrictamente convexa ni estrictamente cóncava sobre R^n. Por consiguiente, a consecuencia de la porción lineal, la función mostrada en la figura 4.1.1(a) no es estrictamente convexa sobre R, pero la función mostrada en la figura 4.1.1(b) es estrictamente convexa sobre $[-1, \infty\rangle$. Ambas funciones de la figura 4.1.2 son estrictamente cóncava sobre sus dominios de definición.

Una función vectorial $n-$ dimensional f definida sobre un conjunto S en R^n es convexa en $\overline{x} \in S$, convexa sobre S, etc., si cada una de sus componentes f_i, $i = 1, 2, \cdots, m$, es convexa en $\overline{x} \in S$, convexa sobre S, etc.

4.1.5. TEOREMA.

Sea $\boldsymbol{f} = (f_1, f_2, \cdots, f_m)$ una función vectorial $m-$dimensional definida sobre $S \subset R^n$. Si $\boldsymbol{f}$ es convexa en $\overline{x} \in S$ (convexa sobre S), entonces cada combinación lineal no negativa de sus componentes f_i

$$\theta(x) = pf(x), \;\; p \geq 0$$

es convexa en $\bar{\boldsymbol{x}}$ (convexa sobre S)

DEMOSTRACIÓN.

Sea $x \in S, 0 \leq \lambda \leq 1$, y sea $(1-\lambda)\bar{\boldsymbol{x}} + \lambda\boldsymbol{x} \in S$. Entonces

$$\theta[(1-\lambda)\bar{\boldsymbol{x}} + \lambda\boldsymbol{x}] = pf[(1-\lambda)\bar{\boldsymbol{x}} + \lambda\boldsymbol{x}]$$

$$\leq p[(1-\lambda)f(\bar{\boldsymbol{x}}) + \lambda f(\boldsymbol{x})]$$

$$= (1-\lambda)f(\bar{\boldsymbol{x}}) + \lambda f(\boldsymbol{x}) = (1-\lambda)\theta(\bar{\boldsymbol{x}}) + \lambda\theta(\boldsymbol{x})$$

4.1.6. TEOREMA.

Para que una función numérica θ definida sobre un conjunto convexo $S \subset R^n$ sea convexa sobre S es necesario y suficiente que su Epigrafo (Conjunto de puntos situados en o sobre el grafo de la función)

$$G_\theta = \{(x,\xi)/\, x \in S, \xi \in R, \theta(x) \leq \xi\} \subset R^{n+1}$$

sea un conjunto convexo en R^{n+1}

DEMOSTRACIÓN.

($\Leftarrow$) Asuma que G_θ es un conjunto convexo. Sean $x_1, x_2 \in S$ entonces $[x_1, \theta(x_1)] \in G_\theta$ y $[x_2, \theta(x_2)] \in G_\theta$. Por la convexidad de G_θ tenemos que

$[(1-\lambda)x_1 + \lambda x_2, (1-\lambda)\theta(x_1) + \lambda\theta(x_2)] \in G_\theta$ para $0 \leq \lambda \leq 1$, ó

$\theta[(1-\lambda)x_1 + \lambda\theta(x_2)] \leq (1-\lambda)\theta(x_1) + \lambda\theta(x_2)$ para $0 \leq \lambda \leq 1$

y por consiguiente θ es convexa sobre S. ■

($\Rightarrow$) Asuma que θ es convexa sobre S. Sea $x_1, \xi_1 \in G_\theta$ y $x_2, \xi_2 \in G_\theta$. Por la convexidad de θ sobre S tenemos que para $0 \leq \lambda \leq 1$

$$\theta[(1-\lambda)x_1 + \lambda x_2] \leq (1-\lambda)\theta(x_1) + \lambda\theta(x_2) \leq (1-\lambda)\xi_1 + \lambda\xi_2$$

$$\therefore\ [(1-\lambda)x_1 + \lambda x_2, (1-\lambda)\xi_1 + \lambda\xi_2] \in G_\theta$$

y G_θ es un conjunto convexo en R^{n+1} ■

4.1.7. COROLARIO.

Para que una función numérica θ definida sobre un conjunto convexo $S \subset R^n$ sea convexa sobre S es necesario y suficiente que su Hipografo (Conjunto de puntos situados en o bajo el grafo de la función)

$H_\theta = \{(x,\xi)/\ x \in S, \xi \in R, \theta(x) \geq \xi\} \subset R^{n+1}$, sea un conjunto

convexo en R^{n+1}

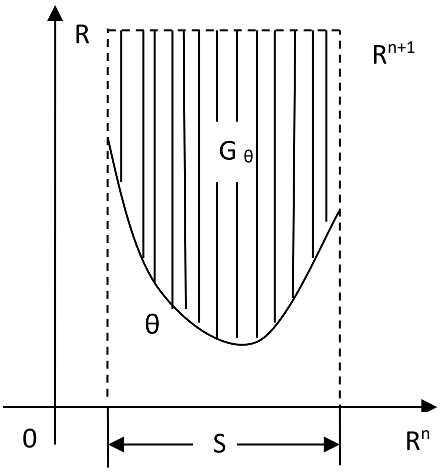


Fig. 4.1.7. Función convexa y su Epígrafo convexo G_θ

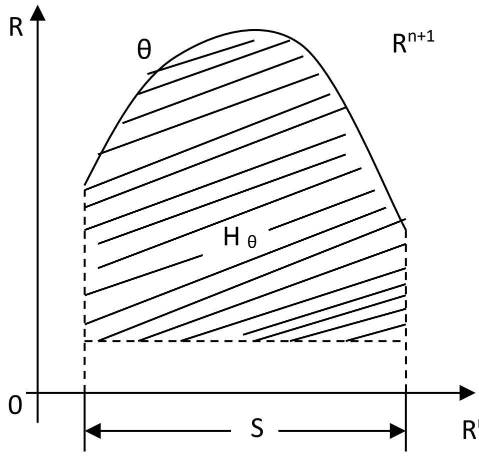


Fig. 4.1.7. Función convexa y su Hipográfo convexo H_θ

La figura 4.1.7 a) muestra una función convexa sobre S y su Epigrafo convexo G_θ. La figura 4.1.7 b)

una función cóncava y su Hipográfo convexo H_θ.

4.1.8. TEOREMA.

Sea θ una función numérica definida sobre un conjunto convexo $S \subset R^n$. Una condición necesaria pero no suficiente para que θ sea convexa sobre S es que el conjunto

$$A_\alpha = \{\boldsymbol{x} \in S\ /\ \theta(\boldsymbol{x}) \leq \alpha\} \subset S \subset R^n$$

sea convexo para cada número real α.

DEMOSTRACIÓN.

($\Longrightarrow$) Sea θ una función convexa sobre S y sean $x_1,\ x_2 \in A_\alpha$. Entonces:

$\theta[(1-\lambda)x_1 + \lambda x_2] \leq (1-\lambda)\theta(x_1) + \lambda\theta(x_2)$ (Por convexidad de θ)

$\leq (1-\lambda)\alpha + \lambda\alpha = \alpha$ (Porque $x_1,\ x_2 \in A_\alpha$) $\leq \alpha$

Por consiguiente

$(1-\lambda)x_1 + \lambda x_2 \in A_\alpha$ y A_α es convexo.■

(⟸) Hacemos la demostración por el absurdo, de la siguiente manera:

Supongamos que: si A_α es convexo para cada α , no se sigue que θ sea una función convexa sobre S. Esto lo haremos a través de un contraejemplo. Consideremos la función θ sobre R definida por $\theta(x) = x^3$ Tenemos que la función $\theta(x) = x^3$ no es convexa sobre R, sin embargo el conjunto $A_\alpha = \{\boldsymbol{x}/\boldsymbol{x} \in R, x^3 \leq \alpha\} = \{x/x \in R, x \leq \alpha^{1/3}\}$ es obviamente convexo para cualquier α. (Ver la definición de semiespacio).■

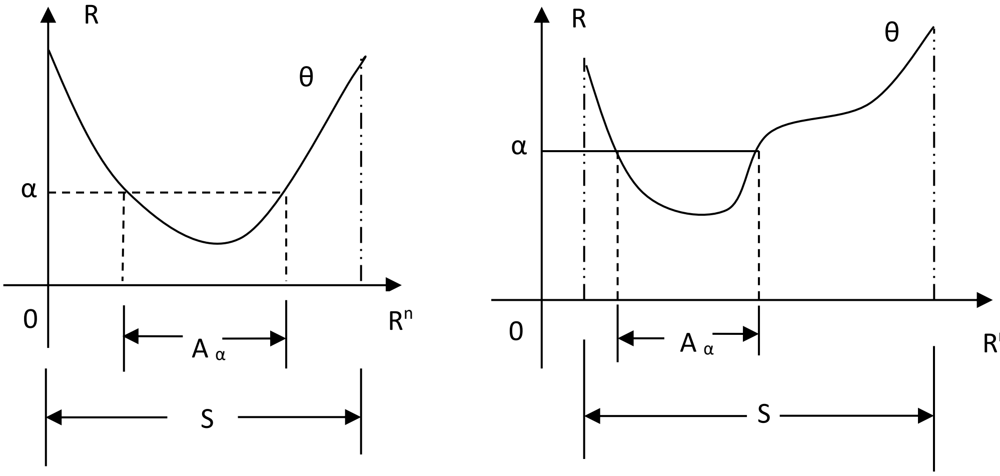
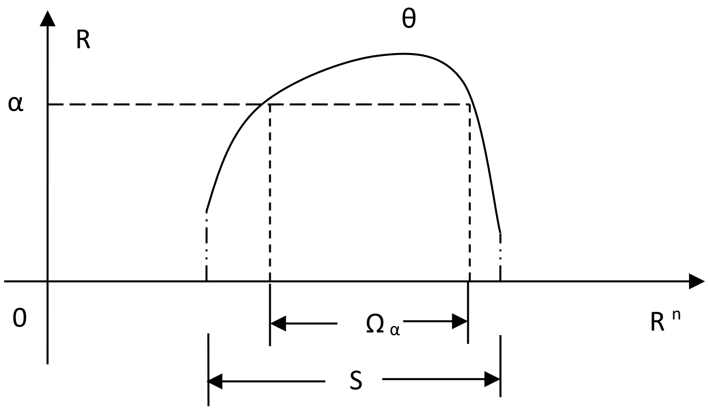


Fig. 4.1.8: El conjunto convexo A $_\alpha$ asociado a la función θ



Graf. 4.1.9: El conjunto convexo Ω $_\alpha$ asociado a la función θ

4.1.9. COROLARIO.

Sea θ una función numérica definida sobre el conjunto $S \subset R^n$. Una condición necesaria pero no suficiente para que θ sea convexa sobre S es que el conjunto:

$$\Omega_\alpha = \{\boldsymbol{x} \in S \ / \ \theta(\boldsymbol{x}) \geq \alpha\} \subset S \subset R^n$$

sea convexo para cada número real α.

4.1.10. TEOREMA.

Si $(\theta_i)_{i\in I}$ es una familia (finita o infinita) de funciones numéricas las cuales son convexas acotadas superiormente sobre un conjunto convexo $S \subset R^n$, entonces la función numérica

$$\theta(x) = \operatorname*{Sup}_{i\in I} \theta_i(x)$$

es una función convexa sobre S.

DEMOSTRACIÓN.

Puesto que cada θ_i es una función convexa sobre S, entonces por el teorema 3.1.6, sus Epigrafos

$$G_{\theta_i} = \{(x,\xi)/\ x \in S, \xi \in R, \theta_i(x) \leq \xi\}$$

son conjuntos convexos en R^{n+1}, y por consiguiente sus intersecciones

$$\bigcap_{i\in I} G_{\theta_i} = \{(\boldsymbol{x},\xi)/\boldsymbol{x} \in S, \xi \in R, \theta_i(\boldsymbol{x}) \leq \xi, \forall\ i \in I\}$$

$$= \{(x,\xi)/\ x \in S, \xi \in R, \theta(x) \leq \xi\}$$

es también un conjunto convexo en R^{n+1}. Pero esta intersección convexa es la epigrafo de θ. Por consiguiente, por teorema 4.1.6, θ es una función convexa sobre S ■

4.1.11. COROLARIO.

Si $(\theta_i)_{i\in I}$ es una familia (finita o infinita) de funciones numéricas que son cóncavas acotadas inferiormente sobre un conjunto convexo $S \subset R^n$, entonces la función numérica

$$\theta(x) = \operatorname*{Inf}_{i\in I} \theta_i(x)$$

es una función cóncava sobre S.

Finalizamos esta sección remarcando que una función θ que es convexa sobre un conjunto convexo $S \subset R^n$ no es necesariamente función continua. Por ejemplo sobre la media recta $S = \{\boldsymbol{x}/\boldsymbol{x} \in R, \boldsymbol{x} \geq -1\}$, la función numérica

$$\theta(x) = \begin{cases} 2 & para\ x = -1 \\ x^2 & para\, x > -1 \end{cases}$$

es una función convexa sobre S, pero obviamente no es continua en $x = -1$. Sin embargo, si S es un conjunto convexo abierto, entonces una función convexa θ sobre S es realmente continua. Este hecho se establece en el siguiente teorema.

4.1.12. TEOREMA.

Sea S un conjunto convexo abierto en R^n. Si θ es una función numérica convexa sobre S, entonces θ es continua sobre S.

DEMOSTRACIÓN.

Sea $x_0 \in S,$ y sea $\propto$ la distancia de x_0 al punto más cercano de R^n que no está en S ($\propto = +\infty\ si\ S = R^n$) Sea C un n-cubo con centro x_0 y longitud de lado 2δ, esto es

$$C = \{\boldsymbol{x}/\boldsymbol{x} \in R^n, -\delta \leq x_i - x_{0i} \leq \delta, i = 1, 2, \cdots, n\}$$

Haciendo $x^{1/2}\delta < \propto$, tenemos que $C \subset S$. Si V denota el conjunto de 2^n vértices de C. Sea

$$\beta = \max_{x \in V} \theta(x)$$

Por teorema 3.1.8, el conjunto $A_\beta = \{\boldsymbol{x}/\boldsymbol{x} \in S, \theta(x) \leq \beta\}$ es convexo. Puesto que C es el casco convexo de V (esto puede ser probado fácilmente por inducción sobre n) y $V \subset A_\beta$, por el teorema 3.1.10, seguimos que $C \subset A_\beta$ (Ver figura 4.1.12)

Sea $\boldsymbol{x}$ cualquier punto tal que $0 < \|\boldsymbol{x} - \boldsymbol{x_0}\| < \delta$, y definamos $\boldsymbol{x_0} + \boldsymbol{u}$, $\boldsymbol{x_0} - \boldsymbol{u}$ sobre la recta que pasa por los puntos $\boldsymbol{x_0}$ y $\boldsymbol{x}$ como se muestra en la figura 3.1.12. Ahora escribimos $\boldsymbol{x}$ como una combinación convexa de $\boldsymbol{x_0}$ y de $\boldsymbol{x_0} + \boldsymbol{u}$, y $\boldsymbol{x_0}$ como una combinación convexa de $\boldsymbol{x}$ y de $\boldsymbol{x_0} - \boldsymbol{u}$. Si $\lambda = \|x - x_0\|/\delta$, entonces

$$x = x_0 + \lambda u = \lambda(x_0 + u) + (1 - \lambda)x_0$$

$$x_0 = x - \lambda u = x + \lambda(x_0 - u) - \lambda x_0$$

$$= \frac{1}{1+\lambda}x + \frac{\lambda}{1+\lambda}(x_0 - u)$$

Puesto que θ es convexa sobre S

$$\theta(x) \leq \lambda\theta(x_0 + u) + (1-\lambda)\theta(x_0) \leq \lambda\beta + (1-\lambda)\theta(x_0)$$

$$\theta(x_0) \leq \frac{1}{1+\lambda}\theta(x) + \frac{\lambda}{1+\lambda}\theta(x_0 - u) \leq \frac{\theta(x) + \lambda\beta}{1+\lambda}$$

Estas ecuaciones dan

$$-\lambda[\beta - \theta(x_0)] \leq \theta(x) - \theta(x_0) \leq \lambda[\beta - \theta(x_0)] \qquad \text{ó}$$

$$|\theta(x) - \theta(x_0)| \leq \frac{\beta - \theta(x_0)}{\delta}\|x - x_0\|$$

Así para cualquier $\varepsilon > 0$ dado seguimos que $|\theta(x) - \theta(x_0)| < \varepsilon$ para todo x que satisface $[\beta - \theta(x_0)]\|x - x_0\| < \delta$, y por consiguiente $\theta(x)$ es continua en x_0 ■

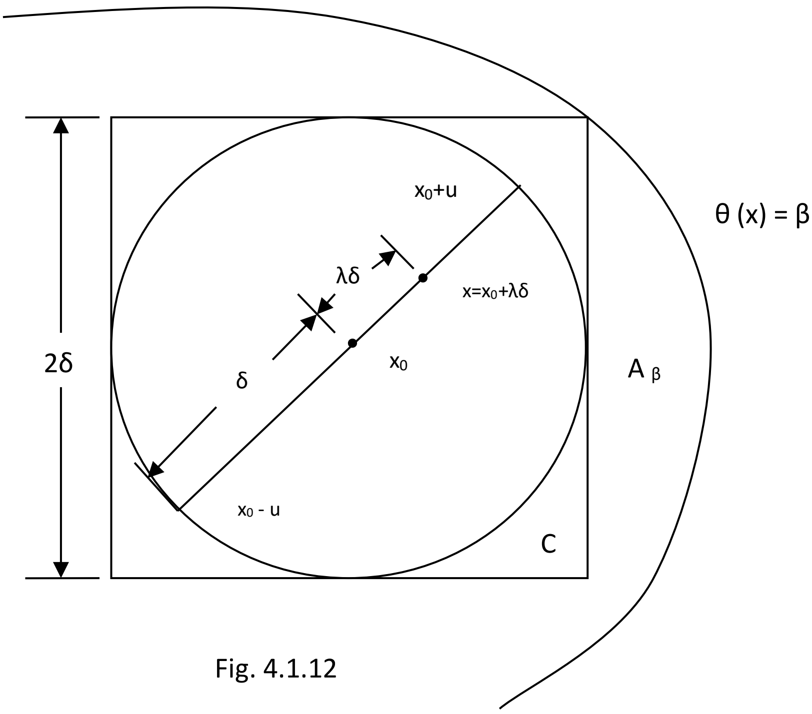


Fig. 4.1.12

esto que el interior de cada conjunto $S \subset R^n$ es abierto, seguimos que si θ es una función convexa sobre un conjunto convexo $S \subset R^n$, es continua en su interior.

4.2. TEOREMAS FUNDAMENTALES PARA FUNCIONES CONVEXAS.

En esta sección presentaremos algunos teoremas de vital importancia para derivar condiciones de optimización en programación no lineal. Estos teoremas son considerados como extensión de los teoremas de la alternativa.

4.2.1. TEOREMA.

Sea S un conjunto convexo no vacío en R^n, sea f una función vectorial m-dimensional sobre S, y sea h una función vectorial lineal k-dimensional en R^n. Si

$$\left\langle \begin{matrix} f(x) < 0 \\ h(x) = 0 \end{matrix} \right\rangle \; no\ tiene\ solución\ x \in S$$

entonces existe $p \in R^m$ y $q \in R^k$ tal que

$$\left\langle \begin{matrix} p \geqq 0,\ (p,q) \neq 0 \\ pf(x) + qh(x) \geq 0,\ \forall\, x \in S \end{matrix} \right\rangle$$

OBSERVACIÓN. $p \geqq 0$ y $(p,q) \neq 0$ no implica que $p \geq 0$ y $q \neq 0$, sino que implica $p \geq 0$ ó $q \neq 0$ ó $ambos$. Sin embargo si se suprime la ecuación lineal $h(x) = 0$, entonces $p \geq 0$.

DEMOSTRACIÓN.

Definamos los conjuntos

$$A(x) = \{(\boldsymbol{x}, \boldsymbol{y}) / \boldsymbol{y} \in R^m, \boldsymbol{z} \in R^k, \boldsymbol{y} > f(\boldsymbol{x}), \boldsymbol{z} = h(\boldsymbol{x})\} \ \ \boldsymbol{x} \in S$$

y

$$A = \bigcup_{x \in S} A(\boldsymbol{x})$$

Por Hipótesis A no contiene el origen $0 \in R^{m+k}$. También, A es convexo, porque si $(\boldsymbol{y_1}, \boldsymbol{z_1})$ y $(\boldsymbol{y_2}, \boldsymbol{z_2})$ están en A, entonces para $0 \leqq \lambda \leqq 1$

$$(1-\lambda)y_1 + \lambda \boldsymbol{y_2} > (1-\lambda) f(\boldsymbol{x_1}) + \lambda f(\boldsymbol{x_2}) \geqq f[(1-\lambda)\boldsymbol{x}_1 + \lambda \boldsymbol{x_2}], \quad \text{y}$$

$$(1-\lambda)\boldsymbol{z_1} + \lambda \boldsymbol{z_2} = (1-\lambda)h(\boldsymbol{x_1}) + \lambda h(\boldsymbol{x_2}) = h[(1-\lambda)\boldsymbol{x_1} + \lambda \boldsymbol{x_2}]$$

Puesto que A es un conjunto convexo no vacío que no contiene el origen, por el lema 2.2.2 seguimos que existe $p \in R^m, q \in R^k, (p,q) \neq 0$ tal que $(u,v) \in A \implies pu + qv \geqq 0$.

Puesto que cada u_i puede tomarse tan grande como se desee, $p \geqq 0$.

Sea $\varepsilon > 0, u = f(x) + e\varepsilon, v = h(x), x \in S$, donde e es un vector de unos en R^m. Por consiguiente $(u,v) \in A(x) \subset A$, y

$pu + qv = pf(x) + \varepsilon pe + qh(x) \geqq 0;\ x \in S$, ó

$pf(x) + qh(x) \geqq -\varepsilon pe$ para $x \in S$.

Ahora, si

$$\operatorname*{Inf}_{x \in S} pf(x) + qh(x) = -\delta < 0.$$

escogiendo ε tal que $\varepsilon pe < \delta$, tenemos que

$$\operatorname*{Inf}_{x \in S} pf(x) + qh(x) = -\delta < -\varepsilon pe$$

El cual es una contradicción con el hecho que $pf(x) + qh(x) \geq -\varepsilon pe$ para todo $x \in S$. Por consiguiente

$\operatorname*{Inf}_{x \in S} pf(x) + qh(x) \geqq 0.$ ■

Si observamos que para una función vectorial m-dimensional f definida sobre $S \subset R^n$ tenemos que

$$\left\langle \begin{matrix} f(x) < 0 \\ \text{tiene una solución} \\ x \in S \end{matrix} \right\rangle \implies \left\langle \begin{matrix} f(x) \leq 0 \\ \text{tiene una solución} \\ x \in S \end{matrix} \right\rangle \implies \left\langle \begin{matrix} f(x) \leqq 0 \\ \text{tiene una solución} \\ x \in S \end{matrix} \right\rangle$$

y

$$\left\langle \begin{matrix} f(x) < 0 \\ \text{no tiene solución} \\ x \in S \end{matrix} \right\rangle \impliedby \left\langle \begin{matrix} f(x) \leq 0 \\ \text{no tiene solución} \\ x \in S \end{matrix} \right\rangle \impliedby \left\langle \begin{matrix} f(x) \leqq 0 \\ \text{no tiene solución} \\ x \in S \end{matrix} \right\rangle$$

entonces, el siguiente corolario es consecuencia directa del teorema 4.2.1.

4.2.2. COROLARIO.

Sea S un conjunto convexo no vacío en R^n., sean f_1, f_2, f_3 funciones vectoriales convexas $m_1-, m_2-, y\, m_3$ –dimensional sobre S, y h una función vectorial lineal $k-$ dimensional sobre R^n. Si

$$\left\langle \begin{array}{c} f_1(x) < 0, f_2(x) \leq 0, f_3(x) \leqq 0 \\ h(x) = 0 \end{array} \right\rangle$$

no tiene solución $x \in S$, entonces existe $p_1 \in R^{m_1}, p_2 \in R^{m_2}, p_3 \in R^{m_3}$, y

$q \in R^k$ tal que

$$\left\langle \begin{array}{c} p_1, p_2, p_3 \geqq 0, (p_1, p_2, p_3, q) \neq 0 \\ p_1 f_1(x) + p_2 f_2(x) + p_3 f_3(x) + q h(x) \geqq 0, \forall x \in S \end{array} \right\rangle$$

Ahora enunciamos una generalización del teorema de Gordan de la alternativa para funciones convexas sobre un conjunto convexo arbitrario en R^n.

4.2.3. TEOREMA GENERALIZADO DE GORDAN.

Sea f una función vectorial convexa $m-$ dimensional sobre el conjunto convexo $S \subset R^n$. Entonces, cualquiera de las dos proposiciones se cumple, pero no ambos:

(I) $f(x) < 0$ tiene una solución $x \in S$

(II) $pf(x) \geqq 0 \; para \; todo \; x \in S \; para \; alg\acute{u}n \; p \geq 0, p \in R^m$

DEMOSTRACIÓN.

$(I \Rightarrow \overline{II})$ Sea $\overline{x} \in S$ una solución de $f(x) < 0$. Entonces para cualquier $p \geq 0$ en R^n, $pf(\overline{x}) < 0$, y por consiguiente II no puede ser válido.

$(\overline{I} \Longrightarrow II)$ Esto lo seguimos directamente del teorema 3.2.1 eliminando $h(x) = 0$ del teorema.■

Para ver que 3.2.3 es realmente una generalización del teorema de Gordan hagamos $f(x) = Ax$, donde A es una matriz $m \times n$. Entonces

$$\left\langle \begin{array}{c} Ax > 0 \\ no \; tiene \; soluci\acute{o}n \\ x \in R^n \end{array} \right\rangle \Leftrightarrow \left\langle \begin{array}{c} pAx \geqq 0 \; para \; todo \; x \in R^n \\ para \; alg\acute{u}n \; p \geq 0, p \in R^m \end{array} \right\rangle$$

$$\Leftrightarrow \left\langle \begin{array}{c} A^T p = 0, p \geq 0 \\ para \; alg\acute{u}n \; p \in R^m \end{array} \right\rangle$$

donde la última equivalencia se sigue tomando $x = \pm e^i, i = 1, 2, \cdots, n$, donde $e^i \in R^n$ tiene ceros para todos sus elementos excepto 1 para el i-ésimo elemento.

4.2.4. TEOREMA.

Sea f una función convexa $m-$ dimensional dada sobre R^n, sea B una matriz dada de orden $k \times n$ con renglones linealmente independientes, y sea $\boldsymbol{d}$ un vector $k-$ dimensional dado. Entonces, cualquiera de las dos proposiciones se cumple, pero no ambos:

(I) $f(x) < 0, B\boldsymbol{x} = \boldsymbol{d}$ tiene una solución $\boldsymbol{x} \in r^n$

(II) $pf(x) + q(B\boldsymbol{x} - \boldsymbol{d}) \geq 0$ para todo $x \in R^n$ para algún $p \geq 0, p \in R^m, q \in R^k$

DEMOSTRACIÓN

$(I \Rightarrow \overline{II})$ Sea $\overline{x} \in R^n$ una solución de $f(x) < 0$ y $Bx = d$. Entonces para cualquier $p \geq 0, p \in R^m, q \in R^k$ respectivamente,

$$pf(\overline{x}) + q(B\overline{x} - d) < 0$$

Por consiguiente II no puede ser válido

$(\overline{I} \Longrightarrow II)$ Si I no tiene solución, entonces por el teorema 3.2.1 existe $p \geq 0,\ (p, q) \neq 0$ tal que

$$pf(x) + q(Bx - d) \geq 0 \ \ \forall x \in R^n$$

Si $p \geq 0$, el teorema queda probado. Por lo contrario, asumamos que $p = 0$, y hallemos una contradicción. Si $p = 0$, entonces

$q(Bx - d) \geq 0 \ \ \forall x \in R^n$ para algún $q \neq 0$

En seguida mostraremos que $B^T q = 0$. Pero, si $B^T q \neq 0$, entonces elegimos $x = -qB$ para el caso que $qd \geq 0$, y $x = 2(qd)qB/qBB^T q$ para el caso cuando $qd < 0$, obtenemos que $q(Bx - d) < 0$. Por consiguiente $B^T q = 0$ para algún $q \neq 0$, lo cual contradice la suposición que los renglones de B son linealmente independientes.

4.2.5. TEOREMA.

Sea S un conjunto convexo compacto no vacío en R^n y sea $(f_i)_{i\in M}$ una familia (finita o infinita) de funciones numéricas que son convexas y semicontinua inferior sobre S, y sea $(h_i)_{i\in K}$ una familia (finita o infinita) de funciones numéricas lineales sobre R^n. Si

$$\left\langle \begin{matrix} f_i(x) \leq 0, i \in M \\ h_i(x) = 0,\ i \in K \end{matrix} \right\rangle \textit{no tiene solución } x \in S$$

Entonces para alguna subfamilia finita $\left(f_{i_1,} \cdots, f_{i_n}\right)$ de $(f_i)_{i\in M}$ y alguna subfamilia $\left(h_{i_1,} \cdots, h_{i_n}\right)$ de $(h_i)_{i\in K}$ existe $p \in R^m$ y $q \in R^k$ tal que

$$\left\langle \begin{matrix} p \geq 0, (p,q) \neq 0, y \\ \sum_{j=1}^{m} p_j f_{i_j}(x) + \sum_{j=1}^{k} q_j h_{i_j}(x) \geq 0 \ \ \forall x \in S \end{matrix} \right\rangle$$

Si K es vacío, esto es si toda igualdad $h_i(x) = 0$ son suprimidas, entonces las últimas inecuaciones anteriores (≥ 0) se convierten en inecuaciones estrictas (> 0).

DEMOSTRACIÓN.

El sistema

$$\left\langle \begin{matrix} f_i(x) \leq 0, i \in M, \forall \epsilon > 0 \\ h_i(x) = 0,\ i \in K \end{matrix} \right\rangle \textit{no tiene solución } x \in S$$

[Pero si tuviera una solución $\overline{x}$, entonces:

$$f_i(\overline{x}) \leq \epsilon, \qquad \forall \epsilon > 0 \ \wedge \forall\, i \in M, \qquad \wedge\ h_i(\overline{x}) = 0 \ \forall\, i \in K.$$

Esto a su vez implica que $f_i(\overline{x}) \leq 0 \ \forall\, i \in M \ \wedge\ h_i(\overline{x}) = 0 \ \ \forall\, i \in K$ (por otro lado $f_i(\overline{x}) > 0$ *para algún* $i \in M,$ entonces haciendo $\epsilon = \frac{1}{2} f_i(\overline{x}) > 0$ tenemos una contradicción) Sin embargo, esto contradice las hipótesis del teorema]. Los conjuntos

$$S(i,j,\epsilon) = \left\{x /\ x \in S, f_i(x) \leq \epsilon, h_j(x) = 0\right\}$$

son conjuntos cerrados (a causa de la semi continuidad inferior de $f_i,$ la linealidad de h_j, y la compacidad de S) contenida en el conjunto compacto S y su intersección es vacío. Por consiguiente por el teorema de la intersección finita, existe un número finito de tales conjuntos de modo que su intersección es vacía. Así obtenemos índices $(i_1, i_2, \cdots, i_m) \in M,\ (i_1, i_2, \cdots, i_k) \in M,$ $y\ números\ reales\ \epsilon_1, \epsilon_2, \cdots, \epsilon_m > 0$ tal que el sistema

$$\left\langle \begin{matrix} f_{i_j}(x) - \epsilon_j \leq 0,\ j = 1, 2, 3, \cdots, m \\ h_{i_j}(x) = 0,\ \ j = 1, 2, 3,\ \cdots, k \end{matrix} \right\rangle\ no\ tiene\ solución\ x \in S$$

Por consiguiente por corolario 3.2.2 existe $p \in R^m, q \in R^k$ tal que $p \geq 0\ \ (p, q) \neq 0$, y $\sum_{j=1}^{m} p_j f_{i_j}(x) + \sum_{j=1}^{k} q_j h_{i_j}(x) \geq \sum_{j=1}^{m} p_j \epsilon_j\ \ \forall x \in S$

De aquí seguimos la conclusión del teorema, si observamos que $\sum_{j=1}^{m} p_j \epsilon_j \geq 0,$ y si $K = \phi$, entonces $p \geq 0\ y\ \sum_{j=1}^{m} p_j \epsilon_j > 0$ ■

Establecemos relaciones entre las funciones cóncavas, convexas, estrictamente cóncavas y estrictamente convexas.

a) f es cóncava si sólo si $Hf(X)$ es SEMI DEFINIDA NEGATIVA $\forall X\ \in S$
b) Si $Hf/X)\ \forall X \in S$ es DEFINIFA NEGATIVA, entonces f es estrictamente cóncava
c) Si $Hf(X)$ es SEMI DEFINIDA NEGATIVA $\forall X\ \in S$, entonces f es estrictamente cóncava
d) f es cinvexa si sólo si $Hf(X)$ es SEMI DEFINIDA POSITIVA $\forall X\ \in S$
e) Si $Hf/X)\ \forall X \in S$ es DEFINIFA POSITIVA, entonces f es estrictamente convexa
f) Si $H\ f(X)\ \forall X\ \in S$ es SEMIDEFINIDA POSITIVA entonces f es estrictamente convexa

5. MODELOS DE OPTIMIZACIÓN CONVEXA.

En esta sección se presenta modelos o programas de optimización, extraídos de

www.ub.edu/matheopt/optimizacion-economica que corresponde a los profesores de optimización económica I de la UNIVERSIDAD DE BARCELONA.

.Iniciamos considerando los siguientes problemas.

a) PROGRAMA CONVEXO.

Sea f una función convexa y S un conjunto convexo, se desea resolver el siguiente programa

$$\begin{cases} mín\, f(X) \\ X \in S \end{cases}$$

b) PROGRAMA CÓNCAVO.

Sea f una función cóncava y S un conjunto convexo, se desea resolver el siguiente programa

$$\begin{cases} max\, f(X) \\ X \in S \end{cases}$$

Óptimo y extremo se usan para indicar Mínimo o máximo

Definición (máximo global o absoluto). X_0 es un máximo global o absoluto de f en S sy Sólo si $f(X) \leq F(X_0)\ \ \forall X \in S$

X_0 es un máximo global o absoluto de f en S sy Sólo si $f(X) \leq F(X_0)\ \ \forall X \in S$

X_0 es un máximo local o relativo de f en S sy Sólo si $\exists \delta > 0$ tal que$f(X) \leq F(X_0)\ \ \forall X \in S$

Definición (Mínimo global o absoluto). X_0 es un máximo global o absoluto de f en S si Sólo si $f(X) \geq F(X_0)\ \ \forall X \in S$

X_0 es un mínimo local o relativo de f en S sy Sólo si $\exists \delta > 0$ tal que$f(X) \geq F(X_0)\ \ \forall X \in S$

Si las desigualdades se cumplen en forma estricta, se dice que los óptimos son estrictos

Todo máximo (mínimo) global es local

Todo máximo (mínimo) global estricto es único.

Definición (Punto de inflexión) X_0 es un punto de inflexión de f, si en este punto la función cambia de convexa a cóncava o viceversa.

Si X_0 es un punto de inflexión de f entonces $f''^{(X_0)} = 0$

Definición (Punto de silla) Es aquel punto o puntos que anulan el gradiente de la función objetivo f pero que no son ni máximos ni mínimos.

5.1. TEOREMA FUNDAMENTAL DE LA PROGRAMACIÓN CONVEXA.

A. Sea S un conjunto convexo no vacío y $f: S \subset \mathbb{R}^n \to \mathbb{R}$ una función convexa

Entonces

a) Si $X_0 \in S$ es mínimo local de f en S, también es mínimo global

b) El conjunto de mínimos locales de f en S es un conjunto convexo

B. Sea S un conjunto convexo no vacío y $f: S \subset \mathbb{R}^n \to \mathbb{R}$ una función cóncava

Entonces

c) Si $X_0 \in S$ es máximo local de f en S, también es máximo global

d) El conjunto de máximos locales de f en S es un conjunto convexo

Para resolver un programa convexo (cóncavo) se calcula el punto donde la derivada de la función se anula, es decir se trata de resolver la ecuación algebraica $\nabla \mathrm{f}(X_0) = 0$

6. OPTIMIZACIÓN SIN RESTRICCIONES

6.1. Condiciones necesarias y suficientes.

1. Formulación del programa

Sea $f: \mathbb{R}^n \to \mathbb{R}$ una función convexa. El problema se formula como.

$$opt\ f(X)\ \ X \in \mathbb{R}^n$$

2. Solución del programa.

Resolver este problema consiste en determinar el punto o los puntos de $\mathbb{R}^n$, de tal manera que la función alcance su valor máximo o mínimo.

CONDICIÓN NECESARIA. (C.N) El punto o los puntos donde la función alcanza su máximo o mínimo son aquellos que anulan el gradiente de f, esto es se cumple

$\nabla f(X_0) = 0$ (C.N)

Los puntos X_0 se llaman puntos críticos o estacionarios.

CONDICIÓN SUFICIENTE (C.S)

- Si $Hf(X_0)$ es D.P entonces X_0 es MÍNIMO LOCAL ESTRICTO
- Si $Hf(X_0)$ es DN entonces X_0 es MÁXIMO LOCAL ESTRICTO
- Si $Hf(X_0)$ es INDEFINIDA entonces X_0 es PUNTO DE SILLA

Esta condición se aplica a los puntosX_0 que satisfacen la C.N

Tenemos que.

Si X_0 es un ´0ptimo entonces X_0 satisface la C.N o equivalentemente decimos

Si X_0 NO satisface la C.N entonces X_0 no es óptimo.

E1ejemplo.

$$opt\ (x^2 + y^2 - 2x + 4y + 7)$$

Solución

Por la C.N. $\nabla f(X_0) = 0 \iff [2x - 2 \quad 2y + 4] = [0 \quad 0]$

Resolviendo $x = 1,\ y = -2$

El punto crítico $(1, -2)$ satisface la condición necesaria

Ahora calculamos la matriz Hessiana y lp evaluamos en el punto crítico

$$Hf(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Entonces

$$Hf(1,-2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Corresponde a una forma cuadrática $q(x.y) = 2x^2 + 2y^2$

´Por el criterio de Jacobi

$\Delta_1 = 2 > 0 \qquad \Delta_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 4 > 0$

La forma cuadrática es DEFINIDA POSOTIVA, entonces por la C.S , El punto crítico es UN Mínimo Local estricto

OBSERVACIÚN. Para X_0 que satisface la C.N $(\nabla f(X_0) = 0)$ tenemos:

- ✓ Si $Hf(X)$ es SDP $\forall X \in \mathbb{R}^n$ entonces X_0esa mínimo local y global
- ✓ Si $Hf(X)$ es SDN $\forall X \in \mathbb{R}^n$ entonces X_0 esa máximo local y global
- ✓ Si $Hf(X)$ es SDP $\forall X$ suficientemente cercano a X_0 entonces X_0 esa mínimo local
- ✓ Si $Hf(X)$ es SDN $\forall X$ suficientemente cercano a X_0 entonces X_0 esa máximo local

En seguida planteamos y resolvemos un programa de optimización para una función de una variable real (Caso $n = 1$)

Sea $f: \mathbb{R} \to \mathbb{R}$ una función continua y derivable en $\mathbb{R}$. $X_0 \in \mathbb{R}$

Para resolver este programa usamos

CONDICIÓN NECESARIA (C.N)

X_0 es óptimo local entonces $f'(X_0) = 0$

CONDICIÓN SUFICIENTE (C.S)

Sea X_0un punto que satisface la condición necesaria, entonces:

Si $f''(X_0) > 0$ entonces X_0 es mínimo local

Si $f''(X_0) < 0$ entonces X_0 es máximo local

Ejemplo

Opt$(3x^2 + 6x + 2)$

Solución

C.N

$f'(X_0) = 0 \;\Leftrightarrow 6x + 6 = 0 \;\Leftrightarrow\; X_0 = -1$

C.S.

$$f''(x) = 6$$

$$f''(-1) = 6 > 0$$

Por lo tanto $X_0 = -1$ es un mínimo local

6.2. Optimización con restricciones de igualdad

En muchas aplicaciones prácticas se presenta el problema de optimizar funciones sujeta a condiciones o restricciones en las variables involucradas

mSea $f: \mathbb{R}^n \to \mathbb{R}$ una función diferenciable

$g_i: \mathbb{R}^n \to \mathbb{R}$; $\quad i = 1, 2, 3, \cdots, m$ Funciones diferenciables

$$b_1, b_2, b_3, \cdots, b_m \in \mathbb{R}$$

El programa de optimización con restricciones de igualdad se formula de la siguiente manera:

$$Opt\ f(X),\ \ X \in \mathbb{R}^n$$

$$\begin{matrix} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \vdots \\ g_m(X) = b_m \end{matrix}\ ; \quad m < n \quad X \in \mathbb{R}^m$$

En este problema $f(X)$ es la función objetivo y $g_j(X)\ j = 1, 2, \cdots, m$ son las restricciones $b_1, b_2, b_3, \cdots, b_m \in \mathbb{R}$ son constantes fijas.

OBSERVACIÓN.

1. Las restricciones del tipo de igualdad n restablece fronteras al conjunto de las soluciones factibles del problema sino que reduce las espacio donde el programa está definido
2. Los óptimos que se obtienen serán débiles en el sentido que una pequeña variación en las restricciones hará que dejan de ser óptimos condicionados o restringidos.

MÉTODOS DE SOLUCIÓN.

Para resolver un programa de optimización con restricciones de igualdad se usan los siguientes métodos.

6.2.1. **Gráficamente usando curvas de nivel**. Este método consiste en determinar el punto de la restricción por el que pasa la curva de nivel más baja.

6.2.2. **Eliminación o sustitución de variables.** El método consiste en despejare una variable de la restricción y remplazarlo en la función objetivo transformando el problema en un programa de optimización libre con la función objetivo con una variable menos.

Ejemplo.

Opt $\boldsymbol{f(x,y,)} = \boldsymbol{Opt}\ (\mathbf{12}\boldsymbol{xy} - \mathbf{3}\boldsymbol{y}^{\mathbf{2}} - \boldsymbol{x}^{\mathbf{2}})$

Sujeto a:

$$\boldsymbol{x} + \boldsymbol{y} = \mathbf{16}$$

Solución.

Despejamos la variable x de la restricción $x + y = 16$, tenemos

$$x = 16 - y$$

Remplazando en la función objetivo

$$f(x,y,) = 12xy - 3y^2 - x^2$$

Se obtiene una mueva función de una variable

$$g(y) = 12(16-y)y - 4y^2 - (16-y)^2$$

$$= 192y - 12y^2 - 3y^2 - (256 - 32y + y^2)$$

$$g(y) = -256 + 224y - 16y^2$$

Así tenemos un programa de óptimo libre, para resolverlo seguimos el método correspondiente

C.N $\quad g'(x) = 0 \iff 224 - 32y = 0$

$y = 7$, entonces $x = 16 - 7 = 9$

CS.

$$g''(7) = -32 < 0$$

Por tanto existe un máximo local en $X_0 = (9, 7)$

6.2.3. **Método de Lagrange**. El método consiste en convertir el problema de optimización con restricciones de igualdad en uno de óptimo libre, para esto se introduce nuevas variables adicionales $\lambda_1, \lambda_2, \cdots, \lambda_m$ llamados multiplicadores de Lagrange formar la Lagrangiana la que está formada por la función objetivo más la suma de la multiplicación de cada restricción por cada multiplicador.

Según lo anunciado anteriormente, el problema de optimización restringido Opt$f(x_1, x_2, \cdots, x_n)$

Sujeto a las restricciones

$$
\begin{aligned}
(g_1 x_1, x_2, \cdots, x_n) &= b_1 \\
g_1(x_1, x_2, \cdots, x_n) &= b_2 \qquad\qquad m < n \\
&\vdots \\
g_m(x_1, x_2, \cdots, x_n) &= b_m
\end{aligned}
$$

Se sustituye por el problema de óptimo libre

Opt $L(X,\ \lambda)$

Donde

$X = (x_1, x_2, \cdots, x_n)$ y $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)$

$$
\begin{aligned}
L(X, \lambda) &= f(X) + \lambda_i(b_i - g_i) \\
&= f(x_1, x_2, \cdots, x_n) + \lambda_1\big(b_1 - (g_1 x_1, x_2, \cdots, x_n)\big) + \cdots \\
&\quad + \lambda_m\big(b_m - (g_1 x_1, x_2, \cdots, x_n)\big)
\end{aligned}
$$

Para la resolución del problema tenemos:

CONDICIÓN NECESARIA (C.N)

$$\nabla L(X, \lambda) = 0 \quad \Leftrightarrow \quad \left[\frac{\partial L}{\partial x_1} \quad \frac{\partial L}{\partial x_2} \quad \cdots \quad \frac{\partial L}{\partial x_n}; \quad \frac{\partial L}{\partial \lambda}\right] = 0$$

CONDICIÓN SUFICIENTE (C.S)

Sea un punto $(X_0; \lambda_0)$ que cumple las condiciones necesarias

Denotando la HESSIANA de la función de Lagrange en el punto $(X_0; \lambda_0)$ respecto a las variables originales y no respecto a λ, como:

$$H_X L(X_0; \lambda_0)$$

Entonces

a) Si $H_X L(X_0; \lambda_0)$ es D.P $\Longrightarrow X_0$ es un mínimo condicionado o restringido

b) Si $H_X L(X_0; \lambda_0)$ es D.N $\Longrightarrow X_0$ es un máximo condicionado o restringido

En otro caso

$\forall\, \Delta\, X_0$ tal que $\nabla_g^T(X_0) \cdot \Delta\, X_0 = 0$

a) Si $\Delta X_0^T H_X L(X_0; \lambda_0) \Delta X_0 > 0 \implies X_0$ es un mínimo condicionado

b) Si $\Delta X_0^T H_X L(X_0; \lambda_0) \Delta X_0 < 0 \implies X_0$ es un máximo condicionado

Ejemplo.

$$Opt\ (x^2 + y^2 + z^2)$$

Sujeto a la restricción

$$3x - 2y + z - 4 = 0$$

Formación de la función de Lagrange

$$L(x, y, z;\ \lambda) = x^2 + y^2 + z^2 + \lambda(3x - 2y + z - 4)$$

C.N.

$$\nabla L(x, y, z, \lambda) = 0$$

$$\frac{\partial L}{\partial x} = 2x + 3\lambda = 0$$

$$\frac{\partial L}{\partial y} = 2y - 2\lambda = 0$$

$$\frac{\partial L}{\partial z} = 2z + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = 3x - 2y + z - 4 = 0$$

Debemos resolver el sistema

$$\begin{aligned} &2x + 3\lambda = 0 && (1)\\ &2y - 2\lambda = 0 && (2)\\ &2z + \lambda = 0 && (3)\\ &3x - 2y + z - 4 = 0 && (4) \end{aligned}$$

Eliminación del multiplicador

De las ecuaciones (1) y (2)

$$\frac{2x}{2y} = -\frac{3}{2} \quad \Longrightarrow y = -\frac{2x}{3} \qquad (5)$$

De las ecuaciones (2) y (3)

$$\frac{y}{z} = -2 \quad \Longrightarrow \quad z = -\frac{y}{2} \qquad (6)$$

Remplazando (5) en (6)

$$z = \frac{x}{3} \qquad (7)$$

Remplazando (5) y (7) en (4)

$$3x + \frac{4x}{3} + \frac{x}{3} = 4 \quad \Longrightarrow x = \frac{6}{7}$$

$$\therefore\ x = \frac{6}{7}, y = -\frac{4}{7}, z = \frac{2}{7}$$

Así tenemos el punto crítico

$$X_0 = \left(\frac{6}{7}, -\frac{4}{7}, \frac{2}{7}\right)$$

CS.

Escribimos la matriz Hessiana

$$H_X L(X_0;\lambda_0) = \begin{bmatrix} L_{xx} & L_{xy} & L_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

La matriz Hessiana evaluada en el punto crítico es

$$H_X L(X_0;\lambda_0) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Según el método de Jacobi la matriz Hessiana es DEFINIDA POSITIVA puesto que:

$$\Delta_1 = 2 > 0; \Delta_2 = 4 >; \Delta_3 = 8 > 0$$

En conclusión, como la Matriz Hessiana es D.P entonces el punto crítico corresponde a un mínimo condicionado.

7. OPTIMIZACIÓN CON RESTRICCIONES DE DESIGUALDAD

1. Formulación del problema

El método de los multiplicadores de Lagrange puede ser modificado cambiando las restricciones de igualdad por restricciones de desigualdad

Sea $f: \mathbb{R}^n \to \mathbb{R}$ una función diferenciable

$g_i: \mathbb{R}^n \to \mathbb{R}; \quad i = 1, 2, 3, \cdots, m$ funciones diferenciables

El programa de optimización con restricciones de desigualdad se formula de la siguiente manera:

$$Opt\ f(X), \ \ X \in \mathbb{R}^n$$

$$\begin{matrix} g_1(X) \leq 0 \\ g_2(X) \leq 0 \\ \vdots \\ g_m(X) \leq 0 \end{matrix} ; \quad X \in \mathbb{R}^m$$

Las restricciones de desigualdad definen fronteras al dominio de soluciones factibles.

2. Resolución del problema.

El problema de optimización con restricciones de desigualad se resuelve siguiendo los siguientes conceptos y teoremas

7.1. Aplicación convexidad. Teorema local global.

7.2. Aplicación curvas de nivel. Sirven para resolver el problema en forma geométrica.

7.3. Aplicación Teorema de WEIERSTRASS. Toda función continua sobre un dominio compacto alcanza máximo y mínimo globales.

7.4. Aplicación condiciones de Kuhn-TUCKER. Para el problema de optimización con restricciones de desigualdad planteado antes se tiene las condiciones siguientes:

7.5. CONDICIONES NECESARIAS DE PRIMER ORDEN DE OPTIMALIDAD LOCAL.

Las condiciones de K-T son necesarias de optimalidad local, es decir:

a) Si X_0 es máximo local $\Longrightarrow$ X_0 satisface K-T para máximo

b) Si X_0 es mínimo local $\Longrightarrow$ X_0 satisface K-T para mínimo

Equivalentemente podemos decir:

a) Si X_0 no satisface K-T para máximo $\Longrightarrow$ X_0 NO es máximo local

b) Si X_0 no satisface K-T para mínimo $\Longrightarrow$ X_0 NO es mínimo local

En resumen podemos decir: Si un punto satisface las condiciones de K-T para máximo, solo podemos decir que es un posible máximo local, en cambio si no lo satisface podemos asegurar que NO es máximo local

7.6. ENUNCIADO DE LAS CONDICIONES

Las condiciones de K-T son las siguientes:

1. $\nabla f(X_0) = \sum_{i=1}^{m} \lambda_i \nabla g_i(X_0)$

   Donde los λ_i se conocen como multiplicadores de K-T

   Esta condición dice que el gradiente de ls función objetivo en X_0es una combinación lineal de los gradientes de las restricciones en X_0

2. $\lambda_i g_i(X_0) = 0 \quad \forall i.$

3. $\lambda_i \geq 0 \quad \forall i$ (Si se trata de máximo)

   Cuando el punto satisface K-T para máximo los multiplicadores deben ser positivos

   $\lambda_i \leq 0 \quad \forall i$ (Si se trata de mínimo)

   Cuando el punto satisface K-T para mínimo los multiplicadores deben ser negativos

4. $g_i(X_0) \leq 0 \quad \forall i$

   Esta condición dice que el punto X_0 debe satisfacer todas las restricciones del problema, es decir que X_0 debe pertenecer al conjunto de soluciones factibles del problema-

   Definición. Se dice que un punto satisface las condiciones de K-T para máximo cuando satisface las condiciones de K-T con $\lambda_i \geq 0 \quad \forall i$

   Definición. Se dice que un punto satisface las condiciones de K-T para mínimo cuando satisface las condiciones de K-T con $\lambda_i \leq 0 \quad \forall i$

7.7. SUFICIENCIA Y CONVEXIDAD.

Cuando hay convexidad, las condiciones de K-T que son necesarias en general también son suficientes, en otros términos

i) Cuando Dominio S CONVEXO Y f convexa. Si X_0 satisface K-T para mínimo ⟺ 0 X_0 es mínimo local

ii) Cuando Dominio S CONVEXO Y f cóncava. X_0 satisface K-T para máximo ⟺ 0 X_0 es máximo local

OBSERVACION.

Si se tiene en cuenta el teorema local global, podemos afirmar que:

a) Cuando Dominio S CONVEXO Y f convexa. Si X_0 satisface K-T para mínimo local⟺ X_0 es mínimo global

b) Cuando Dominio S CONVEXO Y f cóncava Si X_0 satisface K-T para máximo local⟺ X_0 es máximo global

Para resolver un problema de optimización con restricciones de desigualdad es necesario primero clasificarlo y luego aplicar el método. Para esto es necesario analizar:

a. La convexidad de la función objetivo.
b. La compacidad del dominio.
c. continuidad de la función objetivo.
d. La convexidad del dominio.

Ejemplo.

Analizar el siguiente problema

$$Opt(x-3)^2+(y-2)^3$$
$$x+y\leq 7$$
$$x\geq 0$$
$$y\geq 0$$

Solución.

- Veamos la convexidad y la continuidad de la función objetivo

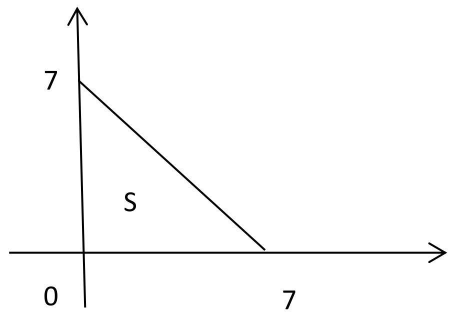
La matriz Hessiana de la función objetivo es

$$Hf(x,y)=\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}=\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Según el método de Jacobi esta matriz corresponde a una forma cuadrática definida positiva, entonces la función objetivo es estrictamente convexa y por lo tanto convexa
La función objetivo es continua.

- Veamos la convexidad del dominio.
  El dominio S está formado por las desigualdades
$$x + y \leq 7$$
$$x \geq 0$$
$$y \geq 0$$

  Una gráfica del dominio en el plano cartesiano es



- Veamos si el dominio es compacto. Tenemos que el dominio es cerrado y acotado, por tanto, es compacto
  Según todo lo analizado antes, el problema se puede resolver por el teorema de Weirstrass porque se cumplen todos los requisitos del teorema y deducir que el problema tiene mínimo global y máximo global; también se puede usar el teorema local global puesto que la función es convexa y el dominio también es convexo y este teorema dirá que el mínimo local será global y único

2.3. DEFINICIÓN DE TÉRMINOS BÁSICOS

Los siguientes términos matemáticos son tomados de Diccionario Ilustrado de conceptos matemáticos, de Soto Apolinar. E 2011

1. Axioma Una verdad tan evidente que no requiere demostrarse. Por ejemplo, «la suma de dos números reales es otro número real», es un axioma.
2. Cálculo Rama de las matemáticas que se encarga del estudio de las cantidades que varían continuamente y las relaciones entre ellas.
3. Cerradura Un conjunto A presenta la propiedad de cerradura bajo una operación cuando al realizar esa operación
4. Cerradura Un conjunto A presenta la propiedad de cerradura bajo una operación cuando al realizar esa operación
5. Conclusión Es el resultado de una implicación lógica.
6. Condición necesaria En la implicación: $p \rightarrow q$, q es la condición necesaria. Por ejemplo, una condición necesaria para que un cuadrilátero sea cuadrado es que todos sus ángulos midan lo mismo. Sin embargo, esta condición no es suficiente.
7. Condición suficiente Condición que requiere cumplir un objeto matemático para satisfacer una implicación en ambos sentidos. $p \leftrightarrow q$
8. Conjunto cerrado Conjunto que contiene todos sus puntos frontera
9. Crítico, punto En una curva, el punto crítico es el punto donde una recta tangente a la curva es horizontal.
10. Demostración Justificación de una afirmación, premisa o sentencia de una manera estructurada, lógica e irrefutable a partir de otras sentencias verdaderas
11. Dependiente, variable Una variable es dependiente si su valor depende del valor de otra u otras variables.
12. Dependencia funcional Se dice que la variable y depende funcionalmente de la variable x si es posible escribir la relación que existe entre ellas en forma de ecuación. En ese caso, y es la variable dependiente (depende de x) y x es la variable independiente
13. Desigualdad Una desigualdad es una relación matemática que compara el valor de dos números o expresiones algebraicas (del tipo mayor o menor)
14. Dominio El dominio D de una función es el conjunto formado por todos los valores que la función puede aceptar para devolver un único valor por cada uno de ellos.
15. Máximo Valor más grande que toma o puede tomar una variable.
16. Mínimo Valor más pequeño que acepta o puede tomar una variable.

17. Optimización Un problema es de optimización cuando se requiere maximizar o minimizar una cantidad.
18. Programa Listado de instrucciones que permite la solución de un problema a través de la computadora. Generalmente los programas se escriben en algún lenguaje de programación para que la computadora pueda entender las instrucciones.
19. Punto crítico En una curva, el punto crítico es el punto donde una recta tangente a la curva es horizontal.
20. Satisfacer Decimos que un valor satisface a una ecuación o a una función cuando al sustituir este valor en la ecuación o función ésta se reduce a una igualdad válida

2.4. Formulación de la Hipótesis

2.4.1. Hipótesis general

Si es necesario y suficiente la convexidad en la optimización matemática.

2.4.2. Hipótesis específicas

HE1. Si la convexidad es aplicable para un sistema de optimización libre.

HE2. Si la convexidad es necesaria para un programa con restricciones de igualdad.

HE3. Si la convexidad es condición necesaria para optimizar un programa con restricciones de desigualdad.

CAPITULO III

3. METODOLOGÍA GENERAL
   3.1. DISEÑO METODOLÓGICO.

      3.3.1. TIPO. Teórico –documental y descriptivo

      3.3.2. ENFOQUE. El desarrollo de la presente tesis es analítico cualitativo. No se usa estadística.

   3.2. POBLACIÓN Y MUESTRA

   De acuerdo al tipo y enfoque de la tesis esta carece de población y muestra.

   3.3. OPERACIONALIZACIÓN DE VARIABLES E INDICADORES.

   Puesto que la tesis es teórica y analítica y en ella no se busca relación entre variables no se consideran variables dependiente e independiente.

   3.4. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS
      3.4.1. Técnicas. Recolección de información bibliográfica.
      3.4.2. Descripción de los instrumentos.
      Recolección y uso de temas necesarios para elaborar la tesis.
   3.5. TÉCNICAS PARA EL PROCESAMIENTO DE LA INFORMACIÓN.

   Análisis, procedimientos algebraicos y técnicas de optimización.

CAPITULO IV

4. RESULTADOS GENERALES.

Entre los resultados fundamentales se presenta y explica la optimización libre y restringida
haciendo uso de las funciones cóncavas y convexas, así como los conjuntos convexos y las formas cuadráticas.

CAPITULO V

5. DISCUSIÓN, CONCLUSIONES Y RECOMENDACIONES

5.1. DISCUSÍONES

Entre las discusiones se han discutido teoremas relativos a funciones convexas y conjuntos convexos, así como teoremas para clasificar las formas cuadráticas a fin de utilizarlo en la teoría de optimización de convexidad´ libre y restringido.

5.2. CONCLUSIONES

Una de las conclusiones de la investigación es que la convexidad de las funciones, los conjuntos convexos, la matriz Hessiana y las formas cuadráticas son herramientas fundamentales para el estudio de optimización y para una generalización al cálculo de variaciones

5.3. RECOMENDACIONES

Se recomienda estudiar la optimización dentro del cálculo de variaciones y la programación dinámica y otros temas más avanzados.

CAPITULO VI

6. FUENTES DE INFORMACIÓN.

6.1. FUENTES DOCUMENTALES

No fueron necesarios

6.2. FUENTE HEMEROGRÁFICAS

No fueron utilizados

6.3. FUENTES BIBLIOGRÁFICAS

1. Arrow, K. – Hurwicz, and A. (1961) *QUASICONCAVE PROGRAMMING ECONOMÉTRICA. Ed, Stanford University Press. California.*

2. Arrow, K. Hirwicz. AND h. USAWA (1958). *Estudies in Linearand Nonlinear Proghramming Ed, Stanford University Press. California.*

3. Abadie, J. (1967) *ONTHE KUHN-TUCKER TEOREM. NONLINEAR PROGRAMMING. Ed. North Holland Publishing Company Amsterdam*

4. Danzing.G. *LINEAR PROGRAMMING AND EXTENSIONS. Ed. Princeton University Press, Princeton N.J.*

5. Hadley, G (1962) *LINEAR PROGRAMMING. . Ed. Adisson Wesley Publisshing Company, inc , Reading g Mass.*

6. Hadley, G. (1964) *NONLINEAR AND DINAMIC PROGRAMMING. Ed. Adisson Wesley Publisshing Company, inc , Reading g Mass*

7. Masanao AOKI (1971) *INTRODUCTION TO OPTIMIZATION TECNIQUES: Fundamentals and applicatiobs of Non Linear Programming. Ed. Macmillan Company.* Nev York

8. Alphaa C. (1992) *ELEMENTS OF DYNAMIC OPTIMIZATION.Ed. Mc Graw- Hill. Nev York*

9. Alpha C. (1984) *FUNDAMENTAL METODS OF MATHEMATIVAL E CONOMIC. Ed. Mc Graw- Hill. Nev York*

10. Draper. J. (1976) *MATEMÁTICAS PARA ADMINISTRACIÓN Y ECONOMÍA. Ed. Harla de C.V. México*.

11. Gass.S. (1973) *PROGRAMACIÓN Lineal. Ed. Continental. México*

12. Phlppi B.(1982) *INTRODUCCIÓN A LA OPTIMIZACIÓMN DE SISTYEMAS*. Ed. Universidad Católica de Chile. Escuela de Ingeniería. Chile.

13. Bonifaz J. (1999) *OPTIMIZACIÓN DINÁMICA Y TEORÍA ECONÓMICA Ed.* Universidad del Pacifico. Perú

14. Bonifaz J. (2003) *MATEMÁTICAS PARA LA ECONOMÍA DINÁMICA la ed.* Corregida Lima CIUP

15. Mangasarian.O. (1969) *NON LINEAR PROGRAMMING Ed.Tata MCc Graw Publishing Company Limited.* Nev York.

16. Leoneli, H.(2001*). MÉTODOS DINÁMICOS EN ECONOMÍA. Otra búsqueda del tiempo perdidoEd. Istituto Tecnológico de México.*

6.4. FUENTES ELECTRÓNICAS.

1 . Barrendero J. Tema 4. Funciones convexas y optimización convexa. Disponible en

verso.mat.uam.es/~joser.berrendero/cursos/Matematicas-IO/io-tema4-16.pdf

2. Funciones convexas. 3. Problemas de optimización convexa. 4. Dualidad. 5. Aplicaciones y algoritmos. Optimización Convexa en Ingeniería, Disponible en:

portal.uned.es/EadmonGuiasWeb/htdocs/abrir_fichero/abrir_fichero.jsp? idGuia...

3. Luna J. Optimización Convexa. Juan Pablo Luna. Programa de Engenharia de Produç˜ao. COPPE – UFRJ – Brasil. Disponible en

www.opmat.org/emalcabolivia2015/docs/Resumen_convexa.pdf jpluna@po.coppe.ufrj.br. Resumen.

4. Optimización económica. Conocimientos Iniciales · Forma Cuadrática · Optimización sin restricciones · Optimización con restricciones de igualdad. Disponible en:

www.ub.edu/matheopt/optimizacion-economica

ANEXOS

MATRIZ DE CONSISTENCIA

PROBLEMAS	OBJETIVOS	HIPÓTESIS	METODOLOGÍA
1. Problema general. ¿Es importante la convexidad en la optimización matemática? 2. Problemas específicos. **PE1**. ¿Cómo resolver un programa de optimización libre? **PE2**. ¿Cómo resolver un programa de optimización con restricciones de igualdad? **PE3**. ¿Cómo resolver un programa de optimización con restricciones de desigualdad?	1. Objetivo general Resolver un programa de optimización convexa. 2. objetivos específicos. **OE1**. Resolver un programa de optimización libre. **OE2**. Resolver un programa de optimización con restricciones de igualdad. **OE3**. Resolver un programa de optimización con restricciones de desigualdad.	1. Hipótesis general Si es necesario y suficiente la convexidad en la optimización matemática. 2. Hipótesis específicas. **HE1**. Si la convexidad es aplicable para un sistema de optimización libre. **HE2**. Si la convexidad es necesaria para un programa con restricciones de igualdad. **HE3**. Si la convexidad es condición necesaria para optimizar un programa con restricciones de desigualdad.	1. Tipo de investigación. Documental –Descriptivo 2. Enfoque. El desarrollo de la presente tesis es analítico cualitativo. No se usa estadística 3. Población y muestra. De acuerdo al tipo y enfoque de la tesis esta carece de población y muestra. 4. Técnicas empleadas. Recolección y uso de temas necesarios para elaborar la tesis.